

ЛЕНИНГРАДСКИЙ ОРДЕНА ЛЕНИНА
И ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ им. А. А. ЖДАНОВА

Л. М. АБРАМОВ, В. Ф. КАПУСТИН

МАТЕМАТИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

Допущено Министерством высшего и среднего
специального образования СССР
в качестве учебного пособия для студентов вузов,
обучающихся по специальности «Экономическая кибернетика»



ИЗДАТЕЛЬСТВО ЛЕНИНГРАДСКОГО УНИВЕРСИТЕТА
ЛЕНИНГРАД 1981

*Печатается по постановлению
Редакционно-издательского совета
Ленинградского университета*

УДК 519.95

Абрамов Л. М., Капустин В. Ф. Математическое программирование: Учеб. пособие. Л.: Изд-во Ленингр. ун-та, 1981. 328 с. Ил. — 39, библиогр. — 52 назв.

Книга написана на основе курса «Математические методы исследования операций», который читается на отделении экономической кибернетики экономического факультета Ленинградского университета. В ней содержатся теоретические основы выпуклого и дискретного программирования и алгоритмы решения соответствующих задач. Рассматриваются некоторые экономические ситуации, которые формализуются как задачи математического программирования. Строгость изложения сочетается с возможностью освоения вычислительных алгоритмов без предварительного разбора их обоснования и изучения теоретического материала.

Книга рассчитана на студентов экономических вузов и факультетов, а также на специалистов экономических служб предприятий и ведомств.

Рецензенты: кафедра исследования операций ф-та вычислительной математики и кибернетики Московского университета (зав. кафедрой проф. П. С. Краснощеков), проф. В. А. Залгаллер (Ленинградское отделение математического института им. В. А. Стеклова)

А 20204—009
076(02)—81 65—80.1502000000

© Издательство
Ленинградского университета, 1981 г.

ИБ № 1038

*Абрамов Леонид Михайлович
Капустин Виталий Федорович*

Математическое программирование

Редакторы *А. А. Гранаткина, Г. И. Чередниченко*

Художественный редактор *В. Н. Васильев*

Технический редактор *А. В. Борщева*

Корректор *Л. И. Чулкова*

Сдано в набор 29.02.80. Подписано в печать 11.12.80. М 33560. Гарнитура литературная.
Печать высокая. Печ. л. 20,5. Формат 60×90/16. Бум. тип. № 1. Уч.-изд. л. 19,61.
Тираж 10000 экз. Заказ 631. Цена 95 коп.
Издательство ЛГУ им. А. А. Жданова, 199164,
Ленинград, В-164, Университетская наб., 7/9.

Ленинградская типография № 2 головное предприятие ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» им. Евгения Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательства, полиграфии и книжной торговли. 198052, г. Ленинград, Л-52, Измайловский проспект, 29.

ОГЛАВЛЕНИЕ

Предисловие	5
Глава 1. Теоретические основы выпуклого программирования	7
§ 1. Некоторые обозначения, понятия и термины линейной алгебры	—
§ 2. Некоторые обозначения, понятия и термины математического анализа	12
§ 3. Выпуклые множества в R^n	17
§ 4. Теорема отделимости	26
§ 5. Многогранные выпуклые множества и выпуклые многогранники	31
§ 6. Конусы. Неограниченные выпуклые множества	37
§ 7. Многогранные выпуклые конусы. Неограниченные многогранные выпуклые множества	43
§ 8. Выпуклые функции	52
Упражнения	64
Глава 2. Экстремальные задачи	66
§ 1. Определение экстремальной задачи	—
§ 2. Некоторые классы экстремальных задач	68
§ 3. Экономико-математическое моделирование. Примеры экстремальных задач	71
§ 4. Эквивалентные экстремальные задачи	83
Упражнения	91
Глава 3. Критерии оптимальности	92
§ 1. Критерии оптимальности при условиях дифференцируемости	93
§ 2. Теорема Куна — Таккера	111
§ 3. Двойственность в математическом программировании	117
Упражнения	134
Глава 4. Основные вычислительные методы линейного программирования	136
§ 1. Предварительные замечания	—
§ 2. Методы допустимых направлений для решения задач линейного программирования	138
§ 3. Симплекс-метод	148
§ 4. Симплекс-метод — вычислительная схема, связанная с преобразованием обратных матриц	162
§ 5. Двойственный симплекс-метод	168
§ 6. Градиентный метод	180

§ 7. Вырожденность. Решение вырожденных задач	189
Упражнения	192
Глава 5. Применение процедуры симплекс-метода для решения некоторых экстремальных задач	193
§ 1. Метод решения задачи линейного программирования с ограниченными сверху переменными	—
§ 2. О задачах квадратичного программирования и одном методе их решения	202
§ 3. Метод Трубина решения задачи целочисленного линейного программирования специального вида	208
Упражнения	211
Глава 6. Специальные классы задач линейного программирования и методы их решения	212
§ 1. Специальные классы задач линейного программирования. Проблемы генерирования их условий и вычислительной информации. Симуляция допустимых планов	—
§ 2. Блочная задача. Метод декомпозиции (разложения) Данигиса — Вулфа	217
§ 3. Графы. Транспортные сети и транспортные задачи	229
§ 4. Метод потенциалов	239
§ 5. Транспортная задача в матричной постановке. Венгерский метод	247
Упражнения	253
Глава 7. Параметрические задачи	254
§ 1. Параметрические задачи линейного программирования	—
§ 2. Алгоритм решения параметрической задачи линейного программирования с параметром в целевой функции	257
Упражнения	267
Глава 8. О методах решения задач дискретного программирования	268
§ 1. Унимодулярные матрицы и целочисленность решения задач транспортного типа	—
§ 2. Методы ветвей и границ	272
Упражнения	282
Глава 9. Методы решения нелинейных экстремальных задач выпуклого программирования	283
§ 1. Общие схемы итеративных методов	284
§ 2. Задачи безусловной оптимизации	287
§ 3. Методы штрафных функций	298
§ 4. Методы допустимых направлений	306
§ 5. Итеративный метод решения задачи квадратичного программирования — метод сопряженных направлений	316
Упражнения	—
Указатель литературы	324
Предметный указатель	325

ПРЕДИСЛОВИЕ

Эта книга возникла на основе курса лекций, который авторы в течение ряда лет читали на отделении экономической кибернетики экономического факультета Ленинградского университета. В ней содержатся теория выпуклого программирования, алгоритмы решения линейных, дискретных и нелинейных экстремальных задач. Обсуждаются вопросы, связанные с экономико-математическим моделированием.

Уже издано довольно много книг и статей, которые полезно использовать как учебные пособия или справочный материал по отдельным разделам курса математического программирования. Однако некоторые из них (например, [6, 11, 30, 48]) давно стали библиографической редкостью, другие изданы малым тиражом и практически недоступны студентам; наконец, только в [31, 32] обсуждаются все разделы типовой программы курса математического программирования для специальности «Экономическая кибернетика». Все это вызвало необходимость в создании настоящего учебного пособия.

Содержание книги в основном соответствует учебным программам по указанному курсу. Особую роль играет первая глава, включающая элементы выпуклого анализа. Ее материал используется как для обоснования теоретических положений и методов, так и для геометрических интерпретаций, проясняющих основные идеи выпуклого программирования. Мы старались сделать изложение более доступным, но, тем не менее, предупреждаем читателя, что изучение курса потребует от него некоторых усилий, совершенно естественных при работе с математической литературой. Особо отметим, что предполагается знакомство читателя с определенными разделами курсов линейной алгебры и математического анализа.

Форма изложения не является традиционной. Здесь мы продолжаем следовать идеям, положенным в основу [1]. Отметим наиболее важные моменты.

В книге рассмотрена теория выпуклого программирования (глава 3). Теория линейного программирования, в частности теоремы двойственности для задач линейного программирования, следует из общей теории. Однако при изложении вычислительных методов мы вынуждены были отказаться от такого подхода: алгоритмы решения задач линейного программирования предшествоуют вычислительным методам решения нелинейных задач, поскольку используются в методах выпуклого, квадратичного и дискретного программирования.

Алгоритмы решения задач следует воспринимать не только как вычислительные и логические предписания, но и как конструкции, которые можно использовать для эффективного доказательства некоторых утверждений.

Алгоритмы не доведены до машинной реализации. Они могут служить лишь основой для написания машинных программ. Принципы и методы составления программ не являются предметом изучения настоящего курса.

Предполагается возможным выборочное чтение некоторых глав и отдельных параграфов. Книга написана так, чтобы можно было изолированно изучать теорию и вычислительные методы, знакомиться в первую очередь с вычислительными аспектами и лишь потом — с идеями и точным обоснованием. В какой мере это удалось — судить читателю.

В конце каждой главы приведены упражнения. Некоторые из них сформулированы непосредственно в тексте. Выполнение упражнений мы настоятельно рекомендуем читателю. Для проведения практических занятий можно использовать задачки [14, 17, 24], последний из них согласован с материалом настоящего пособия.

О системе нумерации и ссылок. Пункты параграфа имеют двойную нумерацию: первое число соответствует номеру параграфа в данной главе, второе — номеру пункта в этом параграфе. Ссылки на пункты внутри главы осуществляются естественным образом: например, ссылка «см. п. 3.6» в тексте четвертой главы означает, что мы отсылаем читателя к названному пункту этой главы. При ссылках на пункты других глав перед номером пункта добавляется номер главы. Так, ссылка «см. п. 4.3.6» в тексте шестой главы означает, что мы отсылаем читателя к п. 3.6 гл. 4. Упражнения нумеруются двумя числами: первое указывает номер главы, второе — номер упражнения в этой главе.

Об одной особенности: для j -й компоненты вектора $x \in R^n$ мы ввели обозначение $x[j]$ (в примерах, как правило, используется обычное обозначение x_j). Это позволяет, как нам кажется, упростить изложение и обоснование некоторых алгоритмов.

Конец доказательства, а также окончание алгоритма отмечается знаком ■.

Указатель литературы не претендует на полноту. На большинство работ имеются прямые ссылки в тексте учебного пособия, остальные [4, 5, 7—9, 12, 25, 26, 36—40, 50, 51] также включены в указатель, так как мы считали полезным обратить на них внимание читателя.

Мы благодарим сотрудников кафедры исследования операций Ленинградского университета и участников семинара этой кафедры за полезное обсуждение книги [1], а также С. А. Ашманова, П. С. Краснощекова и В. Ю. Лисицына за ценные советы и рекомендации. Мы глубоко благодарны В. А. Залгаллеру, особенно тщательно изучившему рукопись. Его замечания и советы, несомненно, способствовали улучшению книги.

Глава I

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

Содержание этой главы несколько разнородно. Первые два параграфа включают обозначения, понятия и термины из линейной алгебры и математического анализа. В этих параграфах сформулированы также некоторые утверждения. Поскольку предполагается, что читатель знаком с названными курсами, утверждения приведены здесь без доказательства. В остальных параграфах достаточно подробно и с полными доказательствами рассматривается теория выпуклого анализа в R^n — теоретическая основа настоящего курса математического программирования. Понимание этой теории требует определенной математической культуры. Поэтому читателям, интересующимся, главным образом, вычислительными проблемами математического программирования, следует начать чтение книги со второй главы, используя первую только как необходимый справочный материал.

§ 1. Некоторые обозначения, понятия и термины линейной алгебры

1.1. Через N обозначается множество всех натуральных чисел, через R , как обычно, — множество вещественных чисел или вещественная прямая, через R^n — n -мерное вещественное векторное (или аффинное) пространство, элементами которого являются упорядоченные n -ки вещественных чисел ($n \in N$). Элементы этого пространства будем называть точками, векторами или планами и обозначать строчными буквами латинского алфавита $a, b, \dots, x, y, \dots$. Если $x \in R^n$, то $x = (x[1], x[2], \dots, x[n])$ или $x^T = (x[1], x[2], \dots, x[n])$ (T — знак транспонирования). Иногда там, где это не может вызвать недоразумений, главным образом в примерах, будем писать $x = (x_1, \dots, x_n)$. Через $e_i, i \in 1:n$, всегда будем обозначать элементы канониче-

ского базиса пространства R^n :

$$e_i[j] = \begin{cases} 1, & \text{если } j=i, \\ 0, & \text{если } j \in 1:n \setminus \{i\}. \end{cases}$$

Числа $x[j]$ — координаты вектора x в каноническом базисе — будем называть *компонентами* вектора x .

Если $a \in R^n$, $b \in R^n$, то через ab обозначаем скалярное произведение векторов a и b : $ab = \sum_{i=1}^n a[i] b[i]$; произведение матриц $A = (a[i, j], i \in 1:k, j \in 1:m)$ и $B = (b[i, j], i \in 1:m, j \in 1:n)$ обозначается через AB ; если $k = 1$ или $n = 1$, будем писать aB или Ab . Знак транспонирования «т» в записи векторов и при умножении векторов и матриц ставится лишь тогда, когда его отсутствие может вызвать недоразумения; в записи скалярного произведения знак транспонирования не ставится. Матрица, обратная к квадратной невырожденной матрице A , обозначается через A^{-1} , единичная матрица — через I или I_n , если возникает необходимость указать ее порядок. Строку матрицы A с номером i будем обозначать через a_i , а ее j -й столбец — через a_j^t .

1.2. *Размерность подпространства T пространства R^n обозначается через $\dim T$. Она равна числу векторов базиса пространства T ($\dim \{0\} = 0$ по определению).*

Пересечение любого множества подпространств пространства R^n есть подпространство в R^n . Если T_1 и T_2 — подпространства в R^n , то $T_1 + T_2 = \{x \in R^n: x = t_1 + t_2, t_1 \in T_1, t_2 \in T_2\}$ — *сумма подпространств T_1 и T_2* — подпространство в R^n .

Всякое подпространство T пространства R^n может быть представлено как множество решений некоторой однородной системы линейных уравнений

$$\begin{matrix} a[1, 1]x[1] + \dots + a[1, n]x[n] = 0, \\ \vdots \\ a[m, 1]x[1] + \dots + a[m, n]x[n] = 0, \end{matrix} \quad (I)$$

или в матричной записи $Ax = 0$, где $A = (a[i, j], i \in 1:m, j \in 1:n)$, $x = (\{1\}, \dots, x[n])$. Система (1) называется *уравнением подпространства T* . Верно и обратное: множество решений системы (1) всегда есть подпространство в R^n , и его размерность равна $n - r(A)$, где через $r(A)$ обозначен ранг матрицы A . Заметим, что при построении уравнения подпространства T матрицу A можно выбрать такой, чтобы $m = r(A) = n - \dim T$.

Через $\mathcal{L}(M)$ обозначим линейную оболочку множества $M \subset R^n$ — наименьшее подпространство, содержащее множество M . Линейная оболочка $\mathcal{L}(M)$ состоит из всех (конечных) линейных комбинаций элементов множества M :

$$\mathcal{L}(M) = \{x \in R^n : x = \sum_{i=1}^k a_i a_i, a_i \in M, a_i \in R, k \in N\}.$$

Размерность подпространства $\mathcal{L}(M)$ равна максимальному числу линейно независимых векторов в множестве M .

1.3. Векторы $a, b \in R^n$ называются *ортогональными*, если их скалярное произведение $ab = 0$. Множества M_1 и M_2 в R^n называются *ортогональными*, если каждый вектор из M_1 ортогонален каждому вектору из M_2 . Множество всех векторов, ортогональных непустому множеству $M \subset R^n$, называется *ортогональным дополнением* множества M и обозначается через $M^\perp$. Ортогональное дополнение непустого подмножества пространства R^n есть подпространство в R^n . Если T — подпространство, то $(T^\perp)^\perp = T$. Если T — подпространство в R^n и $T^\perp$ — его ортогональное дополнение, то $T \cap T^\perp = \{0\}$, $T + T^\perp = R^n$.

Сумма попарно ортогональных подпространств $T_1, T_2, \dots, T_k$ называется *ортогональной суммой подпространств*. Для обозначения ортогональной суммы применяется знак $\oplus$: $T_1 \oplus T_2 \oplus \dots \oplus T_k = \sum_{i=1}^k \oplus T_i$. Если $T = \sum_{i=1}^k \oplus T_i$, то $\dim T = \sum_{i=1}^k \dim T_i$.

Представление пространства R^n в виде ортогональной суммы его подпространств T_i , $i \in 1:k$, называется *разложением пространства R^n в ортогональную сумму подпространств*. При этом если $R^n = \sum_{i=1}^k \oplus T_i$, то $\sum_{i=1}^k \dim T_i = n$ и любой вектор $a \in R^n$ однозначно представим в виде $a = \sum_{i=1}^k a_i$, $a_i \in T_i$. Вектор a_i называется *ортогональной проекцией* вектора a на подпространство T_i . В частности, $R^n = T \oplus T^\perp$, $\dim T^\perp = n - \dim T$ и любой вектор $a \in R^n$ однозначно представим в виде $a = b + c$, $b \in T$, $c \in T^\perp$.

1.4. Пусть $a \in R^n$, $b \in R^n$, $a \neq b$. Множество $L = \{x \in R^n: x = (1 - \alpha)a + \alpha b, \alpha \in R\}$ называется *прямой*, проходящей через точки a и b . Ее подмножества $[a, b] = \{x \in R^n: x = (1 - \alpha)a + \alpha b, \alpha \in [0, 1]\}$ и $L^+ = \{x \in R^n: x = (1 - \alpha)a + \alpha b, \alpha \geq 0\}$ называются соответственно *отрезком*, соединяющим точки a и b , и *лучом*, идущим из a через b . Заметим, что отрезок можно определить и при $a = b$: в этом случае он состоит из единственной точки.

Множество $H \subset R^n$ называется *аффинным*, если $H \neq \emptyset$ и вместе с любыми двумя точками содержит прямую, проходящую через эти точки. Всякое подпространство в R^n является аффинным множеством (множество, состоящее из одной точки, — аффинное).

Если H — аффинное множество и $a \in H$, то множество $H - a = \{x \in R^n: x = h - a, h \in H\}$ — подпространство в R^n . Обратно, если T — подпространство в R^n и $a \in R^n$, то $T + a$ — аффинное множество в R^n . Таким образом, класс аффинных множеств состоит из всех подмножеств пространства R^n , получаемых сдвигами подпространств пространства R^n (на векторы из R^n).

Если H — аффинное множество, то существует единственное подпространство T , сдвигом которого получается H : $T = H - a$,

торы образуют базис подпространства, сдвигом которого получается H).

Если H — гиперплоскость в R^n , то существуют такой вектор $a \in R^n$, $a \neq 0$, и число β , что $H = \{x \in R^n : ax = \beta\}$. Уравнение $ax = \beta$ ($\sum_{j=1}^n a[j] x[j] = \beta$) называется *уравнением гиперплоскости*, a — *нормальным вектором гиперплоскости H* ; вектор a ортогонален подпространству $T = \{x \in R^n : ax = 0\}$, сдвигом которого получается H . Другими словами, каждая гиперплоскость в R^n есть множество уровня соответствующей линейной функции $f(x) = ax$.

Верно и обратное, всякое уравнение $ax = \beta$, $a \neq 0$, $\beta \in R$, определяет гиперплоскость в R^n , причем равносильные уравнения определяют одну и ту же гиперплоскость.

С каждой гиперплоскостью $H = \{x \in R^n : ax = \beta\}$ связывают два множества: $\Pi^+ = \{x \in R^n : ax > \beta\}$ и $\Pi^- = \{x \in R^n : ax < \beta\}$. Эти множества называются *полупространствами*. Говорят, что гиперплоскость H *делит* пространство R^n на два полупространства Π^+ и Π^- : если точки c и d принадлежат одному полупространству (Π^+ или Π^-), то $[c, d] \cap H = \emptyset$; если точки c и d принадлежат разным полупространствам, то $[c, d] \cap H \neq \emptyset$. Очевидно также, что $\Pi^+ \cap \Pi^- = \emptyset$, $\Pi^+ \cup \Pi^- \cup H = R^n$.

Множества $\bar{\Pi}^+ = \Pi^+ \cup H$ и $\bar{\Pi}^- = \Pi^- \cup H$ называются *замкнутыми полупространствами*, порожденными гиперплоскостью H .

1.8. Симметричная квадратная матрица $A = (a[i, j], i, j \in \overline{1:n})$ называется *неотрицательно определенной*, если для любого $x \in R^n$

$$x^T A x = \sum_{i=1}^n \sum_{j=1}^n a[i, j] x[i] x[j] \geq 0,$$

и *положительно определенной*, если $x^T A x > 0$ для любого $x \in R^n$, $x \neq 0$. Аналогично симметричная матрица A называется *неположительно определенной*, если $x^T A x \leq 0$ для любого $x \in R^n$, и *отрицательно определенной*, если $x^T A x < 0$ для любого $x \in R^n$, $x \neq 0$.

Матрица A положительно (неотрицательно) определена в том и только том случае, если матрица $-A$ отрицательно (неположительно) определена. Известны эффективные критерии положительной (отрицательной) и неотрицательной (неположительной) определенности матрицы.

Пусть A — квадратная симметричная матрица порядка n . Определители

$$\Delta_1 = a[1, 1], \quad \Delta_2 = \begin{vmatrix} a[1, 1] & a[1, 2] \\ a[2, 1] & a[2, 2] \end{vmatrix}, \quad \dots$$

называются *главными минорами матрицы A* . Следующее условие положительной (отрицательной) определенности матрицы называется *условием Сильвестра*:

для положительной определенности матрицы A необходимо и достаточно выполнение неравенств $\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0$; для отрицательной определенности матрицы A необходимо и достаточно выполнение неравенств $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots$

Условия $\Delta_1 \geq 0, \Delta_2 \geq 0, \dots, \Delta_n \geq 0$ ($\Delta_1 \leq 0, \Delta_2 \geq 0, \dots$) являются необходимыми, но не достаточными для неотрицательной (неположительной) определенности матрицы. Вычислительные процедуры проверки неотрицательной (неположительной) определенности матрицы обычно более трудоемки, чем процедуры проверки положительной (отрицательной) определенности. Например, матрица является неотрицательно определенной (неположительно определенной), если все корни ее характеристического многочлена (они все вещественные и их число с учетом кратностей равно n) неотрицательны (неположительны). Есть и другие способы проверки неотрицательной (неположительной) определенности матрицы.

§ 2. Некоторые обозначения, понятия и термины математического анализа

2.1. В R^n вводится следующее отношение порядка ($\Leftrightarrow$ — знак равносильности):

$$a < b \Leftrightarrow a[i] < b[i] \text{ для любого } i \in 1:n;$$

$$a \leq b \Leftrightarrow a[i] \leq b[i] \text{ для любого } i \in 1:n \text{ и } a \neq b;$$

$$a \leq b \Leftrightarrow a[i] \leq b[i] \text{ для любого } i \in 1:n.$$

Пусть $J \subset 1:n$ (J — некоторое подмножество индексов из $1:n$). Обозначим

$$R_J^{n+} = \{x \in R^n : x[j] \geq 0, j \in J\}, \quad R^{n+} = \{x \in R^n : x \geq 0\},$$

в частном случае $R^+ = \{x \in R : x \geq 0\}$.

2.2. Нормой, или длиной, вектора $a \in R^n$ называется число $\|a\| = \sqrt{aa}$. Расстояние $\|a - b\|$ между двумя точками a и b из R^n обозначим через $\rho(a, b)$. Шар $\{x \in R^n : \rho(a, x) < r\}$ с центром в точке $a \in R^n$ и радиусом $r > 0$ обозначается через $\mathcal{D}(a, r)$, замкнутый шар $\{x \in R^n : \rho(a, x) \leq r\}$ — через $\overline{\mathcal{D}}(a, r)$, сфера $\{x \in R^n : \rho(a, x) = r\}$ — через $S(a, r)$.

Пусть $a \in R^n$, $b \in R^n$ и $a < b$. Множество $P(a, b) = \{x \in R^n : a < x < b\}$ называется прямоугольником, множество $P[a, b] = \{x \in R^n : a \leq x \leq b\}$ — замкнутым прямоугольником (замкнутый прямоугольник может быть определен и при $a \leq b$, $a \leq b$). Если $a \in R^n$, $d \in R^n$, $d > 0$, то $P(a-d, a+d)$ — прямоугольник с центром в точке a ; $P[a-d, a+d]$ — замкнутый прямоугольник с центром в точке a , если $d \geq 0$.

2.3. Обычным образом определяется топология в R^n , согласованная с его метрической структурой.

Пространство R^n и пустое множество $\emptyset$ являются одновременно замкнутыми и открытыми множествами в R^n . Других подмножеств в R^n , отличных от R^n и $\emptyset$, являющихся одновременно открытыми и замкнутыми, не существует.

Шары и прямоугольники являются открытыми множествами. Замкнутые шары и замкнутые прямоугольники — замкнутые множества. Границей шара $\mathcal{D}(a, r)$ является сфера $S(a, r)$. Аффинные множества в R^n (в частности, подпространства пространства R^n) — замкнутые множества. Полупространства в R^n — открытые множества, замкнутые полупространства — замкнутые множества. Границей полупространства (замкнутого полупространства) является гиперплоскость, которая порождает это полупространство. Всякое множество, состоящее из конечного числа точек, замкнуто.

Множество всех внутренних точек множества M — его *открытое ядро* — обозначается через M° или $\mathring{M}$. Граница множества M обозначается через ∂M . *Замыкание* множества M обозначается через $\bar{M}$.

2.4. Отображение множества N всех натуральных чисел в R^n называется *последовательностью* в R^n . Последовательности будем обозначать через $\{x_k\}_{k=1}^\infty$, или $x_1, x_2, \dots$. Если $M \subset R^n$ и $x_k \in M$ при любом $k \in N$, то будем говорить, что $\{x_k\}_{k=1}^\infty$ — *последовательность элементов множества M* .

Если $\{x_k\}_{k=1}^\infty$ — последовательность в R^n (или в M) и K — бесконечное подмножество множества N , то ограничение отображения $\{x_k\}_{k=1}^\infty$ на множестве K называется *подпоследовательностью* (или *частичной последовательностью*) последовательности $\{x_k\}_{k=1}^\infty$ и обозначается через $\{x_k\}_{k \in K}$. Если последовательность $\{x_k\}_{k=1}^\infty$ имеет предел x_0 (сходится к x_0), то будем писать $x_k \rightarrow x_0$ при $k \rightarrow \infty$, или $\lim_{k \rightarrow \infty} x_k = x_0$. Если x_0 есть предел подпоследовательности $\{x_k\}_{k \in K}$, то будем писать $x_k \rightarrow x_0$ при $k \rightarrow \infty$, $k \in K$.

2.5. Пусть A и B — непустые множества в R^n . *Расстоянием между множествами A и B* называется число $\rho(A, B) = \inf \{\rho(a, b), a \in A, b \in B\}$. Если существуют такие точки $\bar{a} \in A$, $\bar{b} \in B$, что $\rho(\bar{a}, \bar{b}) = \rho(A, B)$, то говорят, что расстояние между множествами A и B *осуществляется*. Некоторые достаточные условия осуществления расстояния будут приведены в п. 4.5.

2.6. Множество $M \subset R^n$ называется *компактным*, если из любой последовательности точек этого множества можно выделить подпоследовательность, сходящуюся к точке этого множества.

Необходимым и достаточным условием компактности множества $M \subset R^n$ является его ограниченность и замкнутость.

Пересечение любого множества компактных множеств есть множество компактное. Пересечение компактного и замкнутого множеств является множеством компактным. Объединение конечного числа компактных множеств — множество компактное. В частности, всякое конечное множество компактно.

2.7. Пусть $\{x_k\}_{k=1}^{\infty}$ — последовательность векторов из R^n и $a_1, a_2, \dots, a_n$ — базис пространства R^n . Каждый вектор x_k представим в виде линейной комбинации векторов $a_i, i \in 1:n$:

$$x_k = \sum_{i=1}^n \alpha_k^{(i)} a_i.$$

Рассмотрим n числовых последовательностей $\{\alpha_k^{(i)}\}_{k=1}^{\infty}, i \in 1:n$.

Пусть $x_0 \in R^n, x_0 = \sum_{i=1}^n \alpha_0^{(i)} a_i$.

Последовательность $\{x_k\}_{k=1}^{\infty}$ в том и только в том случае сходится к вектору x_0 при $k \rightarrow \infty$, если для любого $i \in 1:n$ числовая последовательность $\{\alpha_k^{(i)}\}_{k=1}^{\infty}$ сходится к $\alpha_0^{(i)}$ при $k \rightarrow \infty$. Иными словами, *сходимость в R^n эквивалентна покоординатной сходимости*.

2.8. Функция (отображение) $f: R^n \rightarrow R^m$ называется *непрерывной в точке $x_0 \in R^n$* , если полный прообраз любой окрестности точки $f(x_0) \in R^m$ содержит некоторую окрестность точки x_0 . Функция f *непрерывна на множестве $M \subset R^n$* , если она непрерывна в каждой точке этого множества.

Пусть $f: R^n \rightarrow R^m, M \subset R^n$ и $x_0 \in M$. Функция f *непрерывна в точке $x_0 \in M$ относительно множества M (или относительно непрерывна в точке x_0 множества M)*, если полный прообраз любой окрестности точки $f(x_0)$ содержит пересечение некоторой окрестности точки x_0 с множеством M . Функция f *относительно непрерывна на множестве M* , если f относительно непрерывна в каждой точке множества M .

Аналогично определяется непрерывность (относительная непрерывность) в точке $x_0 \in M$ и на множестве M функции, определенной не на всем пространстве R^n , а лишь на множестве $M' \supset M$.

Функция $f: R^n \rightarrow R^m$ в том и только в том случае непрерывна в точке x_0 (относительно непрерывна в точке $x_0 \in M$), если для любой последовательности $\{x_k\}_{k=1}^{\infty}$, сходящейся к точке x_0 (любой последовательности $\{x_k\}_{k=1}^{\infty}$ точек множества M , сходящейся к точке $x_0 \in M$), последовательность $\{f(x_k)\}_{k=1}^{\infty}$ сходится к $f(x_0)$.

Некоторые примеры. *Аффинное отображение $f: R^n \rightarrow R^m, f(x) = Ax + b$, где $A = (a[i, j]) — m \times n$ матрица и $b \in R^m$, непрерывно на R^n . В частности, непрерывна на R^n всякая аффинная функция $f: R^n \rightarrow R, f(x) = ax + \beta, a \in R^n, \beta \in R$.*

Если $f: R^n \rightarrow R^m$ — непрерывное на R^n отображение, то $\|f\|: R^n \rightarrow R, \|f\|(x) = \|f(x)\|$ (здесь $\|\cdot\|$ — норма в R^m), — непре-

рывная числовая функция на R^n . Функция $\rho: R^n \times R^n \rightarrow R$, $\rho(x, y) = \|x - y\|$, — непрерывная функция на $R^n \times R^n = R^{2n}$.

Если M — компактное множество в R^n и f непрерывна (относительно непрерывна) на M , то $f(M)$ — компактное множество в R^m . В частности, числовая функция f , непрерывная на компактном множестве $M \subset R^n$, принимает на этом множестве наибольшее и наименьшее значения, т. е. в M существуют такие точки $\bar{x}$ и $\underline{x}$, что $f(x) \leq f(\bar{x})$ и $f(x) \geq f(\underline{x})$ для любого $x \in M$ (теорема Вейерштрасса). Наибольшее значение функции f на множестве M обозначается $\max_{x \in M} f(x)$, а наименьшее $\min_{x \in M} f(x)$.

2.9. Пусть M — открытое множество в R^n и f — числовая функция на M . Функция f называется *дифференцируемой в точке* $x \in M$, если существует такой вектор $x^* \in R^n$, что для любого $\varepsilon > 0$ найдется такое $r > 0$, что шар $\mathcal{D}(x, r) \subset M$ и

$$|f(x + s) - f(x) - x^*s| \leq \varepsilon \|s\|$$

для любого $s \in \mathcal{D}(0, r)$. Если функция f дифференцируема в точке x , то вектор x^* определяется однозначно и называется *градиентом* (вектором градиента) функции f в точке x . Градиент функции f в точке x обозначается $\nabla f(x)$, или $\text{grad } f(x)$ (мы будем пользоваться первым обозначением). При этом

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x[1]}, \frac{\partial f(x)}{\partial x[2]}, \dots, \frac{\partial f(x)}{\partial x[n]} \right),$$

где $\partial f(x)/\partial x[j]$ — частная производная функции f по j -й координате, $j \in 1:n$. Линейная функция $g_{f,x}: R^n \rightarrow R$, $g_{f,x}(s) = \nabla f(x)s$, где $\nabla f(x)s$ — скалярное произведение векторов $\nabla f(x)$ и s , называется *дифференциалом функции f в точке x* . Функция f называется *дифференцируемой на множестве M* , если она дифференцируема в каждой точке этого множества. Отображение $x \mapsto g_{f,x}$ множества M в множество $(R^n)^*$ линейных функций на R^n (или отображение $x \mapsto \nabla f(x)$ множества M в R^n) называется *производным отображением функции f* . Функция f называется *непрерывно дифференцируемой на множестве M* , если ее производное отображение непрерывно на этом множестве. Если $M = R^n$, то говорят, что f *дифференцируема* (непрерывно дифференцируема).

Необходимым и достаточным условием непрерывной дифференцируемости функции f на M является существование и непрерывность на M частных производных $\partial f/\partial x[j]$, $j \in 1:n$, функции f .

Функция $f: M \rightarrow R$, где M — открытое множество в R^n , является *дважды дифференцируемой на M* (дважды непрерывно дифференцируемой на M), если f непрерывно дифференцируема на M и каждая из функций $\partial f/\partial x[j]$, $j \in 1:n$, дифференцируема (непрерывно дифференцируема на M). Функция f дважды непрерывно дифференцируема на M в том и только том случае,

если существуют и непрерывны на M все вторые частные производные $\partial^2 f / \partial x[i] \partial x[j]$, $i \in 1:n$, $j \in 1:n$, функции f .

С каждой дважды непрерывно дифференцируемой на M функцией f связывается симметричная квадратная матрица $G(x) = (g(x) [i, j])$ — матрица Гессе, строками которой являются градиенты частных производных функции f , т. е. $g(x) [i, j] = \partial^2 f(x) / \partial x[i] \partial x[j]$.

2.10. Пусть M — открытое множество в R^n и $f: M \rightarrow R$. При фиксированных $x \in M$, $s \in R^n$ рассмотрим функцию $\varphi(\lambda) = f(x + \lambda s)$, определенную на множестве тех $\lambda \in R$, для которых $x + \lambda s \in M$. Если f непрерывно дифференцируема на M , то для любого $s \in R^n$ существует предел

$$f'(x, s) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} (f(x + \lambda s) - f(x))$$

(т. е. предел при $\lambda \rightarrow 0$ по множеству положительных чисел λ). Этот предел называется производной функции f в точке x по направлению s . Кроме того,

$$f'(x, s) = \varphi'(0) = \nabla f(x) s = \sum_{j=1}^n \frac{\partial f(x)}{\partial x[j]} s[j]. \quad (1)$$

Обратное неверно: функция f может иметь в точке x производную по любому направлению s и не быть ни дифференцируемой, ни даже непрерывной в этой точке.

Пусть теперь f дважды непрерывно дифференцируема на M , $x \in M$ и $s \in R^n$. Для таких $\lambda \in R$, что $x + \lambda s \in M$, положим $\psi(\lambda) = \nabla f(x + \lambda s) s = \varphi'(\lambda)$. Тогда

$$\psi'(\lambda) = s^T G(x + \lambda s) s = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(x + \lambda s)}{\partial x[i] \partial x[j]} s[i] s[j]. \quad (2)$$

2.11. Точка $\bar{x} \in R^n$ называется точкой локального максимума (локального минимума) функции $f: R^n \rightarrow R$, если существует такая окрестность $\mathcal{D}(\bar{x}, r)$ точки $\bar{x}$, что $f(x) \leq f(\bar{x})$ ($f(x) \geq f(\bar{x})$) для любого $x \in \mathcal{D}(\bar{x}, r)$. Точки локального максимума и локального минимума функции f называются точками экстремума функции f .

Если f — непрерывно дифференцируемая функция и $\bar{x}$ — точка экстремума функции f , то $\nabla f(\bar{x}) = 0$. Таким образом, точки экстремума непрерывно дифференцируемой функции f находятся в множестве решений системы уравнений

$$\partial f(x) / \partial x[j] = 0, \quad j \in 1:n. \quad (3)$$

Векторы из R^n , являющиеся решениями системы (3), называются стационарными точками функции f . Стационарные точки функции f могут не быть ее точками экстремума.

Пусть f — дважды непрерывно дифференцируемая функция. Для того чтобы стационарная точка x_0 функции f была точкой ее

локального максимума (локального минимума), достаточна положительная определенность (отрицательная определенность) матрицы Гессе $G(x_0)$ или неотрицательная (неположительная) определенность матрицы $G(x)$ в некоторой окрестности точки x_0 .

§ 3. Выпуклые множества в R^n

3.1. Определение. Некоторые свойства. Множество $V \subset R^n$ называется *выпуклым*, если вместе с двумя любыми его точками оно содержит и весь отрезок, соединяющий эти точки, а именно V выпукло в том и только том случае, если из условий $a \in V$, $b \in V$ следует $[a, b] = \{(1-\alpha)a + \alpha b, \alpha \in [0, 1]\} \in V$.

В частности, пустое множество, множество, состоящее из одной точки, и все R^n являются выпуклыми множествами. Нетрудно показать, что *подпространства R^n , аффинные подмножества в R^n , шары и прямоугольники являются выпуклыми множествами.* — Упражнение 1.13. Естественно, этими примерами не исчерпывается совокупность выпуклых подмножеств пространства R^n .

Очевидно, что *пересечение любого числа выпуклых множеств есть выпуклое множество.*

Справедливы следующие утверждения.

Замыкание выпуклого множества есть множество выпуклое.

Доказательство. Пусть V — выпуклое множество, $\bar{V}$ — его замыкание, $a \in \bar{V}$, $b \in \bar{V}$, $\alpha \in [0, 1]$. Покажем, что $(1-\alpha)a + \alpha b \in \bar{V}$. Существуют такие последовательности $\{a_k\}_{k=1}^{\infty}$, $\{b_k\}_{k=1}^{\infty}$ точек множества V , что $\lim_{k \rightarrow \infty} a_k = a$, $\lim_{k \rightarrow \infty} b_k = b$.

Так как V выпукло, то все точки последовательности $\{(1-\alpha)a_k + \alpha b_k\}_{k=1}^{\infty}$ принадлежат V . Но $\|(1-\alpha)a_k + \alpha b_k - ((1-\alpha)a + \alpha b)\| \leq (1-\alpha)\|a_k - a\| + \alpha\|b_k - b\| \rightarrow 0$ при $k \rightarrow \infty$, так что $(1-\alpha)a + \alpha b \in \bar{V}$. ■

Открытое ядро выпуклого множества есть множество выпуклое.

Доказательство. Пусть V — выпуклое множество, V^0 — его открытое ядро, $a \in V^0$, $b \in V^0$, $\alpha \in [0, 1]$. Покажем, что $(1-\alpha)a + \alpha b \in V^0$. Существуют такие положительные числа r_a , r_b , что $\mathcal{D}(a, r_a) \subset V$, $\mathcal{D}(b, r_b) \subset V$. Положим $r = \min\{r_a, r_b\}$ и покажем, что $\mathcal{D}((1-\alpha)a + \alpha b, r) \subset V$.

Пусть $z \in \mathcal{D}((1-\alpha)a + \alpha b, r)$. Положим $z_1 = a + z - (1-\alpha)a - \alpha b$, $z_2 = b + z - (1-\alpha)a - \alpha b$. Тогда $\|z_1 - a\| = \|z - (1-\alpha)a - \alpha b\| < r \leq r_a$, $\|z_2 - b\| = \|z - (1-\alpha)a - \alpha b\| < r \leq r_b$, т. е. $z_1 \in \mathcal{D}(a, r_a) \subset V$, $z_2 \in \mathcal{D}(b, r_b) \subset V$. Так как V — выпуклое множество и $\alpha \in [0, 1]$, то $(1-\alpha)z_1 + \alpha z_2 \in V$. Но $(1-\alpha)z_1 + \alpha z_2 = (1-\alpha)(a + z - (1-\alpha)a - \alpha b) + \alpha(b + z - (1-\alpha)a - \alpha b) = z$. Поэтому $z \in V$ и $\mathcal{D}((1-\alpha)a + \alpha b, r) \subset V$. ■

Иллюстрацией этого доказательства является рис. 1.

Если V_1 и V_2 — выпуклые множества в R^n , то $V_1 + V_2 = \{v \in R^n : v = v_1 + v_2, v_1 \in V_1, v_2 \in V_2\}$ — выпуклое множество.

Если V — выпуклое множество в R^n и $a \in R$, то $aV = \{w \in R^n : w = av, v \in V\}$ — выпуклое множество.

Если V_1 и V_2 — выпуклые множества в R^n , то $V_1 - V_2 = \{v \in R^n : v = v_1 - v_2, v_1 \in V_1, v_2 \in V_2\}$ — выпуклое множество.

Докажем первое из этих утверждений. Доказательство второго и третьего утверждений предлагаем читателю провести самостоятельно. — Упражнения 1.15, 1.16.

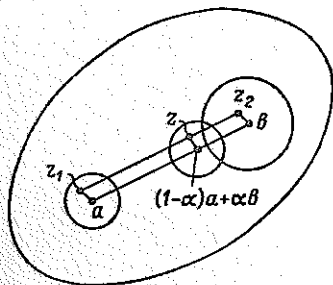


Рис. 1.

Пусть V_1 и V_2 — выпуклые множества, $W = V_1 + V_2$, $w' \in W$, $w'' \in W$, $\alpha \in [0, 1]$. Покажем, что $(1 - \alpha)w' + \alpha w'' \in W$. Действительно, существуют такие v'_1, v'_2 из V_1 и v''_1, v''_2 из V_2 , что $w' = v'_1 + v'_2$, $w'' = v''_1 + v''_2$. Тогда $(1 - \alpha)w' + \alpha w'' = (1 - \alpha)(v'_1 + v'_2) + \alpha(v''_1 + v''_2) = ((1 - \alpha)v'_1 + \alpha v''_1) + ((1 - \alpha)v'_2 + \alpha v''_2)$. Но $(1 - \alpha)v'_1 + \alpha v''_1 \in V_1$, $(1 - \alpha)v'_2 + \alpha v''_2 \in V_2$, и, следовательно, $(1 - \alpha)w' + \alpha w'' \in W$. ■

3.2. Выпуклая оболочка. Выпуклой оболочкой множества $M \subset R^n$ назовем наименьшее выпуклое множество в R^n , содержащее M . Точнее, выпуклая оболочка множества $M \subset R^n$ есть такое множество $v(M)$, что: 1) $M \subset v(M)$; 2) $v(M)$ — выпуклое множество; 3) если V' — какое-нибудь выпуклое множество, содержащее M , то $V' \supset v(M)$.

Выпуклая оболочка множества M очевидно совпадает с пересечением всех выпуклых множеств, содержащих M .

Некоторые примеры: отрезок $[a, b]$ есть выпуклая оболочка точек a и b ; выпуклая оболочка трех точек, не лежащих на одной прямой, — треугольник; замкнутый шар $\mathcal{D}(a, r)$ есть выпуклая оболочка сферы $S(a, r)$; выпуклая оболочка множества неотрицательных координатных лучей в R^n есть R^n .

Линейная комбинация $\alpha_1 a_1 + \dots + \alpha_k a_k$ элементов $a_1, \dots, a_k$ называется *выпуклой линейной комбинацией*, если $\alpha_i \geq 0$, $i \in \{1, \dots, k\}$, и $\sum_{i=1}^k \alpha_i = 1$.

Выпуклая оболочка множества M совпадает с множеством всех выпуклых линейных комбинаций элементов множества M .

Доказательство. Обозначим множество всех выпуклых линейных комбинаций элементов множества M через V и покажем, что $V = v(M)$. Для этого проверим, что для V выполняются свойства 1) — 3) определения выпуклой оболочки.

Для любого $a \in M$ линейная комбинация $1 \cdot a$ есть выпуклая линейная комбинация, и, следовательно, $a \in V$ и $M \subset V$.

Если $v_1 = \sum_{i=1}^k \alpha_i a_i$ и $v_2 = \sum_{j=1}^m \beta_j b_j$ — выпуклые линейные комбинации элементов множества M (т. е. $\alpha_i \geq 0$, $i \in 1:k$, $\sum_{i=1}^k \alpha_i = 1$, $\beta_j \geq 0$, $j \in 1:m$, $\sum_{j=1}^m \beta_j = 1$) и $\lambda \in [0, 1]$, то

$$(1 - \lambda) v_1 + \lambda v_2 = \sum_{i=1}^k (1 - \lambda) \alpha_i a_i + \sum_{j=1}^m \lambda \beta_j b_j$$

— выпуклая линейная комбинация элементов $a_1, \dots, a_k, b_1, \dots, b_m$, так как все коэффициенты этой линейной комбинации неотрицательны, а их сумма

$$\sum_{i=1}^k (1 - \lambda) \alpha_i + \sum_{j=1}^m \lambda \beta_j = (1 - \lambda) \sum_{i=1}^k \alpha_i + \lambda \sum_{j=1}^m \beta_j = 1 - \lambda + \lambda = 1.$$

Следовательно, $(1 - \lambda) v_1 + \lambda v_2 \in V$ и V — выпуклое множество.

Наконец, покажем, что если V' — выпуклое множество, содержащее M , то $V \subset V'$. Доказательство проведем по индукции.

Назовем «длиной» линейной комбинации $\sum_{i=1}^k \alpha_i a_i$ число k . По условию $V' \supset M$, т. е. V' содержит все выпуклые линейные комбинации элементов M длины 1. Предположим, что $m \geq 1$ и V' содержит все выпуклые линейные комбинации, для которых $k \in 1:m$.

Рассмотрим какую-нибудь выпуклую линейную комбинацию $\sum_{i=1}^{m+1} \alpha_i a_i$ ($\alpha_i \geq 0$, $i \in 1:m+1$, $\sum_{i=1}^{m+1} \alpha_i = 1$). Можно считать, что $0 < \alpha_{m+1} < 1$, так как в противном случае $\sum_{i=1}^{m+1} \alpha_i a_i$ можно было бы рассматривать как выпуклую линейную комбинацию, длина которой не превосходит m . Тогда

$$\sum_{i=1}^{m+1} \alpha_i a_i = (1 - \alpha_{m+1}) \sum_{i=1}^m \frac{\alpha_i}{1 - \alpha_{m+1}} a_i + \alpha_{m+1} a_{m+1} \in V',$$

поскольку $\sum_{i=1}^m \frac{\alpha_i}{1 - \alpha_{m+1}} a_i \in V'$ в силу предположения индукции ($\frac{\alpha_i}{1 - \alpha_{m+1}} \geq 0$, $i \in 1:m$; $\sum_{i=1}^m \frac{\alpha_i}{1 - \alpha_{m+1}} = \frac{1}{1 - \alpha_{m+1}} \sum_{i=1}^m \alpha_i = 1$), $a_{m+1} \in V'$, $\alpha_{m+1} \in (0, 1)$ и V' — выпуклое множество. ■

На самом деле для построения выпуклой оболочки множества $M \subset R^n$ нет необходимости рассматривать все выпуклые линейные комбинации элементов множества M : множество $v(M)$ состоит из выпуклых линейных комбинаций, длины которых не превосходят $n + 1$ (теорема Каратеодори).

Доказательство. Предположим, что $v \in v(M)$,

$$v = \sum_{i=0}^m \alpha_i a_i, \quad a_i \in M, \quad \alpha_i > 0, \quad i \in 0:m, \quad \sum_{i=0}^m \alpha_i = 1, \quad (1)$$

и пусть $m > n$. Покажем, как в этом случае заменить представление (1) точки v выпуклой линейной комбинацией элементов множества M , длина которой меньше $m + 1$. Тогда для завершения доказательства теоремы достаточно будет повторять

такое преобразование, пока точка v не будет представлена в виде выпуклой линейной комбинации элементов множества M , длина которой не превосходит $n + 1$.

Так как $m > n$, то система векторов $a_1 - a_0, \dots, a_m - a_0$ линейно зависима и найдутся такие числа $\beta_i, i \in 1:m$, что $\sum_{i=1}^m \beta_i (a_i - a_0) = (-\sum_{i=1}^m \beta_i) a_0 + \sum_{i=1}^m \beta_i a_i = 0$ и не все β_i равны нулю. Но тогда

$$v = (a_0 + \lambda \sum_{i=1}^m \beta_i) a_0 + \sum_{i=1}^m (\alpha_i - \lambda \beta_i) a_i \quad (2)$$

при любом $\lambda \in R$, и нетрудно указать такое $\bar{\lambda}$, чтобы при $\lambda = \bar{\lambda}$ все коэффициенты разложения (2) были неотрицательны, но по крайней мере один из них равнялся нулю. Так как сумма коэффициентов в (2) равна единице, то при $\lambda = \bar{\lambda}$ получим искомое представление точки v . ■

Установим, кроме того, следующие свойства выпуклой оболочки множества.

Пусть $M \subset R^n$ и $v(M)$ — выпуклая оболочка множества M . Тогда $\text{aff } M = \text{aff } v(M)$. Действительно, поскольку $M \subset v(M)$, то и $\text{aff } M \subset \text{aff } v(M)$. Вместе с тем $v(M) \subset \text{aff } M$, и, следовательно, $\text{aff } v(M) \subset \text{aff } (\text{aff } M) = \text{aff } M$. ■

Выпуклая оболочка $v(M)$ ограниченного множества M есть множество ограниченное. Действительно, если $M \subset \bar{D}(0, r)$ и $\sum_{i=1}^m \alpha_i a_i \in v(M)$ ($a_1, \dots, a_m$ из $M, \alpha_i \geq 0, i \in 1:m, \sum_{i=1}^m \alpha_i = 1$), то

$$\|\sum_{i=1}^m \alpha_i a_i\| \leq \sum_{i=1}^m \alpha_i \|a_i\| \leq r \sum_{i=1}^m \alpha_i = r. \quad \blacksquare$$

3.3. Выпуклые многогранники. Выпуклая оболочка конечного множества точек называется *выпуклым многогранником*. Ясно, что выпуклый многогранник является выпуклым множеством.

Теорема. *Выпуклый многогранник есть компактное выпуклое множество.*

Доказательство. Ограниченность выпуклого многогранника $V = v(\{a_1, \dots, a_m\})$ следует из ограниченности конечного множества $\{a_1, \dots, a_m\}$. Докажем замкнутость множества V .

Пусть $\{v_k\}_{k=1}^\infty$ — последовательность элементов множества V , сходящаяся к v_0 . Докажем, что $v_0 \in V$. Так как $v_k \in V$, то $v_k = \sum_{i=1}^m \alpha_i^{(k)} a_i, \alpha_i^{(k)} \geq 0, i \in 1:m, \sum_{i=1}^m \alpha_i^{(k)} = 1$ при каждом $k \in 1: +\infty$. Из ограниченной последовательности $\{\alpha_1^{(k)}\}_{k=1}^\infty$ выделим сходящуюся подпоследовательность $\{\alpha_1^{(k)}\}_{k \in K_1}, K_1 \subset 1: +\infty$ и пусть $\alpha_1^{(0)}$ — ее предел. Затем из последовательности $\{\alpha_2^{(k)}\}_{k \in K_1}$ выделим сходящуюся подпоследовательность $\{\alpha_2^{(k)}\}_{k \in K_2}, K_2 \subset K_1$ и пусть $\alpha_2^{(0)}$ — ее предел. Отметим, что подпоследовательность $\{\alpha_1^{(k)}\}_{k \in K_2}$ последовательности $\{\alpha_1^{(k)}\}_{k \in K_1}$ сходится к $\alpha_1^{(0)}$.

Продолжая этот процесс, получим такую подпоследовательность K_m последовательности всех натуральных чисел, что при каждом $i \in 1:m$ последовательность $\{\alpha_i^{(k)}\}_{k \in K_m}$ будет сходиться к своему пределу $\alpha_i^{(0)}$. При любом $k \in K_m$ имеем $\alpha_i^{(k)} \geq 0$, $i \in 1:m$, $\sum_{i=1}^m \alpha_i^{(k)} = 1$, и, следовательно, $\alpha_i^{(0)} \geq 0$, $i \in 1:m$, $\sum_{i=1}^m \alpha_i^{(0)} = 1$.

Последовательность $\{v_k\}_{k \in K_m} = \{\sum_{i=1}^m \alpha_i^{(k)} a_i\}_{k \in K_m}$ сходится к v_0 . Вместе с тем

$$\left\| \sum_{i=1}^m \alpha_i^{(k)} a_i - \sum_{i=1}^m \alpha_i^{(0)} a_i \right\| \leq \sum_{i=1}^m |\alpha_i^{(k)} - \alpha_i^{(0)}| \cdot \|a_i\| \rightarrow 0, \\ k \rightarrow \infty, \quad k \in K_m.$$

Поэтому $v_0 = \sum_{i=1}^m \alpha_i^{(0)} a_i \in V$. ■

Замечание. Доказанная выше теорема является частным случаем более общей теоремы: *выпуклая оболочка компактного множества в R^n есть компактное множество*. Доказательство этого утверждения см., например, в [43].

Выпуклая оболочка замкнутого (неограниченного) множества может и не быть замкнутым множеством. Так, *выпуклая оболочка множества на плоскости, состоящего из прямой и не принадлежащей этой прямой точки, не является замкнутым множеством*. — Упражнение 1.17.

3.4. Симплексы. Выпуклый многогранник $C_m = v(\{a_0, a_1, \dots, a_m\})$, аффинная оболочка которого есть аффинное множество размерности m , называется m -симплексом.

Примеры. $C_0 = v(\{a_0\})$ — точка a_0 ; $C_1 = v(\{a_0, a_1\})$, $a_0 \neq a_1$, — отрезок $[a_0, a_1]$; $C_2 = v(\{a_0, a_1, a_2\})$, точки a_0, a_1, a_2 не лежат на одной прямой, — треугольник; $C_3 = v(\{a_0, a_1, a_2, a_3\})$, точки a_0, a_1, a_2, a_3 не лежат в одной плоскости, — тетраэдр (рис. 2).

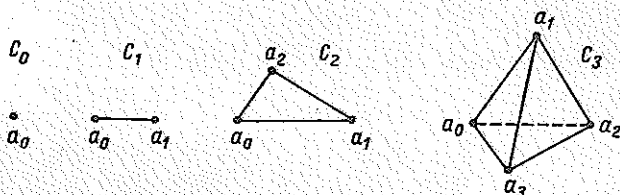


Рис. 2.

Для того чтобы выпуклый многогранник $V = v(\{a_0, a_1, \dots, a_m\})$ был m -симплексом, необходимо и достаточно, чтобы система векторов $a_1 - a_0, a_2 - a_0, \dots, a_m - a_0$ была линейно независимой.

Доказательство. Поскольку $\text{aff } V = \text{aff } \{a_0, a_1, \dots, a_m\}$, то $\dim \text{aff } V$ равна максимальному числу линейно независимых

векторов в системе $a_1 - a_0, \dots, a_m - a_0$. Поэтому $\dim \text{aff } V = m$ в том и только том случае, если система $a_1 - a_0, \dots, a_m - a_0$ линейно независима. ■

3.5. Относительно внутренние точки выпуклого множества. Пусть $M \subset R^n$ и $\text{aff } M$ — аффинная оболочка множества M . Точка $a \in M$ называется *относительно внутренней точкой* множества M , если существует такой шар $\mathcal{D}(a, r)$, что $\mathcal{D}(a, r) \cap \text{aff } M \subset M$.

Пересечение $\mathcal{D}(a, r) \cap \text{aff } M$ будем называть *относительной окрестностью точки a в $\text{aff } M$* (вообще, если H — аффинное множество и $a \in H$, то $\mathcal{D}(a, r) \cap H$ — *относительная окрестность точки a в H*). Множество всех относительно внутренних точек множества M назовем его *относительно открытым ядром* и обозначим через M^{00} .

Примеры. Если $a \in R^3, b \in R^3, a \neq b$, то $[a, b]^0 = \emptyset$. Однако $[a, b]^{00} \neq \emptyset$ и состоит из всех точек $(1-\lambda)a + \lambda b, \lambda \in (0, 1)$ (рис. 3). Аффинной оболочкой множества $\{a_1, a_2, a_3\}$,

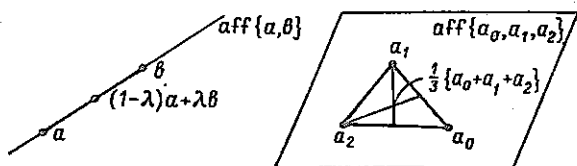


Рис. 3.

которое состоит из трех точек в R^3 , не лежащих на одной прямой, является плоскость. Очевидно, что $C_2^0 = (v(\{a_0, a_1, a_2\}))^0 = \emptyset$, а $C_2^{00} \neq \emptyset$. Так, например, точка $1/3(a_0 + a_1 + a_2)$ заведомо принадлежит C_2^{00} (рис. 3).

Замечание. Удобно считать, что относительно открытое ядро множества, состоящего из единственной точки, совпадает с этим множеством. Поэтому в дальнейшем в дополнение к данному выше определению относительно внутренней точки будем считать, что каждая точка из R^n — относительно внутренняя в своей (нульмерной) аффинной оболочке, состоящей из этой точки.

Если $a \in M^{00}$ (так что $\mathcal{D}(a, r) \cap \text{aff } M \subset M$) и $b \in \text{aff } M$, то найдется такое $\varepsilon > 0$, что $(1-\lambda)b + \lambda a \in M$ при всех λ , для которых $|\lambda - 1| < \varepsilon$. Действительно, $\{(1-\lambda)b + \lambda a, \lambda \in R\} \subset \text{aff } M$, и поскольку $(1-\lambda)b + \lambda a = a + (\lambda - 1)(a - b)$, то достаточно выбрать $\varepsilon > 0$ таким, чтобы $\varepsilon \|a - b\| < r$. ■

Относительно открытое ядро выпуклого множества есть множество выпуклое. — **Упражнение 1.18.** Доказательство этого утверждения по существу не отличается от доказательства аналогичного утверждения п. 3.1.

3.6. Теорема. *Относительно открытое ядро V^{00} непустого выпуклого множества V непусто.*

Доказательство этой теоремы разобьем на две части. Сначала докажем, что $C_m^{00} \neq \emptyset$ для любого m -симплекса C_m , затем установим, что если $\dim \text{aff } V = m$, то V содержит некоторый m -симплекс.

I. Пусть $C_m = v(\{a_0, \dots, a_m\})$ и векторы $a_1 - a_0, \dots, a_m - a_0$ линейно независимы. Возьмем точку $c = \frac{1}{m+1} \sum_{i=0}^m a_i = a_0 + \sum_{i=1}^m \frac{1}{m+1} (a_i - a_0)$. Очевидно, что $c \in C_m$. Покажем, что $c \in C_m^{00}$.

Предположим противное. Тогда существует такая последовательность $\{z_k\}_{k=1}^\infty$, сходящаяся к c при $k \rightarrow \infty$, что $z_k \notin \text{aff}(a_0, \dots, a_m)$, $z_k \notin C_m$ при всех $k \in \mathbb{N}$. Так как $\text{aff}\{a_0, \dots, a_m\} = \{a_0 + \sum_{i=1}^m \alpha_i (a_i - a_0), \alpha_i \in \mathbb{R}, i \in 1:m\}$, то каждый из векторов z_k можно представить в виде $z_k = a_0 + \sum_{i=1}^m \alpha_i^{(k)} (a_i - a_0)$. Поскольку $z_k \rightarrow c$ при $k \rightarrow \infty$, то $z_k - a_0 \rightarrow c - a_0$ при $k \rightarrow \infty$.

Система векторов $a_1 - a_0, a_2 - a_0, \dots, a_m - a_0$ линейно независима. Поэтому ее можно дополнить до базиса пространства $\mathbb{R}^n$. В этом базисе вектор $c - a_0$ имеет координаты $1/(m+1), \dots, 1/(m+1), 0, \dots, 0$, а векторы $z_k - a_0$, $k \in \mathbb{N}$, — координаты $\alpha_1^{(k)}, \dots, \alpha_m^{(k)}, 0, \dots, 0$. Так как сходимость в $\mathbb{R}^n$ эквивалентна покоординатной сходимости (см. п. 2.7), то $\alpha_i^{(k)} \rightarrow 1/(m+1)$ при $k \rightarrow \infty$ для каждого $i \in 1:m$.

Выберем такое k_0 , чтобы неравенства $|\alpha_i^{(k)} - 1/(m+1)| < 1/(m+1)^2$, $i \in 1:m$, выполнялись при любом $k \geq k_0$. Но тогда при $k \geq k_0$ выполняются неравенства $\alpha_i^{(k)} > 0$, $i \in 1:m$, $\sum_{i=1}^m \alpha_i^{(k)} < m/(m+1) + m/(m+1)^2 = [(m+1)^2 - 1]/(m+1)^2 < 1$, и, следовательно,

$$z_k = a_0 + \sum_{i=1}^m \alpha_i^{(k)} (a_i - a_0) = (1 - \sum_{i=1}^m \alpha_i^{(k)}) a_0 + \sum_{i=1}^m \alpha_i^{(k)} a_i$$

есть выпуклая линейная комбинация векторов $a_0, a_1, \dots, a_m$, т. е. $z_k \in C_m$ при $k \geq k_0$, что невозможно, так как $z_k \notin C_m$ при $k \in \mathbb{N}$. ■

II. Докажем теперь, что непустое выпуклое множество V размерности m содержит некоторый m -симплекс. Обозначим через r наибольшее из таких чисел k , что V содержит k -симплекс. Предположим, что $r < m$, и пусть $C_r = v(\{a_0, \dots, a_r\})$ — какой-нибудь r -симплекс, содержащийся в V . Поскольку $\dim \text{aff } V = m > r = \dim \text{aff } C_r$, в множестве V найдется точка $a_{r+1} \in \text{aff } C_r$. Но тогда векторы $a_1 - a_0, \dots, a_r - a_0, a_{r+1} - a_0$ линейно независимы (иначе вектор $a_{r+1} - a_0$ представим в виде линейной комбинации векторов $a_1 - a_0, \dots, a_r - a_0$

и $a_{r+1} \in \text{aff } C_r$ и $C_{r+1} = v(\{a_0, \dots, a_r, a_{r+1}\}) - (r+1)$ -симплекс, содержащийся в V . ■

Размерность аффинной оболочки выпуклого множества V назовем *размерностью выпуклого множества V* . Размерность выпуклого множества V будем обозначать через $\dim V$.

3.7. Введем теперь понятия относительно граничной точки множества M и относительной границы множества M .

Точка $a \in \text{aff } M$ называется *относительно граничной точкой*, если любая относительная окрестность $\mathcal{D}(a, r) \cap \text{aff } M$ содержит как точки множества M , так и точки его дополнения. Множество всех относительно граничных точек множества M будем называть его *относительной границей* и обозначать через $\tilde{\partial}M$. Очевидно, что $\tilde{\partial}M \subseteq \partial M$. Нетрудно также показать, что $M \cup \tilde{\partial}M = M^{00} \cup \tilde{\partial}M = \bar{M}$, т. е. «относительное замыкание» множества совпадает с его замыканием. — Упражнение 1.19.

Теорема (о пересечении выпуклого множества лучом). Пусть V — выпуклое множество, $a \in V^{00}$, $b \in \tilde{\partial}V$. Тогда точки $(1-\lambda)a + \lambda b$ принадлежат V^{00} при $\lambda \in [0, 1]$ и являются внешними точками множества V при $\lambda > 1$.

Другими словами, на луче $L^+ = \{(1-\lambda)a + \lambda b, \lambda \geq 0\}$, идущем из относительно внутренней точки a через относительно граничную точку b , имеется единственная относительно граничная точка — точка b ; все точки луча, лежащие между a и b , — относительно внутренние для V , все точки луча, лежащие «за b », — внешние точки множества V .

Доказательство. Пусть сначала $\lambda \in (0, 1)$. Предположим, что точка $(1-\lambda)a + \lambda b = c \in V^{00}$. Тогда существует такая последовательность $\{z_k\}_{k=1}^\infty$, что $z_k \in \text{aff } V$, $z_k \in V$ и $z_k \rightarrow c$ при $k \rightarrow \infty$. Вместе с тем, так как $b \in \tilde{\partial}V$, существует такая последовательность $\{b_k\}_{k=1}^\infty$, что $b_k \in V$, $b_k \rightarrow b$ при $k \rightarrow \infty$. Положим $a_k = \frac{1}{1-\lambda}(z_k - \lambda b_k)$, $k \in N$ (т. е. $z_k = (1-\lambda)a_k + \lambda b_k$). При любом k точка a_k лежит на прямой, проходящей через z_k и b_k , и, следовательно, $a_k \in \text{aff } V$. Поскольку множество V выпуклое, то из условий $b_k \in V$, $z_k \in V$, $z_k = (1-\lambda)a_k + \lambda b_k$ и $\lambda \in (0, 1)$ следует, что $a_k \in V$. Кроме того,

$$\|a - a_k\| = \left\| \frac{1}{1-\lambda}(c - \lambda b) - \frac{1}{1-\lambda}(z_k - \lambda b_k) \right\| \leq \frac{1}{1-\lambda} \|c - z_k\| + \frac{\lambda}{1-\lambda} \|b - b_k\| \rightarrow 0$$

при $k \rightarrow \infty$. Таким образом, относительно внутренняя точка множества V оказалась пределом последовательности точек $\{a_k\}_{k=1}^\infty$, принадлежащих $\text{aff } V$ и не принадлежащих V , что невозможно.

Покажем теперь, что $c = (1-\lambda)a + \lambda b$ — внешняя точка множества V , если $\lambda > 1$. В этом случае $b = (1-1/\lambda)a +$

$+(1/\lambda)c$, $1/\lambda \in (0, 1)$ и в силу первой части настоящей теоремы c не может быть ни относительно внутренней точкой множества V (V^{00} — множество выпуклое), ни относительно граничной точкой. ■

3.8. Прямыми следствиями доказанных выше теорем являются следующие утверждения.

Теорема 1. Пусть V — выпуклое множество. Тогда $\tilde{\partial}V^{00} = = \partial\tilde{V} = \tilde{\partial}V$.

Доказательство. Утверждение очевидно, если $V = \emptyset$. Пусть $V \neq \emptyset$. Тогда и $V^{00} \neq \emptyset$.

Очевидно, что $\partial\tilde{V} \subset \tilde{\partial}V$ (это включение выполняется для любого множества V). Пусть $v \in \tilde{\partial}V$. Возьмем точку $v_0 \in V^{00}$ и проведем луч $L^+ = \{(1-\lambda)v_0 + \lambda v, \lambda \geq 0\}$ из v_0 через v . По теореме о пересечении выпуклого множества лучом при $\lambda \in [0, 1)$ точки луча L^+ принадлежат $V^{00} \subset V \subset \tilde{V}$, при $\lambda > 1$ точки луча L^+ — внешние для V . Нетрудно проверить, что множество внешних точек любого множества совпадает с множеством внешних точек его замыкания. — Упражнение 1.9. Поэтому при $\lambda > 1$ точки луча L^+ — внешние для V . Таким образом, в любой относительной окрестности точки v есть точки, принадлежащие $V^{00} \subset \tilde{V}$, и точки, не принадлежащие $\tilde{V}$ и тем более не принадлежащие V^{00} . Следовательно, $v \in \partial\tilde{V}$, $v \in \tilde{\partial}V^{00}$, так что $\tilde{\partial}V \subset \partial\tilde{V}$ и $\tilde{\partial}V \subset \tilde{\partial}V^{00}$. Остается проверить включение $\tilde{\partial}V^{00} \subset \tilde{\partial}V$.

Пусть теперь $v \in \tilde{\partial}V^{00}$, $v_0 \in V^{00}$. Применяя снова теорему о пересечении выпуклого множества лучом, получаем, что в любой окрестности v есть точки из V и точки, внешние для V^{00} . Но точка, внешняя для V^{00} , не может принадлежать V , так как уже установлено, что $\tilde{\partial}V \subset \tilde{\partial}V^{00}$, и, следовательно, $V \subset V^{00} \cup \tilde{\partial}V \subset \subset V^{00} \cup \tilde{\partial}V^{00} = \tilde{V}^{00}$. Поэтому в любой окрестности точки v есть точки из V и точки, не принадлежащие V , т. е. $v \in \partial\tilde{V}$, а значит, $\tilde{\partial}V^{00} \subset \partial\tilde{V}$. ■

Теорема 2. Пусть V — выпуклое множество. Тогда $\overline{V}^{00} = \tilde{V}$.

Доказательство. Мы уже отмечали, что $\tilde{V} = V^{00} \cup \tilde{\partial}V$. Но $\tilde{\partial}V = \tilde{\partial}V^{00}$. Поэтому $\tilde{V} = V^{00} \cup \tilde{\partial}V^{00} = \tilde{V}^{00}$. ■

Теорема 3. Пусть V — выпуклое множество. Тогда $V^{00} = = (\tilde{V})^{00}$.

Доказательство. Так как множества внешних точек V и $\tilde{V}$ совпадают и $\tilde{\partial}V = \tilde{\partial}\tilde{V}$, то и $V^{00} = (\tilde{V})^{00}$. ■

Замечание. Все теоремы настоящего пункта можно формулировать для случая $\dim V = n$, опуская везде слово «относительно». Это приведет к уменьшению общности, но никакого «выигрыша» за счет упрощения доказательств мы не получим.

3.9. Относительно открытые выпуклые множества. Множество V называется *относительно открытым*, если все его точки относительно внутренние, т. е. если $V = V^{00}$. (В частности,

пустое множество и множество, состоящее из единственной точки (см. замечание 1 п. 3.4), являются относительно открытыми.)

Очевидно, что относительно открытое ядро выпуклого множества есть относительно открытое выпуклое множество.

Пересечение конечного числа относительно открытых выпуклых множеств есть относительно открытое выпуклое множество.

Доказательство будет проведено для случая двух множеств. Пусть V_1, V_2 — относительно открытые выпуклые множества, $V = V_1 \cap V_2$. Очевидно, что V — выпуклое множество. Пусть $v_0 \in V$. Тогда $v_0 \in V_1$, $v_0 \in V_2$ и существуют относительные окрестности $\mathcal{D}(v_0, r_1) \cap \text{aff } V_1 \subset V_1$, $\mathcal{D}(v_0, r_2) \cap \text{aff } V_2 \subset V_2$. Положим $r = \min \{r_1, r_2\}$. Тогда $\mathcal{D}(v_0, r) \cap \text{aff } V \subset V$, так как $\text{aff } V \subset \text{aff } V_1 \cap \text{aff } V_2$. ■

Всякое аффинное множество в R^n является относительно открытым выпуклым множеством. Поэтому пересечение относительно открытого выпуклого множества с аффинным множеством есть относительно открытое выпуклое множество.

Если V_1 и V_2 — относительно открытые выпуклые множества, то $V_1 + V_2$ — относительно открытое выпуклое множество.

Доказательство. Естественно предполагать, что $V_1 \neq \emptyset$ и $V_2 \neq \emptyset$. Положим $V = V_1 + V_2$. Множество V — выпуклое (см. п. 3.1) и $V \neq \emptyset$. Пусть $v \in V$. Покажем, что $v \in V^{00}$.

Предположим, что $v \notin V^{00}$. Тогда $v \in \tilde{V}$. Так как $V \neq \emptyset$, то существует точка $v_0 \in V^{00}$ (см. п. 3.6). Покажем, что на луче $L^+ = \{(1 - \alpha)v_0 + \alpha v, \alpha \geq 0\}$, идущем из точки $v_0 \in V^{00}$ через точку $v \in \tilde{V}$, есть точки, принадлежащие V при $\alpha > 1$, что противоречит теореме о пересечении выпуклого множества лучом (см. п. 3.7). Действительно, так как $V = V_1 + V_2$, то существуют такие v_1 и v_1^0 из V_1 , v_2 и v_2^0 из V_2 , что $v = v_1 + v_2$, $v_0 = v_1^0 + v_2^0$. Кроме того, V_1 и V_2 — относительно открытые выпуклые множества. Следовательно, найдется такое число $\lambda > 1$, что $(1 - \alpha)v_1^0 + \alpha v_1 \in V_1$, $(1 - \alpha)v_2^0 + \alpha v_2 \in V_2$ при любом $\alpha \in [1, \lambda]$ (см. п. 3.5). Но тогда и $(1 - \alpha)v_1^0 + \alpha v_1 + (1 - \alpha)v_2^0 + \alpha v_2 \in V$ при $\alpha \in [1, \lambda]$. ■

Если V — относительно открытое выпуклое множество и $\lambda \in R$, то λV — относительно открытое выпуклое множество. — Упражнение 1.20.

Если V_1 и V_2 — относительно открытые выпуклые множества, α_1 и α_2 из R , то $\alpha_1 V_1 + \alpha_2 V_2$ (в частности, $V_1 - V_2$) — относительно открытое выпуклое множество. — Упражнение 1.21.

§ 4. Теорема отделимости

4.1. Отделимость. Мы уже отмечали разделяющее свойство гиперплоскости: всякая гиперплоскость $H = \{x \in R^n : ax = \beta, a \in R^n, a \neq 0, \beta \in R\}$ делит пространство R^n на два полупространства Π^+ и Π^- , так что $\Pi^+ \cap \Pi^- = \emptyset$, $\Pi^+ \cup \Pi^- \cup H = R^n$

(см. п. 1.7). Каждое из полупространств Π^+ и Π^- является открытым выпуклым множеством размерности n . Обозначение полупространства (Π^+ или Π^-) зависит от записи уравнения гиперплоскости: если гиперплоскость H задана уравнением $ax = \beta$, то $\Pi^+ = \{x \in R^n: ax > \beta\}$, а $\Pi^- = \{x \in R^n: ax < \beta\}$; если задать ту же гиперплоскость уравнением $(-a)x = -\beta$, то первое из этих полупространств будет обозначаться Π^- , второе — Π^+ . Иногда одно из этих полупространств будем обозначать через Π_n , или просто Π . Замкнутые полупространства обозначим соответственно через $\bar{\Pi}^+$, $\bar{\Pi}^-$, $\bar{\Pi}_n$ или $\bar{\Pi}$. Ясно, что $\bar{\Pi}^+$ и $\bar{\Pi}^-$ — замкнутые выпуклые множества размерности n и $\bar{\Pi}^+ \cup \bar{\Pi}^- = R^n$, $\bar{\Pi}^+ \cap \bar{\Pi}^- = H$.

Пусть S и T — подмножества пространства R^n . Будем говорить, что S и T *отделимы*, если существует такая гиперплоскость $H \subset R^n$, что S содержится в одном из замкнутых полупространств, определяемых этой гиперплоскостью, а T — в другом. Или, что то же, *множества S и T отделимы, если существуют такие вектор $a \in R^n$, $a \neq 0$, и число $\beta \in R$, что $as \geq \beta$ для любого $s \in S$ и $at \leq \beta$ для любого $t \in T$.*

Множества S и T из R^n называются *строго отделимыми*, если существует такая гиперплоскость H , что S содержится в одном из полупространств, определяемых H , а T — в другом, т. е. S и T *строго отделимы, если существуют такие вектор $a \in R^n$, $a \neq 0$, и число $\beta \in R$, что $as > \beta$ для любого $s \in S$ и $at < \beta$ для любого $t \in T$.* При этом говорят, что гиперплоскость H *разделяет* (или *строго разделяет*) множества S и T (гиперплоскость H называется *разделяющей*).

4.2. Следующая теорема является ослабленным вариантом известной теоремы Хана — Банаха.

Теорема. Пусть V — относительно открытое выпуклое множество в R^n , H — аффинное множество, $0 \leq \dim H \leq n-1$, и $V \cap H = \emptyset$. Тогда существует такая гиперплоскость A (т. е. аффинное множество размерности $n-1$), что $A \supset H$, $A \cap V = \emptyset$.

Доказательство. Если $\dim H = n-1$, то $A = H$. Пусть $\dim H = r < n-1$. Докажем, что в этом случае можно построить такое аффинное множество H_1 , что $H_1 \supset H$, $H_1 \cap V = \emptyset$ и $\dim H_1 = r+1$. Если $r+1 = n-1$, то $A = H_1$. В противном случае построим аффинное множество $H_2 \supset H_1$, $H_2 \cap V = \emptyset$, $\dim H_2 = r+2$ и т. д.

Пусть $a \in H$. Тогда $H - a = T$ — подпространство в R^n , $V - a = W$ — относительно открытое выпуклое множество в R^n , $T \cap W = \emptyset$ и $\dim T \leq n-2$. Пусть $T^\perp$ — ортогональное дополнение подпространства T в R^n . Так как $\dim T \leq n-2$, то $\dim T^\perp \geq 2$. Пусть P — какое-нибудь двумерное подпространство пространства $T^\perp$. Пространство P можем отождествить с двумерной плоскостью R^2 . Поскольку W и T — непересекающиеся

относительно открытые выпуклые множества, то $P \cap (W - T)$ — относительно открытое выпуклое множество на плоскости P , не содержащее нуля. Предлагаем читателю самостоятельно убедиться, что на плоскости P существует одномерное подпространство L (проходящая через нуль прямая), не пересекающаяся с $P \cap (W - T)$. — Упражнение 1.22. Тогда $T \oplus L = (r + 1)$ -мерное подпространство в R^n и $(T \oplus L) \cap W = \emptyset$. В самом деле, если $w \in (T \oplus L) \cap W$, то существуют такие $t \in T$, $l \in L$, что $w = t + l$. Но тогда $l = w - t \in P \cap (W - T)$, что невозможно. Положим $H_1 = T \oplus L + a$. Тогда H_1 — аффинное множество размерности $r + 1$ и $H_1 \cap V = (T \oplus L + a) \cap (W + a) = \emptyset$. ■

4.3. Теорема отделимости.

Лемма 1. Пусть V и W — непересекающиеся относительно открытые выпуклые множества. Тогда существует такой вектор $a \in R^n$, $a \neq 0$, что $av < aw$ для любых $v \in V$, $w \in W$.

Доказательство. Так как V и W — относительно открытые выпуклые множества и $V \cap W = \emptyset$, то $V - W$ — относительно открытое выпуклое множество (см. упражнение 1.21), не содержащее нуля. Поэтому $V - W$ не пересекается с аффинным множеством $H = \{0\}$ и по теореме п. 4.2 существует гиперплоскость $\hat{A}$, содержащая H и не пересекающаяся с $V - W$. Множество $V - W$ содержится в одном из полупространств, определяемых гиперплоскостью $\hat{A}$, так как если бы существовали две точки этого множества, принадлежащие разным полупространствам, то отрезок, их соединяющий (он должен принадлежать $V - W$, поскольку $V - W$ — выпуклое множество), пересекался бы с $\hat{A}$. Выберем нормальный вектор $a \neq 0$ гиперплоскости $\hat{A}$ так, чтобы $V - W \subset \Pi^-$. Поскольку $H = \{0\} \subset \hat{A}$, то уравнение гиперплоскости $\hat{A}$ имеет вид $ax = 0$ и $a(v - w) < 0$ для любых $v \in V$, $w \in W$, или, что то же, $av < aw$ для любых $v \in V$, $w \in W$. ■

Лемма 2. Пусть V и W — непустые множества в R^n и $a \neq 0$ — такой вектор из R^n , что $av < aw$ для любых $v \in V$, $w \in W$. Тогда существует такое число $\beta \in R$, что гиперплоскость $H = \{x \in R^n : ax = \beta\}$ разделяет множества V и W и не пересекается по крайней мере с одним из этих множеств.

Доказательство. Положим $\beta = \sup \{av, v \in V\}$. Так как множество чисел $\{av, v \in V\}$ ограничено сверху любым из чисел множества $\{aw, w \in W\}$, то $\beta < +\infty$. Положим $H = \{x \in R^n : ax = \beta\}$. Тогда $av \leq \beta$ для любого $v \in V$ и $\beta \leq aw$ для $w \in W$, поскольку при любом $w \in W$ число aw ограничивает сверху множество $\{av, v \in V\}$. Кроме того, если $\beta \in \{av, v \in V\}$, то $\beta < aw$ для любого $w \in W$. В противном случае $av < \beta$ для любого $v \in V$. ■

Теорема. Непустые выпуклые множества V и W , относительно открытые ядра которых V^{00} и W^{00} не пересекаются, отделимы. При этом разделяющую гиперплоскость можно выбрать так,

чтобы она не пересекалась по крайней мере с одним из множеств V^{00} и W^{00} .

Доказательство. Так как V и W — непустые выпуклые множества, то V^{00} и W^{00} — непустые (см. п. 3.7) выпуклые (см. п. 3.5) множества. В силу леммы 1 существует такой вектор $a \in R^n$, $a \neq 0$, что $av < aw$ для любых $v \in V^{00}$, $w \in W^{00}$. В силу леммы 2 существует такое число $\beta \in R$, что гиперплоскость $H = \{x \in R^n: ax = \beta\}$ разделяет множества V^{00} и W^{00} и не пересекается по крайней мере с одним из этих множеств. Таким образом, $V^{00} \subset \bar{\Pi}^-$, $W^{00} \subset \bar{\Pi}^+$, и, следовательно, $\overline{V^{00}} \subset \bar{\Pi}^-$, $\overline{W^{00}} \subset \bar{\Pi}^+$ (см. упражнение 1.7). Но в силу теоремы о замыкании выпуклого множества (см. п. 3.8) $\overline{V^{00}} = \bar{V}$, $\overline{W^{00}} = \bar{W}$, так что $V \subset \bar{V} \subset \bar{\Pi}^-$, $W \subset \bar{W} \subset \bar{\Pi}^+$. ■

4.4. Существование опорной гиперплоскости. Пусть V — выпуклое множество в R^n и $v_0 \in \partial V$. Гиперплоскость H называется *опорной к множеству V в точке v_0* , если $v_0 \in H$ и $V \subset \bar{\Pi}_H$, где $\bar{\Pi}_H$ — одно из замкнутых полупространств, определенных гиперплоскостью H .

Теорема (Минковского). Пусть V — выпуклое множество в R^n , $\dim V > 0$ и $v_0 \in \partial V$. Тогда существует гиперплоскость H , опорная к V в точке v_0 и не пересекающаяся с относительно открытым ядром V^{00} множества V .

Замечание. Если $\dim V = 0$, т. е. $V = \{v_0\}$ состоит из единственной точки v_0 , то всякая гиперплоскость, проходящая через точку v_0 , будет опорной к множеству V в этой точке.

Доказательство теоремы. Так как $\dim V > 0$ и $v_0 \in \partial V$, то V^{00} и $W = \{v_0\}$ — непустые непересекающиеся относительно открытые выпуклые множества. В силу теоремы п. 4.3 существует гиперплоскость H , разделяющая множества V^{00} и W и не пересекающаяся по крайней мере с одним из этих множеств. Пусть $V^{00} \subset \bar{\Pi}^-$, $v_0 \in \bar{\Pi}^+$. Так как $\overline{V^{00}} = \bar{V}$ (см. теорему 2 п. 3.8), то $V \subset \bar{V} \subset \bar{\Pi}^-$. Кроме того, $v_0 \in \bar{V} \subset \bar{\Pi}^-$, так что $v_0 \in H$. Поэтому $V^{00} \cap H = \emptyset$. ■

4.5. Лемма об осуществлении расстояния. Пусть A и B — непустые замкнутые множества в R^n , по крайней мере одно из которых компактно. Тогда существуют такие точки $\bar{a} \in A$, $\bar{b} \in B$, что $\rho(\bar{a}, \bar{b}) = \rho(A, B)$ (см. п. 2.5).

Доказательство. Для определенности предположим, что A компактно. Так как $\rho(A, B) = \inf \{\rho(a, b), a \in A, b \in B\}$, то для любого натурального k найдутся такие элементы $a_k \in A$, $b_k \in B$, что $\rho(a_k, b_k) < \rho(A, B) + 1/k$. Рассмотрим последовательности $\{a_k\}_{k=1}^\infty$ и $\{b_k\}_{k=1}^\infty$. Так как $\{a_k\}_{k=1}^\infty$ есть последовательность элементов компактного множества A , то из нее можно выделить подпоследовательность $\{a_k\}_{k \in K}$, сходящуюся к пределу $\bar{a} \in A$. Множество значений последовательности $\{b_k\}_{k \in K}$

ограничено, так как $\|b_k\| \leq \|b_k - a_k\| + \|a_k\| < \rho(A, B) + 1 + d$, где через d обозначен радиус шара $\mathcal{D}(0, d)$, содержащего ограниченное множество A . Положим $r = \rho(A, B) + 1 + d$. Тогда $\{b_k\}_{k \in K}$ есть последовательность точек компактного множества $B \cap \mathcal{D}(0, r)$. Поэтому из последовательности $\{b_k\}_{k \in K}$ можно выделить подпоследовательность $\{b_k\}_{k \in K_1}$, $K_1 \subset K$, сходящуюся к элементу $\bar{b}$ множества $B \cap \overline{\mathcal{D}}(0, r) \subset B$. Подпоследовательность $\{a_k\}_{k \in K_1}$ последовательности $\{a_k\}_{k \in K}$ сходится к $\bar{a}$. При любом $k \in K_1$ имеем

$$\rho(A, B) \leq \rho(\bar{a}, \bar{b}) = \|\bar{a} - \bar{b}\| \leq \|\bar{a} - a_k\| + \|a_k - b_k\| + \|b_k - \bar{b}\| \leq \leq \rho(A, B) + \|\bar{a} - a_k\| + \|b_k - \bar{b}\| + 1/k.$$

Переходя в этом неравенстве к пределу при $k \rightarrow \infty$, $k \in K_1$, получаем $\rho(A, B) \leq \rho(\bar{a}, \bar{b}) \leq \rho(A, B)$, т. е. $\rho(A, B) = \rho(\bar{a}, \bar{b})$. ■

4.6. Теорема о строгой отделимости. *Непустые непересекающиеся замкнутые выпуклые множества V и W , по крайней мере одно из которых ограничено, строго отделимы.*

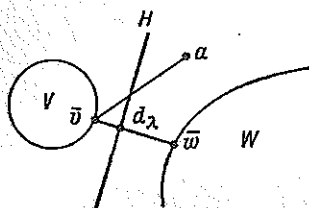


Рис. 4.

Доказательство. В силу предыдущей леммы найдутся такие точки $\bar{v} \in V$, $\bar{w} \in W$, что $\rho(\bar{v}, \bar{w}) = \rho(V, W)$. Поскольку $V \cap W = \emptyset$, имеем $\bar{v} \neq \bar{w}$, так что $\bar{w} - \bar{v} \neq 0$. Положим $d_\lambda = (1 - \lambda)\bar{v} + \lambda\bar{w}$ и покажем, что при любом $\lambda \in (0, 1)$ гиперплоскость $(\bar{w} - \bar{v})x = (\bar{w} - \bar{v})d_\lambda$ строго разделяет множества V и W (рис. 4), а именно

$(\bar{w} - \bar{v})v < (\bar{w} - \bar{v})d_\lambda$ для любого $v \in V$ и $(\bar{w} - \bar{v})w > (\bar{w} - \bar{v})d_\lambda$ для любого $w \in W$. Докажем здесь только первое из этих неравенств. Второе неравенство доказывается аналогично. — **Упражнение 1.23.**

Итак, проверим, что $(\bar{w} - \bar{v})v < (\bar{w} - \bar{v})d_\lambda$ при любом $v \in V$. Для этого достаточно установить, что если $a \in R^n$ и $(\bar{w} - \bar{v})a \geq (\bar{w} - \bar{v})d_\lambda$, то $a \in V$.

Пусть $\alpha \in [0, 1]$. Положим

$$\begin{aligned} \varphi(\alpha) &= (\rho((1 - \alpha)\bar{v} + \alpha a, \bar{w}))^2 = \|(1 - \alpha)\bar{v} + \alpha a - \bar{w}\|^2 = \\ &= \alpha^2 \|a - \bar{v}\|^2 + 2\alpha(\bar{w} - \bar{v})(\bar{v} - a) + \|\bar{v} - \bar{w}\|^2. \end{aligned}$$

Покажем, что $\varphi'(0) < 0$, если $(\bar{w} - \bar{v})a \geq (\bar{w} - \bar{v})d_\lambda$. Действительно,

$$\begin{aligned} \frac{1}{2}\varphi'(0) &= (\bar{w} - \bar{v})(\bar{v} - a) = (\bar{w} - \bar{v})\bar{v} - (\bar{w} - \bar{v})a \leq \\ &\leq (\bar{w} - \bar{v})\bar{v} - (\bar{w} - \bar{v})d_\lambda = (\bar{w} - \bar{v})(\bar{v} - (1 - \lambda)\bar{v} - \lambda\bar{w}) = \\ &= -\lambda\|\bar{w} - \bar{v}\|^2 < 0. \end{aligned}$$

Так как $\varphi'(0) < 0$, то φ — убывающая функция в некоторой окрестности точки 0 и, следовательно, найдется такое $\alpha_0 \in (0, 1)$, что $\varphi(\alpha_0) < \varphi(0)$, т. е.

$$\rho((1 - \alpha_0)\bar{v} - \alpha_0 a, \bar{w}) < \rho(\bar{v}, \bar{w}) = \rho(V, W).$$

Значит, $(1 - \alpha_0)\bar{v} + \alpha_0 a \in V$ и поскольку V — выпуклое множество и $\bar{v} \in V$, то $a \in V$. ■

Замечание. Для доказательства теоремы мы построили разделяющую гиперплоскость с нормальным вектором $\bar{w} - \bar{v}$, проходящую через относительно внутреннюю точку d_λ отрезка $[\bar{v}, \bar{w}]$. Если провести гиперплоскость с тем же нормальным вектором через точку $\bar{v}$, то она будет опорной к множеству V в точке $\bar{v}$, а $W \subset \Pi^+$.

§ 5. Многогранные выпуклые множества и выпуклые многогранники

5.1. Экстремальные точки выпуклых множеств. Точка v_0 выпуклого множества V называется его *экстремальной точкой* (*угловой, крайней*), если она не является относительно внутренней точкой никакого отрезка, концы которого принадлежат множеству V . Другими словами, точка $v_0 \in V$ — экстремальная, если из условий $v_1 \in V$, $v_2 \in V$, $v_0 = (1 - \lambda)v_1 + \lambda v_2$, $\lambda \in (0, 1)$, следует, что $v_1 = v_0 = v_2$.

Для сектора, изображенного на рис. 5, экстремальными являются все точки дуги AB и точка O . Точка c принадлежит ∂V , но не является экстремальной.

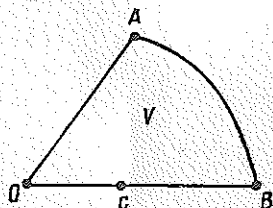


Рис. 5.

Нетрудно показать, что точка $v_0 \in V$ является экстремальной, если она не является серединой никакого отрезка $[v_1, v_2]$, $v_1 \neq v_2$, $v_1 \in V$, $v_2 \in V$. — **Упражнение 1.24.**

Множество всех экстремальных точек выпуклого множества V называется его *экстремальным множеством* и обозначается через $\text{ext } V$.

Экстремальное множество непустого компактного выпуклого множества непусто.

Доказательство. Функция $f(v) = \|v\|$ непрерывна на компактном множестве V , следовательно, существует такая точка $v^* \in V$, что $\|v^*\| = \max_{v \in V} \|v\|$. Точка v^* — экстремальная точка мно-

жества V . Действительно, если $v^* = \frac{1}{2}(v_1 + v_2)$, $v_1 \neq v_2$, $v_1 \in V$, $v_2 \in V$, и точки O , v_1 и v_2 не лежат на одной прямой, то в тре-

угольнике с вершинами в точках O , v_1 и v_2 (рис. 6) медиана Ov^* не меньше каждой из сторон Ov_1 и Ov_2 , что невозможно; очевидно невозможно и расположение точек O , v_1 и v_2 на одной прямой. ■

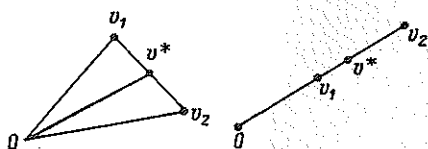


Рис. 6.

5.2. Теорема о представлении компактного выпуклого множества. *Непустое компактное выпуклое множество $V \subset R^n$ совпадает с выпуклой оболочкой $v(\text{ext } V)$ своего экстремального множества $\text{ext } V$.*

Доказательству теоремы предположим две леммы.

Лемма 1. *Пусть V — выпуклое множество и H — гиперплоскость, опорная к V в некоторой точке $v_0 \in \partial V$. Тогда $\text{ext}(V \cap H) = H \cap \text{ext } V$.*

Доказательство. Включение $H \cap \text{ext } V \subset \text{ext}(V \cap H)$ очевидно. Пусть $v \in \text{ext}(V \cap H)$ и $v \notin \text{ext } V$. Тогда существуют такие точки $v_1 \in V$, $v_2 \in V$, $v_1 \neq v_2$, что $v = \frac{1}{2}(v_1 + v_2)$. Пусть $ax = \beta$ — уравнение гиперплоскости H и $V \subset \overline{\Pi^-}$. Тогда $av_1 \leq \beta$, $av_2 \leq \beta$, и так как $v \in H$ и $v \in V$, то $\beta = av = \frac{1}{2}av_1 + \frac{1}{2}av_2 \leq \beta$. Следовательно, $av_1 = \beta$, $av_2 = \beta$, т. е. $v_1 \in H$, $v_2 \in H$ и $v_1 \in V \cap H$, $v_2 \in V \cap H$. Но это противоречит тому, что $v \in \text{ext}(V \cap H)$. ■

Лемма 2. *Пусть V — выпуклое множество и H — такая гиперплоскость в R^n , что $V \cap H \neq \emptyset$ и V не содержится в H . Тогда $\dim(V \cap H) < \dim V$.*

Доказательство. Аффинная оболочка множества V есть аффинное множество A , размерность которого равна $\dim V$. Аффинная оболочка множества $V \cap H$ содержится в непустом аффинном множестве $A \cap H$, и так как V не содержится в H , то $A \cap H \subset A$ и $A \cap H \neq A$. Поэтому $\dim(V \cap H) \leq \dim(A \cap H) < \dim A = \dim V$. ■

Доказательство теоремы. Так как $\text{ext } V \subset V$, то $v(\text{ext } V) \subset V$. Докажем, что $V \subset v(\text{ext } V)$. Доказательство проводится индукцией по размерности $r = \dim V$. Утверждение теоремы очевидно, если $\dim V = 0$. Предположим, что оно справедливо для компактных выпуклых множеств размерности $r \in 0:k$ и V — компактное выпуклое множество в R^n размерности $k+1$.

Пусть $v_0 \in \partial V$. Тогда $v_0 \in V$ (V замкнуто). По теореме об опорной гиперплоскости (см. п. 4.4) существует гиперплоскость

H , опорная к V в точке v_0 и не пересекающаяся с непустым (см. п. 3.6) множеством $V^{00} \subset V$. Гиперплоскость H — замкнутое выпуклое множество. Поэтому $V \cap H$ — компактное выпуклое множество. В силу леммы 2 $\dim(V \cap H) < \dim V$. Итак, $0 \leq \dim(V \cap H) \leq k$ и по предположению индукции $V \cap H = v(\text{ext}(V \cap H))$. Но в силу леммы 1 $\text{ext}(V \cap H) \subset \text{ext } V$. Поэтому $v_0 \in v(\text{ext } V)$, и, следовательно, $\bar{\partial} V \subset v(\text{ext } V)$.

Пусть $v_0 \in V^{00}$. Обозначим через A аффинную оболочку множества V . Тогда $\dim A \geq 1$. Возьмем вектор $a \neq 0$, принадлежащий подпространству $A - v_0$, и рассмотрим прямую $L = \{x \in R^n : (1 - \alpha)v_0 + \alpha a, \alpha \in R\}$. Тогда $L \subset A$, $L \cap V$ — непустое компактное выпуклое множество на прямой L , т. е. $L \cap V = [v_1, v_2]$ — отрезок, содержащий точку v_0 , а v_1 и v_2 — относительно граничные точки множества V (рис. 7). Мы уже доказали, что $\bar{\partial} V \subset v(\text{ext } V)$. Поэтому $[v_1, v_2] \subset v(\text{ext } V)$, и, следовательно, $v_0 \in v(\text{ext } V)$. ■

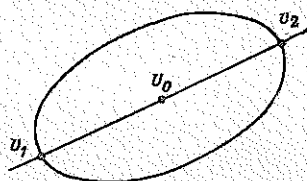


Рис. 7.

5.3. Выпуклые многогранники. Следствием теоремы предыдущего пункта является утверждение: *непустое компактное выпуклое множество, имеющее конечное число экстремальных точек, есть выпуклый многогранник*.

Действительно, экстремальное множество непустого компактного множества V непусто (см. п. 5.1), и если $\text{ext } V = \{a_0, \dots, a_m\}$, то $V = v(\{a_0, \dots, a_m\})$. ■

Таким образом (см. также п. 3.3), выпуклые многогранники суть непустые компактные выпуклые множества, имеющие конечное число экстремальных точек.

Экстремальные точки выпуклого многогранника называются его *вершинами*. Нетрудно показать, что *все вершины выпуклого многогранника $V = v(\{a_0, \dots, a_m\})$ содержатся в множестве $\{a_0, \dots, a_m\}$* . — Упражнение 1.25.

Полезным является следующее утверждение: *сумма $V_1 + V_2$ выпуклых многогранников есть выпуклый многогранник*. Действительно, $V_1 + V_2$ — непустое компактное выпуклое множество и, как легко видеть, *экстремальной точкой множества $V_1 + V_2$ может быть лишь сумма вершин выпуклых многогранников V_1 и V_2* . — Упражнение 1.26.

5.4. Многогранные выпуклые множества. Непустое пересечение конечного числа замкнутых полупространств называется *многогранным выпуклым множеством*. Пусть $\{H_1, H_2, \dots, H_m\}$ — конечное множество гиперплоскостей в R^n и $\{\bar{\Pi}_1, \bar{\Pi}_2, \dots, \bar{\Pi}_m\}$ — множество замкнутых полупространств, определяемых гиперплоскостями H_i , $i \in 1:m$. Множество $V = \bigcap_{i=1}^m \bar{\Pi}_i$ есть мно-

тогранное выпуклое множество, если $V \neq \emptyset$ (рис. 8). Очевидно, что V — замкнутое выпуклое множество в R^n .

Всякое аффинное множество в R^n есть многогранное выпуклое множество (в частности, само R^n удобно считать пересечением пустого множества замкнутых полупространств). Действительно, всякое аффинное множество в R^n можно представить как пересечение конечного множества гиперплоскостей, а любая гиперплоскость в свою очередь есть пересечение двух замкнутых полупространств, которые она определяет. ■

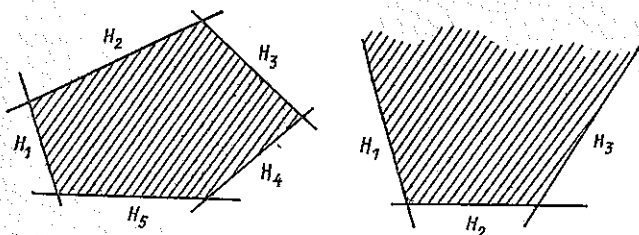


Рис. 8.

Гиперплоскости $H_i, i \in 1:m$, участвующие в определении многогранного выпуклого множества V , будем называть его *ограничивающими гиперплоскостями*.

5.5. Экстремальные точки многогранного выпуклого множества. Здесь будут выяснены условия, при которых точка многогранного выпуклого множества V является его экстремальной точкой.

Пусть $\{H_1, \dots, H_m\}$ — множество гиперплоскостей в R^n , $\{\bar{\Pi}_1, \bar{\Pi}_2, \dots, \bar{\Pi}_m\}$ — множество замкнутых полупространств, определяемых гиперплоскостями $H_i, i \in 1:m$, и $V = \bigcap_{i=1}^m \bar{\Pi}_i$ — многогранное выпуклое множество. Пусть для каждого $i \in 1:m$ уравнение $a_i x = b_i$ есть уравнение гиперплоскости H_i . Обозначим через I некоторое подмножество множества $1:m$. Множество гиперплоскостей $\{H_i, i \in I\}$ назовем *независимым*, если система векторов $\{a_i, i \in I\}$ линейно независима. Пусть $v \in V$. Обозначим через $I(v)$ множество индексов ограничивающих гиперплоскостей, проходящих через точку v .

Теорема. Пусть V — многогранное выпуклое множество в R^n . Точка $v_0 \in V$ в том и только том случае принадлежит $\text{ext } V$, если множество ограничивающих гиперплоскостей $\{H_i, i \in I(v_0)\}$ содержит подмножество из n независимых гиперплоскостей.

Доказательство. Необходимость. Предположим, что максимальное число независимых гиперплоскостей в множестве $\{H_i, i \in I(v_0)\}$ равно $r < n$. Тогда размерность аффинного множества $\bigcap_{i \in I(v_0)} H_i$ равна $n - r \geq 1$. Пусть $i \in I(v_0)$. Обозначим

через r_i расстояние от точки v_0 до гиперплоскости H_i . В силу леммы п. 4.5 расстояние осуществляется, и так как $v_0 \in H_i$ для $i \in I(v_0)$, то $r_i > 0$. Поэтому при каждом $i \in I(v_0)$ шар $\mathcal{D}(v_0, r_i) \subset \bar{\Pi}_i$.

Положим $r = \min \{r_i, i \in I(v_0)\}$. Тогда $r > 0$ и $\mathcal{D}(v_0, r) \subset \cap_{i \in I(v_0)} \bar{\Pi}_i$. Таким образом, относительная окрестность $(\cap_{i \in I(v_0)} H_i) \cap \mathcal{D}(v_0, r)$ точки v_0 в аффинном множестве $\cap_{i \in I(v_0)} H_i$ размерности $n - r \geq 1$ содержится в многогранном выпуклом множестве $V = (\cap_{i \in I(v_0)} \bar{\Pi}_i) \cap (\cap_{i \in I(v_0)} \bar{\Pi}_i)$, и, следовательно, $v_0 \in \text{ext } V$. ■

Достаточность. Пусть множество ограничивающих гиперплоскостей $\{H_i, i \in I(v_0)\}$ содержит n независимых гиперплоскостей. Покажем, что $v_0 \in \text{ext } V$.

Предположим, что $v_0 = 1/2(v_1 + v_2)$, $v_1 \in V$, $v_2 \in V$. Для каждого $i \in I(v_0)$ имеем $a_i v_0 = \beta_i$, $a_i v_1 \leq \beta_i$, $a_i v_2 \leq \beta_i$ (для определенности предполагаем, что $\bar{\Pi}_i = \bar{\Pi}_i^-$). Поэтому $\beta_i = a_i v_0 = 1/2 a_i v_1 + 1/2 a_i v_2 \leq \beta_i$, и, следовательно, $v_1 \in H_i$, $v_2 \in H_i$ при любом $i \in I(v_0)$. Но аффинное множество $\cap_{i \in I(v_0)} H_i$ имеет размерность $n - n = 0$, т. е. состоит из единственной точки. Значит, $v_1 = v_2 = v_0 \in \text{ext } V$. ■

Следствие. Множество $\text{ext } V$ экстремальных точек многогранного выпуклого множества V конечно.

В самом деле, точка $v_0 \in V$ может быть экстремальной только тогда, когда множество $\{H_i, i \in I(v_0)\}$ ограничивающих гиперплоскостей, проходящих через эту точку, содержит n независимых. Но тогда пересечение $\cap_{i \in I(v_0)} H_i$ — аффинное множество размерности нуль. Следовательно, оно состоит из единственной точки v_0 . Остается заметить, что из конечного числа ограничивающих гиперплоскостей можно образовать лишь конечное число таких пересечений. ■

Пример. Пусть $A = (a[i, j], i \in 1:m, j \in 1:n) — m \times n$ матрица и $b \in R^m$. Обозначим через $a_i, i \in 1:m$, строки, а через $a^j, j \in 1:n$, — столбцы матрицы A . Пусть $\mathcal{D} \subset R^n$ — множество решений следующей системы уравнений и неравенств:

$$Ax = b, \quad x[j] = e_j x \geq 0, \quad j \in 1:n$$

(e_j — векторы канонического базиса пространства R^n (см. п. 1.1)). Если $x_0 \in \mathcal{D}$, то обозначим через $J(x_0)$ множество тех $j \in 1:n$, для которых $x_0[j] = 0$. Справедливо следующее утверждение.

Точка $x_0 \in \mathcal{D}$ в том и только том случае является экстремальной точкой множества $\mathcal{D}$, если множество столбцов a^j матрицы A с номерами $j \in 1:n \setminus J(x_0)$ линейно независимо.

Доказательство. Представим множество $\mathcal{D}$ как пересечение конечного множества замкнутых полупространств

$$\bar{\Pi}_i^+ = \{x \in R^n : a_i x \geq b[i]\}, \quad \bar{\Pi}_i^- = \{x \in R^n : a_i x \leq b[i]\}, \quad i \in 1:m,$$

$$\bar{\Pi}_j = \{x \in R^n : e_j x \geq 0\}, \quad j \in 1:n.$$

Точка x_0 принадлежит граничным гиперплоскостям

$$H_i = \{x \in R^n : a_i x = b[i]\}, \quad i \in 1:m,$$

$$\hat{H}_j = \{x \in R^n : e_j x = 0\}, \quad j \in J(x_0),$$

и в силу доказанной выше теоремы является экстремальной точкой множества $\mathcal{D}$ в том и только том случае, если в множестве $\{a_i, i \in 1:m; e_j, j \in J(x_0)\}$ содержится система из n линейно независимых векторов — базис пространства R^n . Система векторов $\{e_j, j \in J(x_0)\}$ линейно независима и, как нетрудно видеть, ее можно дополнить до базиса векторами $a_i, i \in 1:m$, тогда и только тогда, когда система векторов $\{a^j, j \in 1:n \setminus J(x_0)\}$ линейно независима. ■

5.6. В этом и следующем пунктах дается исчерпывающая структурная характеристика ограниченных многогранных выпуклых множеств.

Теорема. Ограниченное многогранное выпуклое множество V является выпуклым многогранником.

Доказательство. Ограниченное многогранное выпуклое множество есть компактное выпуклое множество. Поэтому $\text{ext } V \neq \emptyset$ (см. п. 5.1) и конечно (см. п. 5.5), а значит, V — выпуклый многогранник (см. п. 5.3). ■

5.7. Теорема. Выпуклый многогранник V есть многогранное выпуклое множество.

Докажем сначала следующую лемму.

Лемма. Пусть V — непустое замкнутое выпуклое множество и $\{H_\alpha, \alpha \in A\}$ — такое семейство гиперплоскостей, опорных к множеству V , что $H_\alpha \cap V^{\text{вн}} = \emptyset$ при любом $\alpha \in A$, и для любой точки $v \in \partial V$ найдется такое $\alpha \in A$, что $v \in H_\alpha$. Пусть $\{\bar{\Pi}_\alpha, \alpha \in A\}$ — такое семейство замкнутых полупространств, порожденных гиперплоскостями $H_\alpha, \alpha \in A$, что $V \subset \bar{\Pi}_\alpha$ для каждого $\alpha \in A$. Тогда $(\bigcap_{\alpha \in A} \bar{\Pi}_\alpha) \cap \text{aff } V = V$.

Доказательство. Из определения множества $\{\bar{\Pi}_\alpha, \alpha \in A\}$ следует, что $V \subset (\bigcap_{\alpha \in A} \bar{\Pi}_\alpha) \cap \text{aff } V$. Докажем обратное включение.

Пусть $v \in (\bigcap_{\alpha \in A} \bar{\Pi}_\alpha) \cap \text{aff } V$ и $v \notin V$. Так как множество $V^{\text{вн}} \neq \emptyset$ (см. п. 3.6), то существует точка $v_0 \in V^{\text{вн}}$. Рассмотрим отрезок $[v_0, v]$. Пересечение $[v_0, v] \cap V$ есть замкнутое выпуклое подмножество отрезка, содержащее точку v_0 . Следовательно, это пересечение есть отрезок $[v_0, \bar{v}]$, где $\bar{v} = (1 - \hat{\lambda})v_0 + \hat{\lambda}v$ и

$\bar{\lambda} \in (0, 1)$. Очевидно, что $\bar{v} \in \bar{\partial}V$. Тогда существует такое $\alpha_0 \in A$, что $\bar{v} \in H_{\alpha_0}$. Пусть $a_{\alpha_0}x = \beta_{\alpha_0}$ — уравнение гиперплоскости H_{α_0} и $\bar{P}_{\alpha_0} = \bar{P}_{\alpha_0}^-$. Из условий $a_{\alpha_0}((1 - \bar{\lambda})v_0 + \bar{\lambda}v) = \beta_{\alpha_0}$, $a_{\alpha_0}v_0 < \beta_{\alpha_0}$, $0 < \bar{\lambda} < 1$, следует $a_{\alpha_0}v > \beta_{\alpha_0}$, т. е. $v \notin \bar{P}_{\alpha_0}$, что невозможно, так как $v \in \bigcap_{\alpha \in A} \bar{P}_{\alpha}$. ■

Доказательство теоремы. Пусть $M = \{a_1, a_2, \dots, a_m\}$ и $V = v(M)$ — выпуклый многогранник. Множество $\text{aff } V$ есть пересечение конечного множества гиперплоскостей, а следовательно, пересечение конечного множества замкнутых полупространств (если H — гиперплоскость, то $H = \bar{P}_H^- \cap \bar{P}_H^+$). Поэтому достаточно построить конечное семейство гиперплоскостей, удовлетворяющих условиям леммы.

Возьмем какую-нибудь точку $v_1 \in \bar{\partial}V$. По теореме об опорной гиперплоскости (см. п. 4.4) существует такая гиперплоскость H_1 , опорная к V в точке v_1 , что $V^{00} \cap H_1 = \emptyset$. Положим $V_1 = V \cap H_1$. Тогда V_1 — компактное выпуклое множество и $\text{ext } V_1 = H_1 \cap \text{ext } V$ (см. лемму 1 п. 5.2). Так как $\text{ext } V \subset M$ (см. п. 5.3), то V_1 — выпуклый многогранник, причем $V_1 = v(M_1)$, $M_1 \subset M$.

Если $V_1 \supset \bar{\partial}V$, то процесс построения семейства гиперплоскостей окончен. Пусть существует $v_2 \in \bar{\partial}V$, $v_2 \notin H_1$. Построим такую гиперплоскость H_2 , опорную к V в точке v_2 , что $V^{00} \cap H_2 = \emptyset$. Положим $V_2 = V \cap H_2$. Тогда $V_2 = v(M_2)$ — выпуклый многогранник, $M_2 \subset M$ и $M_2 \neq M_1$ (иначе $v_2 \in V_1$).

Предположим, что уже построены гиперплоскости $H_1, H_2, \dots, H_k$, опорные к V в относительно граничных точках $v_1, v_2, \dots, v_k$ и не пересекающиеся с V^{00} . Пусть $V_1, V_2, \dots, V_k$ — такие выпуклые многогранники, что $V_i = V \cap H_i = v(M_i)$, $M_i \subset M$ и $M_i \neq M_j$ при $i \neq j$, $i \in 1:k$, $j \in 1:k$. Если $\bar{\partial}V \subset \bigcup_{i=1}^k V_i$, то процесс построения гиперплоскостей окончен. В противном случае существует точка $v_{k+1} \in \bar{\partial}V$, $v_{k+1} \notin \bigcup_{i=1}^k V_i$, и можно построить опорную гиперплоскость H_{k+1} к V в точке v_{k+1} , не пересекающуюся с V^{00} . Положим $V_{k+1} = V \cap H_{k+1}$. Тогда V_{k+1} — выпуклый многогранник, $V_{k+1} = v(M_{k+1})$, $M_{k+1} \subset M$, $M_{k+1} \neq M_i$, $i \in 1:k$. Этот процесс построения должен завершиться через конечное число шагов, так как конечное множество M содержит конечное число различных подмножеств M_k . ■

§ 6. Конусы. Неограниченные выпуклые множества

6.1. Конусы. Множество $K_a \subset R^n$ называется конусом с вершиной в точке $a \in R^n$, если $a \in K_a$ и вместе с любой своей точкой $x \neq a$ множество K_a содержит луч $\{(1 - \alpha)a + \alpha x, \alpha \geq 0\}$, идущий из a через x .

Пространство R^n с отмеченной точкой $a \in R^n$, его подпространство T с отмеченной точкой $a \in T$, аффинное множество H с отмеченной точкой $a \in H$ — конусы с вершиной в точке a . Замкнутое полупространство $\bar{\Pi}$ с отмеченной точкой $a \in \partial\bar{\Pi}$ — конус с вершиной в точке a . Если Π — полупространство и $a \in \partial\Pi$, то $\Pi \cup \{a\}$ — конус с вершиной в точке a . Любое объединение лучей, исходящих из некоторой точки, является конусом с вершиной в этой точке (рис. 9).

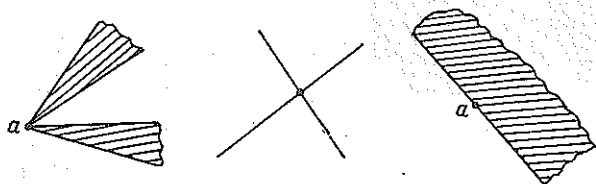


Рис. 9.

В соответствии с определением конус — непустое множество. Множество, состоящее из единственной точки, — конус. Все конусы, содержащие не менее двух точек, — неограниченные множества.

Конус называется *замкнутым*, если он является замкнутым множеством. Конус называется *выпуклым*, если он является выпуклым множеством. Если K_a — конус с вершиной в точке a , то $K_0 = K_a - a$ является конусом с вершиной в точке 0.

Предлагаем читателю доказать, что сумма $K_a + K_b$ конусов с вершинами соответственно в точках a и b есть конус с вершиной в точке $a + b$. В частности, *сумма конусов с вершинами в нуле — конус с вершиной в нуле.* — Упражнение 1.27.

Если K_0 — выпуклый конус с вершиной в нуле, то $K_0 + K_0 = K_0$.

Действительно, так как $0 \in K_0$, то $0 + K_0 = K_0$ и $K_0 \subset K_0 + K_0$. Если $a \in K_0$, $b \in K_0$, то $1/2(a + b) \in K_0$ (K_0 — выпуклое множество), но тогда и $2 \cdot 1/2(a + b) = a + b \in K_0$ (K_0 — конус), т. е. $K_0 + K_0 \subset K_0$. ■

6.2. Конус неограниченности. В дальнейшем вектор $s \in R^n$ иногда будем называть *направлением*. Будем говорить, что вектор $s \in R^n$, $s \neq 0$, есть *направление неограниченности множества M в точке $a \in M$* (или *направление неограниченности в точке a*), если луч $L^+ = \{x \in R^n: x = a + \lambda s, \lambda \geq 0\}$ содержится в множестве M . Очевидно, что для любого вектора b , принадлежащего лучу L^+ и такого, что $b \neq a$, вектор $b - a$ является направлением неограниченности множества M в точке a , поскольку луч, проведенный из a через b , совпадает с лучом L^+ . Множество K_a^∞ всех направлений неограниченности в точке a

с присоединенным к нему нулевым вектором есть конус с вершиной в нуле. Этот конус, который может состоять из единственной точки 0, назовем *конусом неограниченности множества M в точке a* . Очевидно, что если M — ограниченное множество, то $K_a^\infty = \{0\}$ для любой точки $a \in M$. Равенство $K_a^\infty = \{0\}$ может выполняться и для любой точки a неограниченного множества M . Однако это невозможно, если неограниченное множество M выпукло.

Теорема 1. Если V — неограниченное выпуклое множество, то $K_{v_0}^\infty \neq \{0\}$ для любой точки $v_0 \in V^0$ (V^0 — относительно открытое ядро множества V , см. п. 3.5).

Доказательство. Так как множество V не ограничено, то существует такая последовательность $\{v_k\}_{k=1}^\infty$ точек множества V , что $\|v_k - v_0\| = d_k \rightarrow \infty$ при $k \rightarrow \infty$. Элементы последовательности $\{s_k\}_{k=1}^\infty$, где $s_k = (v_k - v_0)/d_k$, принадлежат единичной сфере $S(0, 1)$. Так как $S(0, 1)$ — компактное множество, то из последовательности $\{s_k\}_{k=1}^\infty$ можно выделить подпоследовательность $\{s_k\}_{k \in K}$, сходящуюся к точке $s \in S(0, 1)$ при $k \rightarrow \infty$, $k \in K$. Покажем, что s — направление неограниченности в точке v_0 .

Для любого натурального числа m и любого $\varepsilon > 0$ выберем такой номер $k(m, \varepsilon)$, чтобы при любом $k \geq k(m, \varepsilon)$, $k \in K$, выполнялись неравенства $d_k > m$ и $\|s_k - s\| < \varepsilon/m$. Тогда точка $v_0 + ms_k = (1 - m/d_k)v_0 + (m/d_k)v_k$ принадлежит V (V — выпуклое множество) и $\|v_0 + ms_k - (v_0 + ms)\| = \|ms_k - ms\| < \varepsilon$. Поэтому при любом натуральном m точка $v_0 + ms$ луча $L^+ = \{v_0 + \lambda s, \lambda \geq 0\}$ принадлежит выпуклому множеству $\bar{V}$, а следовательно, и $L^+ \subset \bar{V}$. Так как $v_0 \in V^0 = \bar{V}^0$ (см. п. 3.8), то в силу теоремы о пересечении выпуклого множества лучом (см. п. 3.7) никакая точка луча $L^+ \subset \bar{V}$ не может принадлежать относительной границе $\partial \bar{V}$ множества $\bar{V}$. Значит, $L^+ \subset V^0 \subset V$ и $K_{v_0}^\infty \supset \{\lambda s, \lambda \geq 0, s \neq 0\} \neq \{0\}$. ■

Теорема 2. Если V — замкнутое выпуклое множество, то $K_{v_1}^\infty = K_{v_2}^\infty$ для любых точек $v_1 \in V$, $v_2 \in V$.

Доказательство. Утверждение очевидно, если V — ограниченное множество ($K_{v_1}^\infty = K_{v_2}^\infty = \{0\}$). Пусть s — направление неограниченности в точке v_1 и $L_1^+ = \{v_1 + \lambda s, \lambda \geq 0\}$ — луч, исходящий из точки v_1 в направлении s , содержащийся в V . Достаточно показать, что луч $L_2^+ = \{v_2 + \lambda s, \lambda \geq 0\} \subset V$.

Это очевидно, если v_2 принадлежит прямой $\{v_1 + \lambda s, \lambda \in R\}$ (рис. 10, а). Если $v_2 \notin \{v_1 + \lambda s, \lambda \in R\}$, то луч L_2^+ принадлежит замыканию выпуклой оболочки множества $\{v_2\} \cup L_1^+$ (см. рис. 10, б и упражнение 1.17). Но V — замкнутое выпуклое множество, и, следовательно, $L_2^+ \subset V$. ■

В силу теоремы 2 конус неограниченности замкнутого выпуклого множества V в точке a можно назвать *конусом неограниченности множества V* . Конус неограниченности замкнутого выпуклого множества V будем обозначать через K_V^∞ .

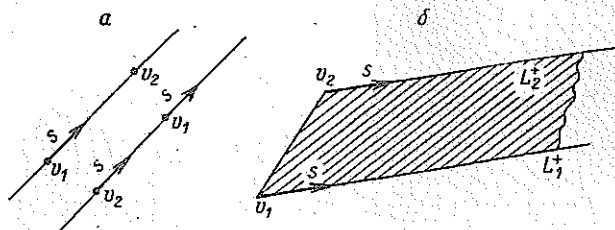


Рис. 10.

Теорема 3. Конус неограниченности K_V^∞ замкнутого выпуклого множества V есть замкнутый выпуклый конус.

Доказательство. Докажем выпуклость конуса K_V^∞ . Пусть $s_1 \in K_V^\infty$, $s_2 \in K_V^\infty$. Тогда лучи $L_1^+ = \{v + \lambda s_1, \lambda \geq 0\}$ и $L_2^+ = \{v + \lambda s_2, \lambda \geq 0\}$ содержатся в V при любом $v \in V$, а

$$v + \lambda [(1 - \alpha)s_1 + \alpha s_2] = (1 - \alpha)(v + \lambda s_1) + \alpha(v + \lambda s_2) \in V$$

при любом $\lambda \geq 0$ и любом $\alpha \in [0, 1]$ в силу выпуклости множества V . Следовательно, $(1 - \alpha)s_1 + \alpha s_2 \in K_V^\infty$ при любом $\alpha \in [0, 1]$. Замкнутость конуса K_V^∞ предлагаем читателю доказать самостоятельно. — Упражнение 1.28.

6.3. Следующее утверждение полезно для уяснения структурной характеристики неограниченных выпуклых множеств.

Теорема. Экстремальное множество $\text{ext } V$ непустого замкнутого выпуклого множества V непусто в том и только том случае, если K_V^∞ не содержит прямых (ср. с утверждением п. 5.1).

Доказательство. Пусть $\text{ext } V \neq \emptyset$ и $v_0 \in \text{ext } V$. Предположим, что множество K_V^∞ содержит некоторую прямую $L = \{x \in R^n: x = \alpha s, \alpha \in R\}$. Тогда множеству K_V^∞ принадлежат векторы s и $-s$, а значит, $v_0 + s$ и $v_0 - s$ из V , т. е. $v_0 \in \text{ext } V$.

Докажем теперь, что $\text{ext } V \neq \emptyset$, если K_V^∞ не содержит прямых. Доказательство проводится индукцией по размерности множества V . Утверждение очевидно, если $\dim V = 0$ или $\dim V = 1$. Предположим, что оно доказано для непустых замкнутых выпуклых множеств, размерность которых не превосходит k . Пусть $\dim V = k + 1$ и K_V^∞ не содержит прямых. Тогда $\partial V \neq \emptyset$: в противном случае V — аффинное множество размерности не меньше единицы, K_V^∞ — подпространство, сдвигом которого получается V , $\dim K_V^\infty \geq 1$, и, следовательно, K_V^∞ содержит прямую. Пусть

$v_0 \in \tilde{\partial}V$. По теореме об опорной гиперплоскости (см. п. 4.4) существует такая гиперплоскость H , опорная к V в точке v_0 , что $H \cap V^{00} = \emptyset$. Тогда $H \cap V$ — непустое замкнутое выпуклое множество, размерность которого не превосходит k . Так как очевидно $K_{H \cap V}^{\infty} \subset K_V^{\infty}$, то $K_{H \cap V}^{\infty}$ не содержит прямых, и в силу предположения индукции $\text{ext}(H \cap V) \neq \emptyset$. Но (см. лемму 1 п. 5.2) $\text{ext}(H \cap V) = H \cap \text{ext} V$, и, следовательно, $\text{ext} V \neq \emptyset$. ■

6.4. Теорема о представлении замкнутого выпуклого множества. Если V — непустое замкнутое выпуклое множество и конус K_V^{∞} не содержит прямых, то $V = v(\text{ext} V) + K_V^{\infty}$.

Доказательство. В силу утверждения предыдущего пункта множество $\text{ext} V \neq \emptyset$ и $v(\text{ext} V)$ — непустое подмножество множества V . Если $K_V^{\infty} = \{0\}$, то V — непустое компактное выпуклое множество (см. теорему 1 п. 6.2) и утверждение теоремы следует из теоремы п. 5.2. Поэтому в дальнейшем предполагаем, что $K_V^{\infty} \neq \{0\}$.

Так как $v(\text{ext} V) \subset V$, то из теоремы 2 п. 6.2 следует, что $v(\text{ext} V) + K_V^{\infty} \subset V$. Обратное включение докажем индукцией по размерности V .

Утверждение очевидно, если $\dim V = 0$ или $\dim V = 1$. Пусть оно верно, если $\dim V \leq k$, и V — замкнутое выпуклое множество размерности $k+1$, конус неограниченности которого K_V^{∞} не содержит прямых. Возьмем точку $v_0 \in V$, и пусть луч $L^+ = \{x \in R^n : x = as, a \geq 0\} \subset K_V^{\infty}$. Так как V не содержит прямых, то пересечение прямой $L = \{x \in R^n : x = v_0 + as, a \in R\}$ с множеством V есть луч $\tilde{v} + L^+$, где $\tilde{v} \in \tilde{\partial}V$ и $v_0 \in \tilde{v} + L^+$. Через точку $\tilde{v}$ проведем гиперплоскость H , опорную к V в точке $\tilde{v}$ и не пересекающуюся с V^{00} . Тогда $V \cap H$ — непустое выпуклое замкнутое множество, $\dim(V \cap H) \leq k$ (см. лемму 2 п. 5.2) и $K_{V \cap H}^{\infty}$ не содержит прямых, поскольку $K_{V \cap H}^{\infty} \subset K_V^{\infty}$. В силу предположения индукции $V \cap H = v(\text{ext}(V \cap H)) + K_{V \cap H}^{\infty}$. Поэтому $\tilde{v} \in v(\text{ext}(V \cap H)) + K_{V \cap H}^{\infty}$ и $v_0 \in v(\text{ext}(V \cap H)) + K_{V \cap H}^{\infty} + L^+$. Так как $K_{V \cap H}^{\infty} \subset K_V^{\infty}$, $L^+ \subset K_V^{\infty}$, то (см. п. 6.1) $K_{V \cap H}^{\infty} + L^+ \subset K_V^{\infty} + K_V^{\infty} = K_V^{\infty}$; $v(\text{ext}(V \cap H)) \subset v(\text{ext} V)$, поскольку $\text{ext}(V \cap H) \subset \text{ext} V$ в силу леммы 1 п. 5.2. Следовательно,

$$v(\text{ext}(V \cap H)) + K_{V \cap H}^{\infty} + L^+ \subset v(\text{ext} V) + K_V^{\infty} \quad \text{и} \\ v_0 \in v(\text{ext} V) + K_V^{\infty}. \quad \blacksquare$$

6.5. Теорема, которую мы здесь докажем, является в некотором смысле обратной к теореме предыдущего пункта (имеется в виду случай, когда $v(\text{ext} V)$ — компактное множество).

Теорема. Если W — непустое компактное выпуклое множество и K_0 — замкнутый выпуклый конус с вершиной в нуле, то $V = W + K_0$ — замкнутое выпуклое множество и K_0 есть конус неограниченности множества V .

В п. 3.1 было установлено, что сумма выпуклых множеств есть множество выпуклое. Докажем теперь три леммы, представляющие самостоятельный интерес, из которых непосредственно следует сформулированная выше теорема.

Лемма 1. *Сумма компактного и замкнутого множеств есть множество замкнутое.*

Доказательство. Пусть A — компактное, B — замкнутое множества и $C = A + B$. Пусть $\{c_k\}_{k=1}^{\infty}$ — последовательность точек множества C , сходящаяся к точке c_0 . Докажем, что $c_0 \in C$. Для любого $k \in N$ имеем $c_k = a_k + b_k$, $a_k \in A$, $b_k \in B$. Из последовательности $\{a_k\}_{k=1}^{\infty}$ точек компактного множества A выделим подпоследовательность $\{a_k\}_{k \in K}$, сходящуюся к точке $a_0 \in A$. Тогда последовательность $\{b_k\}_{k \in K}$ сходится к точке $b_0 = c_0 - a_0$, и так как B замкнуто, то $b_0 \in B$. Следовательно, $c_0 = a_0 + b_0$, $a_0 \in A$, $b_0 \in B$, т. е. $c_0 \in C$. ■

Лемма 2. *Гиперплоскость, опорная к замкнутому выпуклому конусу, проходит через его вершину.*

Доказательство. Пусть K_v — замкнутый выпуклый конус с вершиной в точке v_0 , $\bar{v} \in \tilde{K}_{v_0}$ и $H = \{x \in R^n: ax = \beta\}$ — гиперплоскость, опорная к конусу K_v в точке $\bar{v}$, так что $a\bar{v} = \beta$ и $av \leq \beta$ для любого $v \in K_v$. Утверждение очевидно, если $\bar{v} = v_0$. Пусть $\bar{v} \neq v_0$. Тогда векторы $\frac{1}{2}(v_0 + \bar{v})$ и $(1-2)v_0 + 2\bar{v} = 2\bar{v} - v_0$ принадлежат K_v . Поэтому $\frac{1}{2}a(v_0 + \bar{v}) \leq \beta = a\bar{v}$, $a(2\bar{v} - v_0) \leq a\bar{v}$, откуда следует $av_0 \leq a\bar{v}$ и $a\bar{v} \leq av_0$, т. е. $av_0 = a\bar{v} = \beta$. ■

Лемма 3. *Если замкнутое выпуклое множество V является суммой компактного выпуклого множества W и замкнутого выпуклого конуса K_0 с вершиной в нуле, то K_0 — конус неограниченности множества V .*

Доказательство. Очевидно, что $K_0 \subset K_V^{\infty}$. Докажем обратное включение. Пусть $\bar{u} \neq 0$, $\bar{u} \in K_V^{\infty}$ и $\bar{u} \notin K_0$. Из замечания к теореме о строгой отделимости следует, что существует гиперплоскость H , опорная к конусу K_0 , разделяющая компактное множество $\{\bar{u}\}$ и замкнутое множество K_0 и не проходящая через точку $\bar{u}$. В силу леммы 2 гиперплоскость H проходит через вершину конуса K_0 — точку 0. Следовательно, уравнение гиперплоскости имеет вид $ax = 0$, $a \neq 0$. Пусть $K_0 \subset \bar{\Pi}_H$, $\bar{u} \in \Pi_H^+$. Тогда $a\bar{u} > 0$ и $au \leq 0$ для любого $u \in K_0$.

На компактном множестве W рассмотрим непрерывную функцию $f(x) = ax$. По теореме Вейерштрасса существует такая точка $\bar{w} \in W$, что $a\bar{w} \geq aw$ для любого $w \in W$. Тогда для любого $v \in V$, $v = w + u$, $w \in W$, $u \in K_0$, имеем $av = aw + au \leq a\bar{w}$. В то же время $a(\bar{w} + \bar{u}) = a\bar{w} + a\bar{u} > a\bar{w}$. Поэтому $\bar{w} + \bar{u} \notin V$, и, следовательно (см. теорему 2 п. 6.2), $\bar{u} \notin K_V^{\infty}$, что противоречит выбору вектора $\bar{u}$. ■

§ 7. Многогранные выпуклые конусы.

Неограниченные многогранные выпуклые множества

7.1. Многогранные выпуклые конусы. Выпуклый конус K_a с вершиной в точке a назовем *многогранным выпуклым конусом*, если существует такой выпуклый многогранник $V = v(\{a_1, \dots, a_m\})$, что

$$K_a = \{x \in R^n : x = (1 - \alpha)a + \alpha v, v \in V, \alpha \geq 0\}. \quad (1)$$

Если конус K_a представим в виде (1), то говорят, что K_a — многогранный выпуклый конус с вершиной в точке a , натянутый на выпуклый многогранник V .

Для любого множества V множество K_a , определенное формулой (1), есть объединение лучей, идущих из точки a через точки множества V ; следовательно, K_a — конус. Нетрудно проверить, что для любой точки $a \in R^n$ и любого выпуклого множества V множество K_a , определенное формулой (1), выпукло. — Упражнение 1.29. Поэтому для любой точки $a \in R^n$ и любого выпуклого многогранника V формула (1) определяет многогранный выпуклый конус. Ясно, что один и тот же многогранный выпуклый ко-

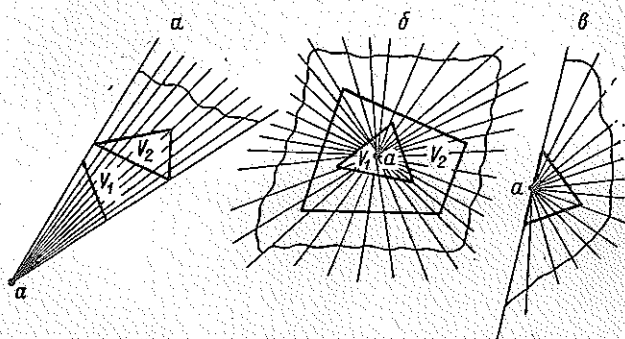


Рис. 11.

нус может быть натянут на разные выпуклые многогранники. На рис. 11 изображены многогранные выпуклые конусы и выпуклые многогранники, на которые они натянуты. Заметим (см. рис. 11, б)), что плоскость с отмеченной точкой есть многогранный выпуклый конус.

Пространство R^n с отмеченной точкой, всякое его аффинное подмножество с отмеченной точкой, полупространство $\Pi \subset R^n$ с отмеченной точкой на граничной гиперплоскости — многогранные выпуклые конусы. Многогранным выпуклым конусом является множество, состоящее из единственной точки (только такие множества суть ограниченные многогранные выпуклые конусы). Конечно, это далеко не полный набор примеров многогранных выпуклых конусов.

7.2. Представление многогранного выпуклого конуса. Будем рассматривать многогранные выпуклые конусы с вершиной в нуле. Это не ограничит общности наших рассмотрений, так как если K_a — многогранный выпуклый конус, натянутый на выпуклый многогранник V , то $K_a - a$ — многогранный выпуклый конус с вершиной в нуле, натянутый на выпуклый многогранник $V - a$, и обратно, если K_0 — многогранный выпуклый конус с вершиной в нуле, натянутый на выпуклый многогранник V , то $K_0 + a$ — многогранный выпуклый конус с вершиной в точке a , натянутый на выпуклый многогранник $V + a$.

Итак, пусть K_0 — многогранный выпуклый конус с вершиной в нуле, натянутый на выпуклый многогранник $V = v(\{a_1, \dots, a_m\})$, т. е.

$$K_0 = \{x \in R^n : x = \alpha v, v \in V, \alpha \geq 0\}. \quad (2)$$

Если $v \in V$, то $v = \sum_{i=1}^m \lambda_i a_i$, где $\lambda_i \geq 0, i \in 1:m$, и $\sum_{i=1}^m \lambda_i = 1$. Поэтому любая точка $x \in K_0$ может быть записана в виде $x = \alpha v = \sum_{i=1}^m \alpha \lambda_i a_i = \sum_{i=1}^m \alpha_i a_i$, где $\alpha_i \geq 0, i \in 1:m$. Другими словами, каждая точка $x \in K_0$ есть неотрицательная линейная комбинация векторов $a_1, a_2, \dots, a_m$. Обратно, пусть $x = \sum_{i=1}^m \alpha_i a_i, \alpha_i \geq 0, i \in 1:m$, — неотрицательная линейная комбинация векторов $a_1, a_2, \dots, a_m$. Если $\sum_{i=1}^m \alpha_i = 0$, то $\alpha_i = 0, i \in 1:m, x = 0$ и $x \in K_0$. Если $\sum_{i=1}^m \alpha_i > 0$, то

$$x = \sum_{i=1}^m \alpha_i a_i = (\sum_{k=1}^m \alpha_k) \sum_{i=1}^m (\alpha_i / \sum_{k=1}^m \alpha_k) a_i = \alpha v,$$

где $\alpha = \sum_{k=1}^m \alpha_k$, а $v = \sum_{i=1}^m (\alpha_i / \alpha) a_i$. Так как $\alpha_i / \alpha \geq 0, i \in 1:m$, и $\sum_{i=1}^m (\alpha_i / \alpha) = 1$, то $v \in V, x \in K_0$. ■

Таким образом, многогранный выпуклый конус с вершиной в нуле, натянутый на выпуклый многогранник $V = v(\{a_1, \dots, a_m\})$, есть множество всех неотрицательных линейных комбинаций векторов $a_1, a_2, \dots, a_m$. Векторы $a_1, a_2, \dots, a_m$ называют образующими векторами многогранного выпуклого конуса K_0 , а лучи $L_k^+ = \{\alpha a_k, \alpha \geq 0\}, k \in 1:m$, — его образующими лучами.

Из доказанного выше утверждения очевидно следует, что многогранный выпуклый конус K_0 является выпуклой оболочкой своих образующих лучей. Верно также, что сумма $K_0^1 + K_0^2$ многогранных выпуклых конусов с вершиной в нуле, натянутых на выпуклые многогранники $V_1 = v(\{a_1, a_2, \dots, a_m\})$ и $V_2 = v(\{b_1, b_2, \dots, b_s\})$, есть многогранный выпуклый конус с вершиной в нуле, натянутый на выпуклый многогранник $V = v(\{a_1, a_2, \dots, a_m\} \cup \{b_1, b_2, \dots, b_s\})$.

Пусть теперь $x = \sum_{i=1}^m \alpha_i a_i \in K_0$ ($\alpha_i \geq 0, i \in 1:m$) и I — множество индексов тех векторов $a_i, i \in 1:m$, при которых в записи

$x = \sum_{i=1}^m \alpha_i a_i$ стоят *положительные* коэффициенты: $x = \sum_{i \in I} \alpha_i a_i$ ($\alpha_i > 0, i \in I$). Тогда если $x \neq 0$, то *можно выбрать такое множество* $I' \subset I$, *что* $x = \sum_{i \in I'} \alpha'_i a_i$, $\alpha'_i > 0$ *для* $i \in I'$ *и система векторов* $\{a_i, i \in I'\}$ *— линейно независима*. (Это утверждение иногда называют *теоремой о независимом представлении многогранного выпуклого конуса*.)

Доказательство. Пусть $x \in K_0$, $x \neq 0$, и $x = \sum_{i \in I} \alpha_i a_i$, где $\alpha_i > 0, i \in I$. Если система векторов $\{a_i, i \in I\}$ линейно независима, то искомое разложение вектора x найдено. Предположим, что система векторов $\{a_i, i \in I\}$ линейно зависима. Тогда существуют такие числа $\beta_i, i \in I$, что $\sum_{i \in I} \beta_i a_i = 0$ и не все β_i равны нулю. Не ограничивая общности, можно считать, что среди чисел β_i есть положительные. Имеем

$$x = \sum_{i \in I} \alpha_i a_i = \sum_{i \in I'} \alpha_i a_i - \delta \sum_{i \in I} \beta_i a_i = \sum_{i \in I'} (\alpha_i - \delta \beta_i) a_i,$$

где δ — некоторое положительное число, которое мы сейчас выберем. Возьмем δ таким, чтобы все коэффициенты $\alpha_i - \delta \beta_i$ оказались неотрицательными, но хотя бы один из них обратился в нуль. Для этого достаточно положить $\delta = \min(\alpha_i / \beta_i)$, где минимум берется по тем индексам $i \in I$, для которых $\beta_i > 0$. Тогда вектор x запишется в виде $x = \sum_{i \in I'} \alpha'_i a_i$, где $\alpha'_i = \alpha_i - \delta \beta_i > 0, i \in I', I' \subset I$, и число элементов в множестве I' по крайней мере на единицу меньше числа элементов в множестве I . Если новая система векторов не является линейно независимой, то повторим описанную выше процедуру. Так как $x \neq 0$, а число векторов в представлении точки x каждый раз уменьшается по крайней мере на единицу, то через конечное число шагов получим искомую линейно независимую систему векторов. ■

З а м е ч а н и е. Возвращаясь к сказанному в начале настоящего пункта, отметим следующее.

Если K_a — многогранный выпуклый конус с вершиной в точке a , натянутый на выпуклый многогранник $V = v(\{a_1, \dots, a_m\})$, то каждая точка $x \in K_a$ может быть записана в виде

$$x = a + \sum_{i=1}^m \alpha_i (a_i - a), \alpha_i \geq 0, i \in 1:m,$$

и для любого $x \in K_a, x \neq a$, можно выбрать такое множество $I \subset 1:m$, что $x = a + \sum_{i \in I} \alpha_i (a_i - a)$, $\alpha_i > 0, i \in I$, и система векторов $\{a_i - a, i \in I\}$ линейно независима.

7.3. Замкнутость многогранного выпуклого конуса.

Теорема. *Многогранный выпуклый конус есть замкнутое множество.*

Доказательство. Пусть $\{v_s\}_{s=1}^\infty$ — последовательность точек многогранного выпуклого конуса K_0 с вершиной в нуле.

натянутого на выпуклый многогранник $V = v(\{a_1, \dots, a_m\})$, сходящаяся к точке v_0 . Покажем, что $v_0 \in K_0$.

По теореме предыдущего пункта для каждого s найдется такое множество $I_s \subset 1:m$, что $v_s = \sum_{i \in I_s} \alpha_i^{(s)} a_i$, $\alpha_i^{(s)} > 0$ для $i \in I_s$,

система векторов $\{a_i, i \in I_s\}$ линейно независима. Так как в множестве $1:m$ есть лишь конечное число подмножеств, то найдется такое бесконечное подмножество S множества натуральных чисел, что для всех $s \in S$ линейно независимые системы векторов $\{a_i, i \in I_s\}$ совпадают, т. е. $I_s = I \subset 1:m$ для любого $s \in S$. Дополним линейно независимую систему векторов $\{a_i, i \in I\}$ до базиса пространства R^n . Так как сходимость в R^n эквивалентна по координатной сходимости и $v_s \rightarrow v_0$ при $s \rightarrow \infty$, $s \in S$, то $v_0 = \sum_{i \in I} \alpha_i a_i$, где $\alpha_i = \lim_{s \rightarrow \infty} \alpha_i^{(s)}$, и, следовательно, $\alpha_i \geq 0$, $i \in I$,

т. е. $v_0 \in K_0$. ■

7.4. Теорема (лемма) Фаркаша и ее приложение к системам линейных уравнений и неравенств.

Теорема. Пусть $A = (a[i, j])$ — $m \times n$ матрица и $c = (c[1], \dots, c[n])$ — вектор из R^n . Тогда имеет место одно и только одно из следующих двух предложений (альтернатив): либо вектор c является неотрицательной комбинацией строчек матрицы A , либо найдется такой вектор $d \in R^n$, что $Ad \leq 0$, $cd > 0$.

Другая формулировка теоремы связана с условием существования неотрицательного решения системы линейных уравнений.

Система линейных уравнений $uA = c$ имеет неотрицательное решение в том и только том случае, если для всякого решения $\bar{x}$ системы линейных неравенств $Ax \leq 0$ выполняется неравенство $c\bar{x} \leq 0$. Равносильность приведенных формулировок нетрудно проверить. — Упражнение 1.30.

Доказательство. Пусть $a_1, a_2, \dots, a_m$ — строки матрицы A . Рассмотрим многогранный выпуклый конус K_0 с вершиной в нуле, натянутый на выпуклый многогранник $V = v(\{a_1, \dots, a_m\})$. Очевидно, что имеет место одно и только одно из следующих утверждений: либо $c \in K_0$, либо $c \notin K_0$. Покажем, что предположение $c \in K_0$ равносильно выполнению первой альтернативы, а $c \notin K_0$ — второй.

Вектор $c \in R^n$ в том и только том случае принадлежит многогранному выпуклому конусу (см. п. 7.2), если c является неотрицательной линейной комбинацией векторов $a_1, a_2, \dots, a_m$.

Покажем теперь, что существование такого вектора $d \in R^n$, для которого $Ad \leq 0$, $cd > 0$, равносильно тому, что $c \notin K_0$. Действительно, конус K_0 — замкнутое выпуклое множество, и если компактное выпуклое множество, состоящее из одной точки c , не содержится в K_0 , то (и только в этом случае) существует гиперплоскость H , опорная к конусу K_0 , разделяющая K_0 и $\{c\}$ и не проходящая через точку c (см. замечание к теореме о стро-

гой отделимости, п. 4.6). Пусть $dx = \gamma$, $d \neq 0$, — уравнение гиперплоскости H . Поскольку H проходит через вершину конуса K_0 (см. лемму 2 п. 6.5), то $\gamma = 0$ и уравнение гиперплоскости H имеет вид $dx = 0$. Пусть $K_0 \subset \bar{P}_H$, $c \in \bar{P}_H$. Тогда $dc > 0$ и для любого вектора из конуса K_0 , а значит, и для каждого вектора a_i , $i \in 1:m$, выполняется неравенство $da_i \leq 0$, т. е. $Ad \leq 0$.

Обратно, пусть существует такой вектор $d \in R^n$, что $dc > 0$ и $Ad \leq 0$ (т. е. $da_i \leq 0$ для $i \in 1:m$). Любой вектор $v \in K_0$ может быть представлен в виде неотрицательной линейной комбинации $v = \sum_{i=1}^m \alpha_i a_i$ векторов $a_1, \dots, a_m$ (см. п. 7.2). Но тогда $dv = \sum_{i=1}^m \alpha_i da_i \leq 0$ для любого $v \in K_0$, и, следовательно, $c \in K_0$. ■

7.5. Представление многогранного выпуклого множества. В настоящем и следующем пунктах будет дана полная структурная характеристика многогранных выпуклых множеств (см. п. 5.4), а именно будет установлено, что всякое многогранное выпуклое множество может быть представлено в виде суммы выпуклого многогранника и многогранного выпуклого конуса с вершиной в нуле, и обратно, сумма многогранного выпуклого конуса с вершиной в нуле и выпуклого многогранника есть многогранное выпуклое множество.

Введем некоторые определения. Пусть G — многогранное выпуклое множество, полученное пересечением замкнутых полупространств $\bar{P}_i$, $i \in 1:m$; H_i , $i \in 1:m$, — гиперплоскости, порождающие полупространства $\bar{P}_i$, $i \in 1:m$; $a_i x = \beta_i$ — уравнение гиперплоскости H_i ; $a_i x \leq \beta_i$ — неравенство, определяющее замкнутое полупространство $\bar{P}_i$. Напомним, что гиперплоскости H_i , $i \in 1:m$, названы ограничивающими.

Пусть I — некоторое подмножество индексов из $1:m$ (множество I может быть пустым). Множество $G_I = (\bigcap_{i \in I} H_i) \cap G$ называется *гранью* множества G , если $G_I \neq \emptyset$ (считаем, что $\bigcap_{i \in \emptyset} H_i = R^n$). Обозначим через I множество всех индексов гиперплоскостей H_i , содержащих грань G_I : $I = \{i \in 1:m: H_i \supset G_I\}$. Очевидно, что $G_{\hat{I}} = G_I$, грань G_I — многогранное выпуклое множество, и ее размерность $\dim G_I$ равна максимальному числу линейно независимых векторов в множестве $\{a_i, i \in I\}$. Грани размерности нуль называются *вершинами* многогранного выпуклого множества G , грани размерности единица — его *ребрами*. Ясно, что множество всех граней многогранного выпуклого множества конечно.

В множестве граней многогранного выпуклого множества выделим грани, имеющие минимальную размерность.

Лемма 1. *Всякая грань G_I минимальной размерности есть аффинное множество и $G_{\hat{I}_0} = \bigcap_{i \in \hat{I}_0} H_i$.*

Доказательство. Достаточно доказать, что $\bigcap_{i \in \bar{I}} H_i \subset G$. Предположим противное, и пусть $v_1 \in \bigcap_{i \in \bar{I}_0} H_i$, но $v_1 \notin G$. По определению грани существует точка $v_0 \in G_I$. Так как $v_0 \in G$, а $v_1 \notin G$, то в множестве ограничивающих гиперплоскостей с индексами $i \in 1:m \setminus I_0$ найдется гиперплоскость, пересекающая отрезок $[v_0, v_1]$. Выберем из множества таких гиперплоскостей гиперплоскость H_k , пересекающую отрезок $[v_0, v_1]$ в точке $\bar{v}$, наиболее близкой к v_0 . Тогда $[v_0, \bar{v}] \subset G$, множество $(\bigcap_{i \in \bar{I}} H_i) \cap H_k = G' \neq \emptyset$ и, следовательно, является гранью многогранного выпуклого множества G . Гиперплоскость H_k является опорной к $G_{\bar{I}_0}$ в точке $\bar{v}$ и не содержит G_I , так как $v_0 \notin H_k$. В силу леммы 2 п. 5.2 $\dim G' = \dim (G_I \cap H_k) < \dim G_I$, что противоречит определению грани G_{I_0} . ■

Следствие. Если $\dim G_{I_0} = 0$, т. е. $G_I = \{v_0\}$ — вершина множества G , то $v_0 \in \text{ext } G$.

Действительно, так как $\{v_0\} = \bigcap_{i \in \bar{I}_0} H_i$, то множество гиперплоскостей $\{H_i, i \in \bar{I}_0\}$ содержит подмножество из n независимых гиперплоскостей и в силу теоремы п. 5.5 точка v_0 является экстремальной точкой множества G . ■

Лемма 2. Всякая относительно граничная точка множества G принадлежит некоторой грани, размерность которой меньше размерности множества G .

Доказательство. Очевидно, что относительно открытое ядро множества G совпадает с пересечением ограничивающих гиперплоскостей, содержащих G , и (открытых) полупространств, порожденных остальными ограничивающими гиперплоскостями. Поэтому если v_0 — относительно граничная точка множества G , то существует ограничивающая гиперплоскость H_k , проходящая через v_0 и не содержащая G . В силу леммы 2 п. 5.2 размерность грани $H_k \cap G$, содержащей точку v_0 , меньше размерности множества G . ■

Лемма 3. Пусть $G_I = G_{\bar{I}_0}$ — грань минимальной размерности многогранного выпуклого множества G ; $T_{\bar{I}_0}$ — подпространство, сдвигом которого получается аффинное (см. лемму 1) множество $G_{\bar{I}_0}$; $T_{\bar{I}_0}^\perp$ — ортогональное дополнение подпространства $T_{\bar{I}_0}$.

Тогда:

- $G' = G \cap T_{\bar{I}_0}^\perp$ — многогранное выпуклое множество;
- грань $G'_{\bar{I}_0} = (\bigcap_{i \in \bar{I}_0} H_i) \cap G' = G_{\bar{I}_0} \cap G'$ многогранного выпуклого множества G' имеет размерность нуль;
- $G = G' + T_{\bar{I}_0}$.

Доказательство. Для удобства введем обозначения $G_{\bar{I}_0} = G_0$, $T_{\bar{I}_0} = T_0$, $T_{\bar{I}_0}^\perp = T_0^\perp$, $G'_{\bar{I}_0} = G'_0$.

а) Очевидно, что G' есть пересечение конечного числа замкнутых полупространств. Покажем, что $G' \neq \emptyset$. Пусть $x \in G_0$. Тогда $G_0 = T_0 + x$. Вектор x можно (однозначно) представить

в виде $x = y + z$, $y \in T_0$, $z \in T_0^\perp$. Так как T_0 — подпространство, то $(-y) \in T_0$, и, следовательно, $z = (-y) + x \in G_0 \subset G$.

Значит, $z \in G \cap T_0^\perp = G'$ и $G' \neq \emptyset$. ■

б) Так как $G_0 \subset G$, то $G'_0 = G_0 \cap G' = G_0 \cap G \cap T_0^\perp = G_0 \cap T_0^\perp$. Предположим, что $\dim G'_0 \geq 1$. Тогда существуют $v_1 \in G'_0$, $v_2 \in G'_0$, $v_1 \neq v_2$. Так как $v_2 \in G_0$, то $G_0 - v_2 = T_0$, поскольку T_0 — подпространство, сдвигом которого получается аффинное множество G_0 . Кроме того, $T_0^\perp - v_2 = T_0^\perp$, так как $T_0^\perp$ — подпространство, а $v_2 \in T_0^\perp$. Поэтому $v_1 - v_2 \in (G_0 - v_2) \cap (T_0^\perp - v_2) = T_0 \cap T_0^\perp = \{0\}$, т. е. $v_1 = v_2$. ■

в) Докажем теперь, что $G = G' + T_0$. Поскольку аффинное множество $G_0 \subset G$, то подпространство T_0 , сдвигом которого получается G_0 , содержится в конусе неограниченности множества G (см. п. 6.2). Поэтому $G' + T_0 \subset G$ (см. теорему 2 п. 6.2).

Вместе с тем если $x \in G$, то $x = y + z$, $y \in T_0$, $z \in T_0^\perp$. Но тогда $(-y) \in T_0$, и, следовательно, $z = x + (-y) \in G + T_0$.

Мы уже отмечали, что T_0 содержится в конусе неограниченности множества G . Тогда $G + T_0 = G$, $z \in G$, и так как $z \in T_0^\perp$ то $z \in G'$. Значит, $x = y + z$, $y \in T_0$, $z \in G'$, т. е. $x \in G' + T_0$, $G \subset G' + T_0$. ■

Теорема 1. *Многогранное выпуклое множество G , не содержащее прямых, представимо в виде $G = V + K_0$, где $V = v(\text{ext } G)$ — выпуклый многогранник, а K_0 — многогранный выпуклый конус с вершиной в нуле, являющийся конусом неограниченности множества G .*

Доказательство. Прежде всего заметим, что $\text{ext } G \neq \emptyset$ (см. п. 6.3) и конечно (см. п. 5.5). Поэтому $V = v(\text{ext } G)$ — выпуклый многогранник. Кроме того, $\text{ext } G \subset \partial G$, и, следовательно (см. лемму 2), все экстремальные точки множества G содержатся в гранях множества G , размерности которых меньше $\dim G$, так что $\text{ext } G$ есть объединение экстремальных множеств таких граней. Теорема верна, если G — ограниченное множество (см. п. 5.6). Предположим, что множество G не ограничено. В этом случае доказательство проведем индукцией по размерности множества G . Утверждение очевидно, если $\dim G = 1$. Предположим, что оно верно для многогранных выпуклых множеств, размерность которых не превосходит k , и пусть $\dim G = k + 1$. В силу сделанного предположения для каждой грани G_i множества G , размерность которой не превосходит k , найдется такое конечное множество векторов B_i , принадлежащих конусу неограниченности этой грани (и, следовательно, конусу неограниченности множества G), что каждый вектор этой грани представим в виде суммы двух элементов: неотрицательной линейной комбинации векторов множества B_i и выпуклой линейной комбинации экстремальных точек этой грани (см. п. 7.2).

Пусть $B = \{b_0, b_1, b_2, \dots, b_s\}$ — объединение всех множеств B_i , отвечающих граням G_i ($\dim G_i \leq k$), дополненное, если в этом есть необходимость, каким-нибудь направлением неограниченности $b_0 \neq 0$ множества G . Утверждается, что $G = V + K_0$, где $V = v(\text{ext } G)$, а K_0 — многогранный выпуклый конус с вершиной в нуле, натянутый на выпуклый многогранник $v(B)$. Действительно, пусть $v \in G$. Прямая $L = \{x \in R^n : x = v + \alpha b_0, \alpha \in R\}$ не содержится в G , и, следовательно, на этой прямой найдется точка $\bar{v} \in \partial \tilde{G}$: $\bar{v} = \bar{v} + \bar{\alpha} b_0$ и $\bar{\alpha} \leq 0$, так как луч $L^+ = \{x \in R^n : x = v + \alpha b_0, \alpha \geq 0\} \subset K_0^\infty$. Поскольку $\bar{v} \in \partial \tilde{G}$, то найдется такая грань G_i ($\dim G_i \leq k$), что $\bar{v} \in G_i$ (см. лемму 2). Поэтому $\bar{v}$ есть сумма двух элементов: выпуклой линейной комбинации точек множества $\text{ext } G_i$ и неотрицательной линейной комбинации векторов множества B_i , а значит, и $v = \bar{v} + (-\bar{\alpha})b_0$ есть сумма вектора из $v(\text{ext } G)$ и вектора из K_0 .

Остается отметить, что в силу леммы 3 п. 6.5 $K_0 = K_0^\infty$. ■

Следствие. Если G — многогранное выпуклое множество, не содержащее прямых, то существуют конечное множество $\{a_1, \dots, a_m\}$ экстремальных точек множества G и такое конечное множество векторов $b_1, \dots, b_s$, что всякая точка $v \in G$ представима в виде $v = \sum_{i=1}^m \alpha_i a_i + \sum_{j=1}^s \beta_j b_j$, $\alpha_i \geq 0$, $i \in 1:m$, $\sum_{i=1}^m \alpha_i = 1$, $\beta_j \geq 0$, $j \in 1:s$.

Теорема 2. Если G — многогранное выпуклое множество, то существуют такие выпуклый многогранник V и многогранный выпуклый конус K_0 с вершиной в нуле, что $G = V + K_0$.

Замечание. В силу леммы 3 п. 6.5 $K_0 = K_0^\infty$, и, следовательно, конус неограниченности многогранного выпуклого множества есть многогранный выпуклый конус.

Доказательство теоремы. В силу леммы 3 многогранное выпуклое множество G представимо в виде $G = G' + T_{\hat{L}}$, где $T_{\hat{L}}$ — подпространство пространства R^n , а G' — многогранное выпуклое множество, не содержащее прямых. В силу теоремы 1 $G' = V + K'_0$, где V — выпуклый многогранник, а K'_0 — многогранный выпуклый конус с вершиной в нуле. Поэтому $G = V + K'_0 + T_{\hat{L}}$. Но подпространство $T_{\hat{L}}$ тоже является многогранным выпуклым конусом с вершиной в нуле (см. п. 7.1), и, следовательно, $K'_0 + T_{\hat{L}} = K_0$ — многогранный выпуклый конус с вершиной в нуле (см. п. 7.2) и $G = V + K_0$. ■

Следствие. Если G — многогранное выпуклое множество, то существуют такие векторы $a_1, \dots, a_m, b_1, \dots, b_s$, что всякая точка $v \in G$ представима в виде

$$v = \sum_{i=1}^m \alpha_i a_i + \sum_{j=1}^s \beta_j b_j, \quad \alpha_i \geq 0, \quad i \in 1:m, \quad \sum_{i=1}^m \alpha_i = 1, \quad \beta_j \geq 0, \quad j \in 1:s.$$

7.6. Теорема. Пусть $V = v(\{a_1, a_2, \dots, a_m\})$ — выпуклый многогранник и K_0 — многогранный выпуклый конус с вершиной в нуле, натянутый на выпуклый многогранник $W = v(\{b_1, b_2, \dots, b_s\})$. Тогда $G = V + K_0$ — многогранное выпуклое множество (частный случай этой теоремы для $K_0 = \{0\}$ установлен в п. 5.7).

Докажем сначала следующую лемму.

Лемма. Пусть $v_0 \in \partial G$, $v_0 = \sum_{i \in I} \alpha_i a_i + \sum_{j \in J} \beta_j b_j$, где $I \subset 1:m$, $J \subset 1:s$; $\alpha_i > 0$, $i \in I$, $\sum_{i \in I} \alpha_i = 1$; $\beta_j > 0$, $j \in J$, и пусть $H = \{x \in R^n : ax = \beta\}$ — гиперплоскость, опорная к G в точке v_0 . Тогда любая точка v , представляемая в виде

$$v = \sum_{i \in I} \alpha'_i a_i + \sum_{j \in J} \beta'_j b_j, \quad \alpha'_i \geq 0, \quad i \in I, \quad \sum_{i \in I} \alpha'_i = 1; \\ \beta'_j \geq 0, \quad j \in J,$$

принадлежит $H \cap G$.

Доказательство. Очевидно, что всякая точка v , удовлетворяющая условию леммы, принадлежит G . Покажем, что $v \in H$. Предположим, что точка $\bar{v} = \sum_{i \in I} \bar{\alpha}_i a_i + \sum_{j \in J} \bar{\beta}_j b_j$, $\bar{\alpha}_i \geq 0$, $i \in I$, $\sum_{i \in I} \bar{\alpha}_i = 1$; $\bar{\beta}_j \geq 0$, $j \in J$, не принадлежит H , и пусть для определенности $G \subset \bar{\Pi}_H$. Тогда $av_0 = \beta$, $a\bar{v} < \beta$ и $(1-\lambda)v_0 + \lambda\bar{v} \in G$ при $\lambda < 0$, поскольку $a[(1-\lambda)v_0 + \lambda\bar{v}] = (1-\lambda)\beta + \lambda a\bar{v} > (1-\lambda)\beta + \lambda\beta = \beta$. Но $(1-\lambda)v_0 + \lambda\bar{v} = \sum_{i \in I} [\alpha_i + \lambda(\bar{\alpha}_i - \alpha_i)] a_i + \sum_{j \in J} [\beta_j + \lambda(\bar{\beta}_j - \beta_j)] b_j$, и так как $\alpha_i > 0$, $i \in I$, $\beta_j > 0$, $j \in J$, то можно подобрать такое $\bar{\lambda} < 0$, что $\alpha_i + \bar{\lambda}(\bar{\alpha}_i - \alpha_i) > 0$, $i \in I$, $\beta_j + \bar{\lambda}(\bar{\beta}_j - \beta_j) > 0$, $j \in J$. К тому же $\sum_{i \in I} [\bar{\alpha}_i + \bar{\lambda}(\bar{\alpha}_i - \alpha_i)] = 1$, так что $(1-\bar{\lambda})v_0 + \bar{\lambda}\bar{v} \in G$. ■

Доказательство теоремы. Так как K_0 — замкнутый выпуклый конус (см. п. 7.3), то G — замкнутое выпуклое множество (см. п. 6.5) и для доказательства теоремы достаточно построить конечное семейство гиперплоскостей, опорных к G и удовлетворяющих условиям леммы п. 5.7.

Пусть v_1 — какая-нибудь точка из ∂G . Построим гиперплоскость H_1 , опорную к G в точке v_1 и не пересекающуюся с G^{00} (п. 4.4). Пусть $v_1 = \sum_{i \in I_1} \alpha_i^{(1)} a_i + \sum_{j \in J_1} \beta_j^{(1)} b_j$, где $\alpha_i^{(1)} > 0$, $i \in I_1$, $\sum_{i \in I_1} \alpha_i^{(1)} = 1$; $\beta_j^{(1)} > 0$, $j \in J_1$. В силу леммы множество $H_1 \cap G$ содержит все относительно граничные точки множества G , представимые в виде

$$v = \sum_{i \in I_1} \alpha_i a_i + \sum_{j \in J_1} \beta_j b_j, \quad \alpha_i \geq 0, \quad i \in I_1, \\ \sum_{i \in I_1} \alpha_i = 1; \quad \beta_j \geq 0, \quad j \in J_1.$$

Если $H_1 \cap G \supset \partial G$, то процесс построения системы гиперплоскостей окончен. В противном случае найдется точка $v_2 \in \partial G$,

$v_2 \in H_1 \cap G$. Тогда $v_2 = \sum_{i \in I_1} \alpha_i^{(2)} a_i + \sum_{j \in I_2} \beta_j^{(2)} b_j$, $\alpha_i^{(2)} > 0$, $i \in I_2$, $\sum_{i \in I_1} \alpha_i^{(2)} = 1$; $\beta_j \geq 0$, $j \in I_2$, и $I_2 \neq I_1$ или $I_2 \neq J_1$. Всякая относительно граничная точка множества G , представимая в виде

$$v = \sum_{i \in I_1} \alpha_i a_i + \sum_{j \in I_2} \beta_j b_j, \quad \alpha_i \geq 0, \quad i \in I_2, \\ \sum_{i \in I_1} \alpha_i = 1; \quad \beta_j \geq 0, \quad j \in I_2,$$

принадлежит $H_2 \cap G$ (см. лемму). Если $(H_1 \cap G) \cup (H_2 \cap G) \supset \bar{\partial} G$, то процесс окончен. В противном случае найдется точка $v_3 \in \bar{\partial} G$, $v_3 \in (H_1 \cap G) \cup (H_2 \cap G)$ и т. д.

Очевидно, что процесс окончится через конечное число шагов, поскольку каждое из множеств $1:m$ и $1:s$ содержит конечное число подмножеств. ■

Следствие. Сумма $G_1 + G_2$ многогранных выпуклых множеств G_1 и G_2 есть многогранное выпуклое множество.

Действительно, в силу теоремы 2 п. 7.5 $G_1 = V_1 + K_0^1$, $G_2 = V_2 + K_0^2$, где V_1 и V_2 — выпуклые многогранники, а K_0^1 и K_0^2 — многогранные выпуклые конусы с вершиной в нуле. Тогда $G_1 + G_2 = (V_1 + V_2) + (K_0^1 + K_0^2)$ и $V_1 + V_2$ — выпуклый многогранник (п. 5.3), а $K_0^1 + K_0^2$ — многогранный выпуклый конус с вершиной в нуле (см. п. 7.2). Из доказанной выше теоремы следует, что $G_1 + G_2$ — многогранное выпуклое множество. ■

§ 8. Выпуклые функции

8.1. Определения. Обозначим через $\bar{R}$ расширенную прямую: множество вещественных чисел R , пополненное точками $-\infty$ и $+\infty$. Как обычно, будем считать, что $-\infty < a < +\infty$ для любого $a \in R$; $a + \infty = +\infty + a = +\infty$, $a \in R$; $-\infty + a = a - \infty = -\infty$, $a \in R$; $a(+\infty) = +\infty$, если $a > 0$; $a(+\infty) = -\infty$, если $a < 0$; $0(+\infty) = 0(-\infty) = 0$ и т. д.

Введем следующее обозначение. Если $x = (x[1], \dots, x[n]) \in R^n$ и $x[n+1] \in R$, то элемент $(x[1], \dots, x[n], x[n+1]) \in R^{n+1}$ будем обозначать через $(x, x[n+1])$.

Пусть $f: R^n \rightarrow \bar{R}$. Свяжем с функцией f множество

$$\text{epi } f = \{(x, x[n+1]) \in R^{n+1} : x \in R^n, x[n+1] \geq f(x)\}.$$

Множество $\text{epi } f$ назовем *надграфиком* функции f . Заметим, что если $(x, x[n+1]) \in \text{epi } f$, то $f(x) < +\infty$, и если $f(x) < +\infty$, то существует такое $x[n+1] \in R$, что $(x, x[n+1]) \in \text{epi } f$.

Функция $f: R^n \rightarrow \bar{R}$ называется *выпуклой*, если $\text{epi } f$ — выпуклое множество в R^{n+1} . Функция $f: R^n \rightarrow \bar{R}$ называется *вогнутой*, если $-f$ — выпуклая функция.

Очевидными примерами выпуклых функций являются аффинные функции $f: R^n \rightarrow R$, $f(x) = ax + \beta$. Для каждой такой функ-

ции f ее надграфик $\text{epi } f$ есть замкнутое полупространство в R^n . Очевидно также, что аффинная функция является функцией и выпуклой, и вогнутой.

В дальнейшем будем рассматривать только такие выпуклые (вогнутые) функции, для которых $f(x) > -\infty$ при любом $x \in R^n$ ($f(x) < +\infty$ соответственно).

С каждой выпуклой (вогнутой) функцией $f: R^n \rightarrow \bar{R}$ свяжем множество $\text{dom } f = \{x \in R^n: f(x) < +\infty\}$ (для вогнутой функции $\text{dom } f = \{x \in R^n: f(x) > -\infty\}$) — эффективную область определения функции f . Множество $\{(x, x[n+1]) \in R^{n+1}: x \in \text{dom } f, x[n+1] = f(x)\}$ называется графиком функции f .

Приведем теперь характеристику выпуклой функции, которая может быть принята за ее определение.

Для того чтобы функция $f: R^n \rightarrow \bar{R}$ была выпуклой, необходимо и достаточно, чтобы множество $\text{dom } f$ было выпуклым подмножеством R^n и для любых $x_1 \in \text{dom } f$, $x_2 \in \text{dom } f$ и любого $\alpha \in (0, 1)$ выполнялось неравенство

$$f((1-\alpha)x_1 + \alpha x_2) \leq (1-\alpha)f(x_1) + \alpha f(x_2). \quad (1)$$

Доказательство. Необходимость. Пусть $x_1 \in \text{dom } f$, $x_2 \in \text{dom } f$, $\alpha \in (0, 1)$. Тогда $(x_1, f(x_1)) \in \text{epi } f$, $(x_2, f(x_2)) \in \text{epi } f$, и так как $(1-\alpha)f(x_1) + \alpha f(x_2) < +\infty$, то $(1-\alpha)(x_1, f(x_1)) + \alpha(x_2, f(x_2)) = ((1-\alpha)x_1 + \alpha x_2, (1-\alpha)f(x_1) + \alpha f(x_2)) \in \text{epi } f$. Следовательно, $(1-\alpha)x_1 + \alpha x_2 \in \text{dom } f$ и выполняется неравенство (1).

Достаточность. Пусть $\text{dom } f$ — выпуклое множество и для любых x_1, x_2 из $\text{dom } f$ и любого $\alpha \in (0, 1)$ выполняется неравенство (1). Если $(x_1, x'[n+1]) \in \text{epi } f$, $(x_2, x''[n+1]) \in \text{epi } f$ и $\alpha \in (0, 1)$, то $x_1 \in \text{dom } f$, $x_2 \in \text{dom } f$, $(1-\alpha)x_1 + \alpha x_2 \in \text{dom } f$ и $f((1-\alpha)x_1 + \alpha x_2) \leq (1-\alpha)f(x_1) + \alpha f(x_2) \leq (1-\alpha)x'[n+1] + \alpha x''[n+1]$, так что $((1-\alpha)x_1 + \alpha x_2, (1-\alpha)x'[n+1] + \alpha x''[n+1]) \in \text{epi } f$. ■

Из доказанного выше утверждения очевидно следует условие вогнутости функции f : для того чтобы функция $f: R^n \rightarrow \bar{R}$ была вогнутой, необходимо и достаточно, чтобы множество $\text{dom } f$ было выпуклым и для любых $x_1 \in \text{dom } f$, $x_2 \in \text{dom } f$ и любого $\alpha \in (0, 1)$ выполнялось неравенство

$$f((1-\alpha)x_1 + \alpha x_2) \geq (1-\alpha)f(x_1) + \alpha f(x_2). \quad (2)$$

Легко видеть, что если f — выпуклая функция на R^n , $x_1, x_2, \dots, x_k$ — система векторов из R^n , $\alpha_1, \alpha_2, \dots, \alpha_k$ — такие числа из R , что $\alpha_i \geq 0$, $i \in 1:k$, $\sum_{i=1}^k \alpha_i = 1$, то

$$f\left(\sum_{i=1}^k \alpha_i x_i\right) \leq \sum_{i=1}^k \alpha_i f(x_i). \quad (3)$$

Это неравенство называется неравенством Иенсена. Оно следует из неравенства (1) и доказывается индукцией по k . — Упражнение 1.31.

Пусть V — выпуклое множество в R^n и $V \times R = \{(x, x[n+1]) \in R^{n+1} : x \in V, x[n+1] \in R\}$. Функция $f: R^n \rightarrow \bar{R}$ называется *выпуклой* (вогнутой) на выпуклом множестве V , если множество $(V \times R) \cap \text{epi } f$ (множество $(V \times R) \cap \text{epi } (-f)$) выпукло. Ясно, что функция f тогда и только тогда выпукла (вогнута) на V , когда для любых $x_1 \in V, x_2 \in V$ и любого $\alpha \in (0, 1)$ выполняется неравенство (1) (неравенство (2)). Если при этом для любых $x_1 \in V, x_2 \in V, x_1 \neq x_2$ и любого $\alpha \in (0, 1)$ неравенство (1) (неравенство (2)) выполняется строго, то говорят, что f — *строго выпуклая* (строго вогнутая) функция на V .

Отметим еще, что неотрицательная линейная комбинация $f = \sum_{k=1}^m \alpha_k f_k, \alpha_k \geq 0, k \in 1:m$, системы $f_1, f_2, \dots, f_k$ выпуклых функций из R^n в $\bar{R}$ есть выпуклая функция из R^n в $\bar{R}$. — Упражнение 1.37.

8.2. Экстремальные свойства выпуклых функций. В этом пункте будут установлены некоторые свойства выпуклых и вогнутых функций, имеющие фундаментальное значение для теории выпуклого программирования. Они сформулированы для выпуклых функций. Предоставляем читателю самостоятельно переформулировать эти свойства для вогнутых функций.

1°. Пусть $f: R^n \rightarrow \bar{R}$ — выпуклая функция и $\alpha \in R$. Тогда множество $L_\alpha = \{x \in R^n : f(x) \leq \alpha\}$ — выпуклое. В частности, если V — выпуклое множество в R^n и $\bar{x}$ — такая точка из V , что $f(\bar{x}) \leq f(x)$ для любого $x \in V$ ($f(\bar{x}) = \min_{x \in V} f(x)$), то множество $L = \{x \in V : f(x) = f(\bar{x})\} = L_{f(\bar{x})} \cap V$ выпуклое.

Действительно, для любых $x_1 \in L_\alpha, x_2 \in L_\alpha$ и $\lambda \in (0, 1)$ выполняется неравенство

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2) \leq (1-\lambda)\alpha + \lambda\alpha = \alpha,$$

и, следовательно, $(1-\lambda)x_1 + \lambda x_2 \in L_\alpha$, L_α — выпуклое множество и L — пересечение двух выпуклых множеств — множество выпуклое. ■

2°. Если функция f строго выпуклая на выпуклом множестве V и $\bar{x}$ — такая точка из V , что $f(\bar{x}) = \min_{x \in V} f(x)$, то множество $L = \{x \in V : f(x) = f(\bar{x})\}$ состоит из единственной точки (точки $\bar{x}$).

В самом деле, если $\bar{x}_1 \in L, \bar{x}_1 \neq \bar{x}$, то $1/2(\bar{x}_1 + \bar{x}) \in L$ и $f(1/2(\bar{x}_1 + \bar{x})) < 1/2(f(\bar{x}_1) + f(\bar{x})) = f(\bar{x})$, что противоречит определению точки $\bar{x}$. ■

3°. Если $f: R^n \rightarrow \bar{R}$ — выпуклая функция и V — непустое компактное выпуклое множество в R^n , то $\sup_{x \in V} f(x) = \sup_{x \in \text{ext } V} f(x)$ ($\text{ext } V$ — экстремальное множество выпуклого множества V (см. п. 1.5)).

Очевидно, что $\sup_{x \in \text{ext } V} f(x) \leq \sup_{x \in V} f(x)$. Вместе с тем всякая точка v непустого компактного множества V является выпуклой

комбинацией его экстремальных точек (см. п. 5.2):

$$v = \sum_{i=1}^k \alpha_i v_i, \quad \alpha_i \geq 0, \quad \sum_{i=1}^k \alpha_i = 1, \quad v_i \in \text{ext } V, \quad i \in 1:k.$$

Поэтому в силу неравенства Йенсена (см. (3))

$$\begin{aligned} f(v) &= f\left(\sum_{i=1}^k \alpha_i v_i\right) \leq \sum_{i=1}^k \alpha_i f(v_i) \leq \\ &\leq \left[\sup_{x \in \text{ext } V} f(x)\right] \sum_{i=1}^k \alpha_i = \sup_{x \in \text{ext } V} f(x), \end{aligned}$$

и, значит, $\sup_{x \in V} f(x) \leq \sup_{x \in \text{ext } V} f(x)$. ■

4°. Если $f: R^n \rightarrow \bar{R}$ — выпуклая функция и $V = v(\{v_1, \dots, v_k\})$ — выпуклый многогранник в R^n , то f принимает наибольшее значение на V в вершине выпуклого многогранника V .

Не ограничивая общности, можно считать, что $\{v_1, \dots, v_k\}$ — множество вершин выпуклого многогранника V (см. п. 5.3). Пусть v_{i_0} — такая вершина, что $f(v_{i_0}) = \max\{f(v_i), i \in 1:k\}$. Так как $\text{ext } V = \{v_1, \dots, v_k\}$, то (см. 3°)

$$\sup_{x \in V} f(x) = \max_{x \in V} f(x) = \max\{f(v_i), i \in 1:k\} = f(v_{i_0}). \quad \blacksquare$$

5°. Если $f: R^n \rightarrow R$, $f(x) = ax$ — линейная функция на R^n и V — выпуклый многогранник, то f принимает наибольшее и наименьшее значения в вершинах многогранника V .

Действительно, так как f — и выпуклая, и вогнутая функция, то для f и $-f$ выполняется свойство 4°. ■

6°. Если $f: R^n \rightarrow R$, $f(x) = ax$ — линейная функция на R^n и G — многогранное выпуклое множество в R^n , то f ограничена сверху (снизу) на G в том и только том случае, если $f(s) \leq 0$ ($f(s) \geq 0$) для любого вектора s из конуса неограниченности K_G^∞ (см. п. 6.2) множества G .

В силу теоремы 2 п. 7.5 $G = V + K_G^\infty$, где V — выпуклый многогранник. Положим $\bar{a} = \max_{v \in V} f(v) = \max_{v \in V} av$. Если $f(s) = as \leq 0$ для любого $s \in K_G^\infty$, то для любого $x \in G$, $x = v + s$, $v \in V$, $s \in K_G^\infty$, имеем $f(x) = ax = av + as \leq \bar{a}$.

Обратно, если существует такой вектор $s_0 \in K_G^\infty$, что $f(s_0) = as_0 > 0$, то для любого $v_0 \in V$ вектор $v_0 + \lambda s_0 \in G$ при любом $\lambda > 0$ и $f(v_0 + \lambda s_0) = a(v_0 + \lambda s_0) = av_0 + \lambda as_0$, а значит, f не ограничена сверху на G . ■

7°. Если $f: R^n \rightarrow R$, $f(x) = ax$ — линейная функция на R^n , G — многогранное выпуклое множество и f ограничена сверху (снизу) на G , то существует такая точка $\bar{x} \in G$, что

$$f(\bar{x}) = \max_{x \in G} f(x) \quad (f(\bar{x}) = \min_{x \in G} f(x)).$$

Если к тому же многогранное выпуклое множество G не содержит прямых, то точку $\bar{x}$ можно выбрать в множестве $\text{ext } G$ экстремальных точек множества G .

В множестве V (см. доказательство свойства 6°) выберем такую точку $\bar{v}$ (точку $\underline{v}$), что $f(\bar{v}) = a\bar{v} = \bar{a} = \max_{v \in V} (f(v)) = \underline{a} = \min_{v \in V} av$ (см. свойство 5°). Тогда $\bar{v} \in G$ ($\underline{v} \in G$), поскольку

$\bar{v} = \bar{v} + 0 \in V + K_G^\infty = G$. Положим $\bar{x} = \bar{v}$ ($\bar{x} = \underline{v}$). В силу свойства 6° $f(s) \leq 0$ ($f(s) \geq 0$) для любого $s \in K_G^\infty$, и, следовательно, для любого $x \in G$, $x = \bar{v} + s$, $v \in V$, $s \in K_G^\infty$, имеем $f(x) = ax = av + as \leq a\bar{v} = f(\bar{x})$ ($f(x) = av + as \geq av = f(\bar{x})$). Если к тому же множество G не содержит прямых, то можно взять $V = v(\text{ext } G)$ (см. теорему 1 п. 7.5) и выбрать точку $\bar{v}$ (точку $\underline{v}$) в множестве $\text{ext } G$ (см. свойство 5°). ■

8.3. Относительно внутренние точки надграфика выпуклой функции. Тривиальным примером выпуклой функции $f: R^n \rightarrow \bar{R}$ является функция, значение которой в каждой точке R^n равно $+\infty$. Для такой функции f очевидно $\text{dom } f = \emptyset$, $\text{epi } f = \emptyset$. Если $f: R^n \rightarrow \bar{R}$ — выпуклая функция, не равная тождественно $+\infty$, то $\text{dom } f$ и $\text{epi } f$ — непустые выпуклые множества в R^n и R^{n+1} соответственно*, причем ясно, что в этом случае $\dim \text{epi } f = \dim \text{dom } f + 1$ и относительно открытые ядра $(\text{dom } f)^{00}$ и $(\text{epi } f)^{00}$ множества $\text{dom } f$ и $\text{epi } f$ — непустые выпуклые множества (см. п. 3.6).

В дальнейшем нам понадобится следующая теорема.

Теорема. Пусть $f: R^n \rightarrow \bar{R}$ — выпуклая функция, не равная тождественно $+\infty$. Точка $(x^0, x^0[n+1]) \in \text{epi } f$ в том и только в том случае принадлежит $(\text{epi } f)^{00}$, если $x^0 \in (\text{dom } f)^{00}$ и $x^0[n+1] > f(x^0)$.

Доказательство. Необходимость. Так как $(x^0, x^0[n+1]) \in \text{epi } f$, то $x^0 \in \text{dom } f$. Если $x^0[n+1] = f(x^0)$, то точка $(x^0, x^0[n+1])$ является пределом последовательности точек $\{z_k = (x^0, x^0[n+1] - 1/k)\}_{k=1}^\infty$, очевидно принадлежащих $\text{aff } \text{epi } f$ и не принадлежащих $\text{epi } f$, и, следовательно, $(x^0, f(x^0)) \in \text{epi } f$ и не принадлежащих $\text{epi } f$, и, следовательно, $(x^0, f(x^0)) \in \text{epi } f$. Если $x^0 \in \text{dom } f$, то найдется последовательность $\{x_k\}_{k=1}^\infty$ точек множества $\text{aff } \text{dom } f$, не принадлежащих $\text{dom } f$, сходящаяся к x^0 при $k \rightarrow \infty$. Тогда $\{y_k = (x_k, x^0[n+1])\}_{k=1}^\infty$ — последовательность точек множества $\text{aff } \text{epi } f$, не принадлежащих $\text{epi } f$, сходящаяся к $(x^0, x^0[n+1])$ при $k \rightarrow \infty$, так что $(x^0, x^0[n+1]) \in (\text{epi } f)^{00}$.

Достаточность. Пусть $x^0 \in (\text{dom } f)^{00}$ и $x^0[n+1] > f(x^0)$. Существует такой шар $\mathcal{D}(x^0, r)$, что $\text{aff } \text{dom } f \cap \mathcal{D}(x^0, r) \subset \text{dom } f$ (см. п. 3.5). Нетрудно построить такой замкнутый пря-

* Напомним, что мы условились (см. п. 8.1) не рассматривать выпуклые функции, принимающие значения $-\infty$.

моугольник $P[x^0 - d, x^0 + d], d > 0$, что $P[x^0 - d, x^0 + d] \subset \mathcal{D}(x^0, r)$ (см. упражнение 1.6.а)). Но тогда $V = P[x^0 - d, x^0 + d] \cap \text{aff dom } f$ — ограниченное многогранное выпуклое множество, $V \subset (\text{dom } f)^{00}$ и содержит некоторую относительную окрестность $\mathcal{D}(x^0, r') \cap \text{aff dom } f$ точки x^0 . В силу теоремы п. 5.6 V — выпуклый многогранник, и, следовательно (см. 4° п. 8.2), выпуклая функция f принимает наибольшее значение на V в одной из его вершин. Пусть это значение равно $\bar{a}$. Тогда множество $\{(x, x[n+1]) \in R^{n+1} : x \in \mathcal{D}(x^0, r') \cap \text{dom } f, x[n+1] > \bar{a}\}$ содержится в $\text{epi } f$ (рис. 12). Выберем $\alpha_0 > \max\{\bar{a}, x^0[n+1]\}$ и $r'' = \min\{r', \alpha_0 - \bar{a}\} > 0$. Тогда относительная окрестность $\mathcal{D}((x^0, \alpha_0), r'') \cap \text{aff epi } f$ точки (x^0, α_0) содержится в $\text{epi } f$ и $(x^0, \alpha_0) \in (\text{epi } f)^{00}$. Точка $(x^0, x^0[n+1])$ является относительно внутренней точкой отрезка $[(x^0, \alpha_0), (x^0, f(x^0))]$ и, следовательно, по

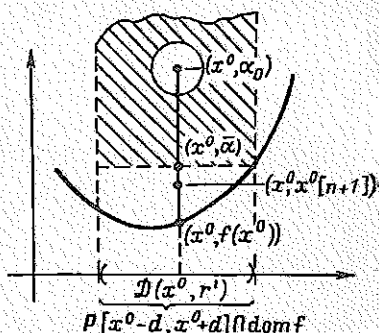


Рис. 12.

теореме о пересечении выпуклого множества лучом (см. п. 3.7) является относительно внутренней точкой множества $\text{epi } f$. ■

8.4. Аффинные касательные к выпуклой функции. Выпуклая функция $f: R^n \rightarrow R$ называется *замкнутой*, если ее надграфик $\text{epi } f$ — замкнутое множество в R^{n+1} . Очевидным примером замкнутой выпуклой функции является аффинная функция $f(x) = ax + \alpha$. Выпуклость и замкнутость аффинной функции следует из того, что ее надграфиком является замкнутое полупространство в R^{n+1} .

Заметим, что из замкнутости множества $\text{epi } f$ еще не следует замкнутость $\text{dom } f$. Так, например, функция $f: R \rightarrow \bar{R}$,

$$f(x) = \begin{cases} 1/x & \text{при } x > 0, \\ +\infty & \text{при } x \leq 0, \end{cases}$$

выпуклая, множество $\text{epi } f$ замкнуто, а $\text{dom } f = \{x \in R : x > 0\}$.

Однако если f — замкнутая выпуклая функция, то при любом $\alpha \in R$ множество $L_\alpha = \{x \in R^n : f(x) \leq \alpha\}$ замкнуто и выпукло. Действительно, выпуклость множества L_α была установлена в 1° п. 8.2, и если $\{x_k\}_{k=1}^\infty$ — последовательность точек множества L_α , сходящаяся к точке x_0 при $k \rightarrow \infty$, то $\{(x_k, \alpha)\}_{k=1}^\infty$ — последовательность точек $\text{epi } f$, сходящаяся к (x_0, α) . Так как множество $\text{epi } f$ замкнуто, то $(x_0, \alpha) \in \text{epi } f$, $f(x_0) \leq \alpha$ и $x_0 \in L_\alpha$. ■

Если задано некоторое семейство $\{f_\lambda, \lambda \in \Lambda\}$ определенных на R^n выпуклых функций f_λ , то функция $f(x) = \sup_{\lambda \in \Lambda} f_\lambda(x)$ из R^n

в $\bar{R}$ является выпуклой. Если к тому же каждая из функций f_λ — замкнутая выпуклая функция, то и f — замкнутая выпуклая функция. В частности, f — замкнутая выпуклая функция из R^n в $\bar{R}$, если $f_\lambda(x) = a_\lambda x + \alpha_\lambda$ — аффинная функция на R^n при каждом $\lambda \in \Lambda$.

Это утверждение следует из того, что $\text{epi } f = \bigcap_{\lambda \in \Lambda} \text{epi } f_\lambda$, так что $\text{epi } f$ — выпуклое множество, если каждое из множеств $\text{epi } f_\lambda$, $\lambda \in \Lambda$, выпукло (см. п. 3.1), и $\text{epi } f$ — замкнутое множество, если каждое из множеств $\text{epi } f_\lambda$, $\lambda \in \Lambda$, замкнуто. ■

В некотором смысле обратной к этому утверждению является следующая теорема.

Теорема. Пусть $f: R^n \rightarrow \bar{R}$ — замкнутая выпуклая функция. Тогда существует такое семейство $\{f_\lambda, \lambda \in \Lambda\}$ аффинных функций $f_\lambda(x) = a_\lambda x + \alpha_\lambda$, $\lambda \in \Lambda$, что $f(x) = \sup_{\lambda \in \Lambda} \{f_\lambda(x)\}$ для любого $x \in \text{dom } f$.

Доказательству этой теоремы предположим две леммы, имеющие и самостоятельное значение.

Лемма 1. Пусть f — выпуклая функция и x_0 — относительно внутренняя точка $\text{dom } f$. Тогда существует такая аффинная функция $\varphi_{x_0}(x) = a_{x_0}x + \alpha_{x_0}$, что $\varphi_{x_0}(x_0) = f(x_0)$ и $\varphi_{x_0}(x) \leq f(x)$ для любого $x \in R^n$.

Доказательство. Точка $(x_0, f(x_0)) \in \tilde{\partial} \text{epi } f$ и по теореме об опорной гиперплоскости (см. п. 4.4) существует гиперплоскость в R^{n+1} , опорная к $\text{epi } f$ в точке $(x_0, f(x_0))$ и не пересекающаяся с $(\text{epi } f)^{00}$. Пусть уравнение этой гиперплоскости $a'x + \gamma'x[n+1] = \alpha'$, где $a' \in R^n$, $\gamma' \in R$, $(a', \gamma') \neq 0$, и пусть

$$a'x_0 + \gamma'f(x_0) = \alpha', \quad (4)$$

$a'x + \gamma'x[n+1] \geq \alpha'$ для любой точки $(x, x[n+1]) \in \text{epi } f$,
 $a'x + \gamma'x[n+1] > \alpha'$ для любой точки $(x, x[n+1]) \in (\text{epi } f)^{00}$.
 В частности,

$$a'x_0 + \gamma'x[n+1] > \alpha' \quad (5)$$

при $x[n+1] > f(x_0)$ (см. теорему п. 8.3). Тогда $\gamma' > 0$, поскольку при $\gamma' = 0$ соотношения (4) и (5) противоречивы, а при $\gamma' < 0$ можно выбрать столь большое $x[n+1]$, что нарушится неравенство (5). Положим $\varphi_{x_0}(x) = (-1/\gamma')a'x + \alpha'/\gamma' = a_{x_0}x + \alpha_{x_0}$, где $a_{x_0} = (-1/\gamma')a'$, $\alpha_{x_0} = \alpha'/\gamma'$. Из равенства (4) следует $f(x_0) = (-1/\gamma')a'x_0 + \alpha'/\gamma' = \varphi_{x_0}(x_0)$. Если $x \in \text{dom } f$, то $(x, f(x)) \in \text{epi } f$. Так что $a'x + \gamma'f(x) \geq \alpha'$ и $f(x) \geq (-1/\gamma')a'x + \alpha'/\gamma' = \varphi_{x_0}(x)$. При $x \in \text{dom } f$ имеем $\varphi_{x_0}(x) < +\infty = f(x)$. ■

Лемма 2. Пусть f — замкнутая выпуклая функция, $x_0 \in \tilde{\partial} \text{dom } f$ и $x_0 \in \text{dom } f$. Тогда для любого $\varepsilon > 0$ существует такая аффинная функция $\varphi_{x_0, \varepsilon}(x) = a_{x_0, \varepsilon}x + \alpha_{x_0, \varepsilon}$, что $\varphi_{x_0, \varepsilon}(x_0) > f(x_0) - \varepsilon$ и $\varphi_{x_0, \varepsilon}(x) < f(x)$ для любого $x \in R^n$.

Доказательство. Рассмотрим в R^{n+1} множества $\{(x_0, f(x_0) - \varepsilon)\}$ и $\text{epi } f$. Это непустые непересекающиеся замкнутые выпуклые множества, одно из которых компактно. По теореме о строгой отделимости (см. п. 4.5) существует гиперплоскость в R^{n+1} , строго разделяющая эти множества. Пусть ее уравнение $a'x + \gamma'x[n+1] = \alpha'$, где $a' \in R^n, \gamma' \in R, (a', \gamma') \neq 0$, и пусть

$$a'x_0 + \gamma'(f(x_0) - \varepsilon) < \alpha', \quad (6)$$

$a'x + \gamma'x[n+1] > \alpha'$ для любой точки $(x, x[n+1]) \in \text{epi } f$. Из условия $x_0 \in \text{dom } f$ следует, что

$$a'x_0 + \gamma'x[n+1] > \alpha', \quad (7)$$

если $x[n+1] \geq f(x_0)$. Тогда $\gamma' > 0$, поскольку при $\gamma' = 0$ неравенства (6) и (7) противоречивы, а при $\gamma' < 0$ неравенство (7) нарушается при достаточно больших значениях $x[n+1]$. Положим

$$\alpha_{x_0, \varepsilon} = (-1/\gamma')\alpha', \quad \alpha_{x_0, \varepsilon} = \alpha'/\gamma' \quad \text{и} \quad \varphi_{x_0, \varepsilon}(x) = \alpha_{x_0, \varepsilon}x + \alpha_{x_0, \varepsilon}.$$

Тогда из (6) следует неравенство $\varphi_{x_0, \varepsilon}(x_0) > f(x_0) - \varepsilon$. Если $x \in \text{dom } f$, то $(x, f(x)) \in \text{epi } f$. Значит, $a'x + \gamma'f(x) > \alpha'$, так что $f(x) > \varphi_{x_0, \varepsilon}(x)$. При $x \in \text{dom } f, \varphi_{x_0, \varepsilon}(x) < +\infty = f(x)$. ■

Доказательство теоремы. Если $\text{dom } f = \emptyset$, то $f(x) = \sup_{m \in N} \{\varphi_m(x) \equiv m\} \equiv +\infty$.

Пусть $\text{dom } f \neq \emptyset$. Построим следующее семейство аффинных функций: для каждой точки $x_0 \in (\text{dom } f)^{00}$ возьмем аффинную функцию φ_{x_0} , существование которой утверждалось в лемме 1; для каждой точки $x_0 \in \text{dom } f \cap \tilde{\text{dom}} f$ и каждого натурального числа m возьмем аффинную функцию $\varphi_{x_0, 1/m}$, существование которой утверждалось леммой 2. Для любого $x \in R^n$ положим

$$g(x) = \sup \{\varphi_{x_0}(x), x_0 \in (\text{dom } f)^{00}; \varphi_{x_0, 1/m}(x), x_0 \in \text{dom } f \cap \tilde{\text{dom}} f, m \in N\}.$$

Тогда очевидно $g(x) \leq f(x)$ для любого $x \in \text{dom } f$, $g(x) \geq \varphi_{x_0}(x) = f(x)$ для любого $x \in (\text{dom } f)^{00}$, $g(x) > \varphi_{x_0, 1/m}(x) > f(x) - 1/m$ для любого $x \in \text{dom } f \cap \tilde{\text{dom}} f$ и любого $m \in N$. Следовательно, $f(x) = g(x)$ для любого $x \in \text{dom } f$. ■

Из доказанной теоремы следует возможность аппроксимации замкнутой выпуклой функции кусочно-аффинными функциями. Эти функции являются поточечными максимумами конечных подсистем семейства аффинных функций.

8.5. Непрерывность выпуклой функции. Условия выпуклости функции оказываются столь сильными, что из них следуют некоторые свойства непрерывности и дифференцируемости. Точные

формулировки соответствующих утверждений будут приведены ниже.

Пусть $f: R^n \rightarrow \bar{R}$ — выпуклая функция, не равная тождественно $+\infty$. Рассмотрим функцию $\tilde{f}: \text{dom } f \rightarrow R$, $\tilde{f}(x) = f(x)$ для $x \in \text{dom } f$ — ограничение функции f на множестве $\text{dom } f$.

Теорема. *Функция $\tilde{f}$ непрерывна в каждой точке $(\text{dom } f)^{\text{oo}}$.*

Доказательство. Пусть $x_0 \in (\text{dom } f)^{\text{oo}}$. Достаточно показать, что для любой последовательности $\{x_k\}_{k=1}^{\infty}$, $x_k \in \text{dom } f$, $x_k \rightarrow x_0$ при $k \rightarrow \infty$, имеет место $\tilde{f}(x_k) \rightarrow \tilde{f}(x_0)$ при $k \rightarrow \infty$.

Пусть φ_{x_0} — аффинная функция, существование которой установлено в лемме 1. Тогда $\varphi_{x_0}(x_k) \leq f(x_k) = \tilde{f}(x_k)$ при $k \in N$, и, следовательно,

$$\tilde{f}(x_0) = f(x_0) = \varphi_{x_0}(x_0) = \lim_{k \rightarrow \infty} \varphi_{x_0}(x_k) = \lim_{k \rightarrow \infty} \varphi_{x_0}(x_k) \leq \lim_{k \rightarrow \infty} \tilde{f}(x_k) \quad (8)$$

(через $\lim_{k \rightarrow \infty} \alpha_k$ обозначается нижний предел числовой последовательности $\{\alpha_k\}_{k=1}^{\infty}$). Вместе с тем всякое число μ , являющееся частичным пределом последовательности $\{\tilde{f}(x_k)\}_{k=1}^{\infty} = \{f(x_k)\}_{k=1}^{\infty}$, не превосходит $\tilde{f}(x_0) = f(x_0)$, так как в противном случае соответствующая подпоследовательность последовательности $\{(x_k, f(x_k))\}_{k=1}^{\infty}$ относительно граничных точек множества $\text{epi } f$ сходилась бы к относительно внутренней точке (x_0, μ) этого множества (см. теорему п. 8.3). Поэтому

$$\overline{\lim}_{k \rightarrow \infty} \tilde{f}(x_k) = \overline{\lim}_{k \rightarrow \infty} f(x_k) \leq f(x_0) = \tilde{f}(x_0) \quad (9)$$

(через $\overline{\lim}_{k \rightarrow \infty} \alpha_k$ обозначается верхний предел числовой последовательности $\{\alpha_k\}_{k=1}^{\infty}$). Из соотношений (8) и (9) следуют существование предела последовательности $\{\tilde{f}(x_k)\}_{k=1}^{\infty}$ при $k \rightarrow \infty$ и равенство $\lim_{k \rightarrow \infty} \tilde{f}(x_k) = f(x_0)$. ■

Следствие. *Выпуклая функция f непрерывна в каждой внутренней точке $\text{dom } f$.*

8.6. Производная выпуклой функции по направлению. Пусть $f: R^n \rightarrow \bar{R}$ — выпуклая функция. Для любого $x_0 \in \text{dom } f$ и любого $s \in R^n$ существует предел (конечный, равный $+\infty$ или $-\infty$)

$$\lim_{\lambda \downarrow 0} (1/\lambda) [f(x_0 + \lambda s) - f(x_0)] = f'(x_0, s) \quad (10)$$

— производная функции f в точке $x_0 \in \text{dom } f$ по направлению s (см. п. 2.10).

Доказательство. Положим $h(\lambda) = (1/\lambda) [f(x_0 + \lambda s) - f(x_0)]$. Докажем, что h — монотонно неубывающая функция

от λ на множестве $\{\lambda \in R: \lambda > 0\}$, откуда и будет следовать существование предела (10).

Пусть $0 < \lambda_1 < \lambda_2$. Тогда $0 < \lambda_1/\lambda_2 < 1$ и в силу выпуклости функций f (см. п. 8.1)

$$f(x_0 + \lambda_1 s) = f\left(\left(1 - \frac{\lambda_1}{\lambda_2}\right)x_0 + \frac{\lambda_1}{\lambda_2}(x_0 + \lambda_2 s)\right) \leq \left(1 - \frac{\lambda_1}{\lambda_2}\right)f(x_0) + \frac{\lambda_1}{\lambda_2}f(x_0 + \lambda_2 s) = f(x_0) + \frac{\lambda_1}{\lambda_2}[f(x_0 + \lambda_2 s) - f(x_0)].$$

Преобразуя это неравенство, получаем

$$(1/\lambda_1)[f(x_0 + \lambda_1 s) - f(x_0)] \leq (1/\lambda_2)[f(x_0 + \lambda_2 s) - f(x_0)],$$

т. е. $h(\lambda_1) \leq h(\lambda_2)$. ■

З а м е ч а н и е. Из приведенного доказательства следует, что

$$f'(x_0, s) = \inf_{\lambda > 0} h(\lambda) = \inf_{\lambda > 0} (1/\lambda)[f(x_0 + \lambda s) - f(x_0)].$$

При каждом фиксированном $x_0 \in \text{dom } f$ функция $f'_{x_0}(s) = f'(x_0, s)$ есть функция от s , $f'_{x_0}: R^n \rightarrow R$.

Если f — непрерывно дифференцируемая функция, то $f'_{x_0}(s) = \nabla f(x_0)s$, где $\nabla f(x_0) = \left(\frac{\partial f(x_0)}{\partial x[1]}, \frac{\partial f(x_0)}{\partial x[2]}, \dots, \frac{\partial f(x_0)}{\partial x[n]}\right)$ — вектор-градиент функции f в точке x_0 .

Действительно, положим $\varphi(\lambda) = f(x_0 + \lambda s)$. Очевидно, что если существует $\varphi'(0)$, то $f'(x_0, s) = \varphi'(0)$. Но в соответствии с правилом дифференцирования сложной функции

$$\varphi'(\lambda) = \nabla f(x_0 + \lambda s)s, \quad (11)$$

и, следовательно, $\varphi'(0) = \nabla f(x_0)s = f'(x_0, s)$. ■

8.7. Эквивалентные условия выпуклости. Для случая, когда f — конечная непрерывно дифференцируемая функция на R^n , можно сформулировать эквивалентные условия выпуклости функции f , полезные как для теоретических исследований, так и для приложений.

Пусть $f: R^n \rightarrow R$ — непрерывно дифференцируемая функция. Для фиксированных $x \in R^n$, $s \in R^n$ определим функции

$$\varphi: R \rightarrow R, \quad \varphi(\lambda) = \varphi(x, s, \lambda) = f(x + \lambda s),$$

$$\psi: R \rightarrow R, \quad \psi(\lambda) = \psi(x, s, \lambda) = \nabla f(x + \lambda s)s.$$

Мы уже отмечали (см. (11) п. 8.6), что $\psi(\lambda) = \varphi'(\lambda)$.

Теорема. Следующие условия эквивалентны:

1) Функция f выпукла на R^n .

2) Для произвольных x_1, x_2 из R^n справедливо неравенство

$$f(x_2) - f(x_1) \geq \nabla f(x_1)(x_2 - x_1).$$

3) Функция $\psi(\lambda) = \nabla f(x + \lambda s)s$, $\lambda \in R$, монотонно не убывает по λ при любых фиксированных x и s из R^n .

Если функция f дважды непрерывно дифференцируема на R^n , то каждое из условий 1)–3) эквивалентно следующему условию, в котором $G(x)$ — матрица порядка n с элементами

$$g(x)[i, j] = \frac{\partial^2 f(x)}{\partial x[i] \partial x[j]}, \quad i \in 1:n, \quad j \in 1:n$$

(матрица Гессе).

4) Матрица $G(x)$ неотрицательно определена при любом $x \in R^n$.

Доказательство. 1) $\Rightarrow$ 2) Пусть $f: R^n \rightarrow R$ — выпуклая функция. Тогда для любых $x_1 \in R^n$, $x_2 \in R^n$ и любого $\lambda \in (0, 1)$

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2)$$

(см. п. 8.1), откуда получаем

$$(1/\lambda)[f(x_1 + \lambda(x_2 - x_1)) - f(x_1)] \leq f(x_2) - f(x_1).$$

При $\lambda \rightarrow 0$ левая часть последнего неравенства стремится к производной $\varphi'(0)$ функции $\varphi(\lambda) = \varphi(x_1, x_2 - x_1, \lambda)$, а правая не зависит от λ . Но $\varphi'(0) = \psi(0) = \psi(x_1, x_2 - x_1, 0) = \nabla f(x_1)(x_2 - x_1)$. Поэтому $\varphi'(0) = \nabla f(x_1)(x_2 - x_1) \leq f(x_2) - f(x_1)$. ■

2) $\Rightarrow$ 3) Пусть для любых $x_1 \in R^n$, $x_2 \in R^n$ справедливо неравенство

$$f(x_2) - f(x_1) \geq \nabla f(x_1)(x_2 - x_1). \quad (12)$$

Покажем, что для $x \in R^n$, $s \in R^n$, λ_1 и λ_2 из R , $\lambda_1 < \lambda_2$ выполняется неравенство $\psi(\lambda_1) = \psi(x, s, \lambda_1) \leq \psi(\lambda_2) = \psi(x, s, \lambda_2)$. Подставляя в (12) сначала $x_1 = x + \lambda_1 s$, $x_2 = x + \lambda_2 s$, а потом $x_1 = x + \lambda_2 s$, $x_2 = x + \lambda_1 s$, получаем

$$f(x + \lambda_2 s) - f(x + \lambda_1 s) \geq (\lambda_2 - \lambda_1) \nabla f(x + \lambda_1 s)s, \quad (13)$$

$$f(x + \lambda_1 s) - f(x + \lambda_2 s) \geq (\lambda_1 - \lambda_2) \nabla f(x + \lambda_2 s)s. \quad (14)$$

При $\lambda_1 < \lambda_2$ из неравенства (13) имеем

$$\psi(\lambda_1) = \nabla f(x + \lambda_1 s)s \leq \frac{1}{\lambda_2 - \lambda_1} [f(x + \lambda_2 s) - f(x + \lambda_1 s)],$$

а из неравенства (14) —

$$\psi(\lambda_2) = \nabla f(x + \lambda_2 s)s \geq \frac{1}{\lambda_2 - \lambda_1} [f(x + \lambda_2 s) - f(x + \lambda_1 s)].$$

Следовательно, $\psi(\lambda_1) \leq \psi(\lambda_2)$. ■

3) $\Rightarrow$ 1) Пусть $\psi(\lambda) = \psi(x, s, \lambda)$ — неубывающая функция от λ при любых $x \in R^n$, $s \in R^n$. Покажем, что для любых $x_1 \in R^n$, $x_2 \in R^n$, $\alpha \in (0, 1)$ выполняется неравенство $f((1-\alpha)x_1 + \alpha x_2) \leq (1-\alpha)f(x_1) + \alpha f(x_2)$. Положим $x = x_1$, $s = x_2 - x_1$, и пусть $\mu \in (0, 1)$. Тогда $\mu\alpha < \mu$ и, следовательно, $\psi(\mu\alpha) \leq \psi(\mu)$.

Поэтому

$$\int_0^1 [\psi(\mu) - \psi(\mu\alpha)] d\mu \geq 0.$$

Но $\int_0^1 \psi(\mu) d\mu = \varphi(\mu) \Big|_0^1 = \varphi(1) - \varphi(0) = f(x_1 + (x_2 - x_1)) - f(x_1) = f(x_2) - f(x_1)$, $\int_0^1 \psi(\mu\alpha) d\mu = (1/\alpha) \varphi(\mu\alpha) \Big|_0^1 = (1/\alpha) [\varphi(\alpha) - \varphi(0)] = (1/\alpha) [f(x_1 + \alpha(x_2 - x_1)) - f(x_1)]$. Таким образом, $f(x_2) - f(x_1) - (1/\alpha) f((1-\alpha)x_1 + \alpha x_2) + (1/\alpha) f(x_1) \geq 0$, или $f((1-\alpha)x_1 + \alpha x_2) \leq (1-\alpha)f(x_1) + \alpha f(x_2)$. ■

3) $\Rightarrow$ 4) Если f дважды непрерывно дифференцируема, то $\psi(\lambda)$ — непрерывно дифференцируемая функция. Необходимым и достаточным условием неубывания функции ψ является неотрицательность ее производной ψ' . Но

$$\begin{aligned} \psi'(\lambda) &= \frac{d}{d\lambda} \nabla f(x + \lambda s) s = \frac{d}{d\lambda} \sum_{k=1}^n \frac{\partial f(x + \lambda s)}{\partial x[k]} s[k] = \\ &= \sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2 f(x + \lambda s)}{\partial x[i] \partial x[k]} s[i] s[k]. \end{aligned}$$

Поэтому если при любых $x \in R^n$, $s \in R^n$ функция $\psi'(\lambda) \geq 0$ для любого $\lambda \in R$, то $\psi'(0) \geq 0$. Значит, для каждого $x \in R^n$ квадратичная функция

$$h(s) = \sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2 f(x)}{\partial x[i] \partial x[k]} s[i] s[k] \geq 0$$

при любом $s \in R^n$, т. е. матрица $G(x) = \left(\frac{\partial^2 f(x)}{\partial x[i] \partial x[k]} \right)$ неотрицательно определена для любого $x \in R^n$. Обратно, если матрица $G(x)$ неотрицательно определена при каждом $x \in R^n$, то неотрицательно определена и матрица $G(x + \lambda s)$ при любых $x \in R^n$, $s \in R^n$, $\lambda \in R$, и, следовательно, для любых $x \in R^n$, $s \in R^n$ функция ψ' неотрицательна. ■

Замечание 1. Некоторым обобщением доказанной теоремы является следующее утверждение.

Если функция $f: R^n \rightarrow \bar{R}$ непрерывно дифференцируема на открытом выпуклом множестве $V \subset \text{dom } f$, то следующие условия эквивалентны.

1') Функция f выпукла на V .

2') Для произвольных $x_1 \in V$, $x_2 \in V$ справедливо неравенство

$$f(x_2) - f(x_1) \geq \nabla f(x_1)(x_2 - x_1).$$

3') Функция $\psi(\lambda) = \nabla f(x + \lambda s)s$ монотонно не убывает по λ на (открытом) интервале тех $\lambda \in R$, при которых $x + \lambda s \in V$ для любых фиксированных x и s из R^n .

Если функция f дважды непрерывно дифференцируема на V , то каждое из условий 1') — 3') эквивалентно следующему условию.

4') Матрица $G(x)$ неотрицательно определена при любом $x \in V$.

Доказательство этого утверждения по существу ничем не отличается от доказательства теоремы. Мы предоставляем читателю возможность провести его самостоятельно. — Упражнение 1.38.

В заключение отметим, что эквивалентное условие 4) (4') можно использовать для эффективной проверки выпуклости функции.

Примеры. 1. Пусть $f: R^n \rightarrow R$, $f(x) = x^T D x + c x$, $D = (d[i, j])$ — квадратная симметричная матрица порядка n , $c \in R^n$. Тогда $G(x) \equiv 2D$ и функция f выпукла в том и только том случае, если D — неотрицательно определенная матрица.

2. Пусть $f: R^2 \rightarrow R$, $f(x) = x_1^3 + 2x_1 + x_2$. Здесь $G(x) = \begin{pmatrix} 6x_1 & 0 \\ 0 & 0 \end{pmatrix}$, и легко видеть, что $G(x)$ неотрицательно определена в полуплоскости $\Pi = \{x \in R^2: x_1 > 0\}$. Поэтому f выпукла на Π .

Замечание 2. Из доказанной выше теоремы (см. также замечание 1) очевидно следуют эквивалентные условия вогнутости функции $f: R^n \rightarrow \bar{R}$.

Если $f: R^n \rightarrow \bar{R}$ непрерывно дифференцируема на открытом множестве $V \subset \text{dom } f$, то следующие условия эквивалентны.

1'') Функция f вогнута на V .

2'') Для произвольных $x_1 \in V, x_2 \in V$ справедливо неравенство

$$f(x_2) - f(x_1) \leq \nabla f(x_1)(x_2 - x_1).$$

3'') Функция $\psi(\lambda) = \nabla f(x + \lambda s)s$ монотонно не возрастает по λ на (открытом) интервале тех $\lambda \in R$, при которых $x + \lambda s \in V$ для любых фиксированных $x \in V$ и $s \in R^n$.

Если функция дважды непрерывно дифференцируема на V , то каждое из условий 1'')—3'') эквивалентно следующему условию.

4'') Матрица $G(x)$ неположительно определена при любом $x \in V$.

Проверку эквивалентности оставляем читателю. — Упражнение 1.39.

УПРАЖНЕНИЯ

1.1. Доказать, что если система векторов $a_1, a_2, \dots, a_k$ линейно независима, $a \in R^n$ и $a = \sum_{i=1}^k \alpha_i a_i$, $\alpha_r \neq 0$, то и система векторов $a_1, a_2, \dots, a_{r-1}, a, a_{r+1}, \dots, a_k$ линейно независима.

1.2. См. с. 10.

1.3. Пусть M_1 и M_2 — подмножества пространства R^n . Верно ли, что $\text{aff}(M_1 + M_2) = \text{aff } M_1 + \text{aff } M_2$?

1.4. Пусть $u \in \Pi^+ = \{x \in R^n: ax > \beta\}$, а $v \in \Pi^- = \{x \in R^n: ax < \beta\}$. Найти точку пересечения отрезка $[u, v]$ с гиперплоскостью $H = \{x \in R^n: ax = \beta\}$.

1.5. Показать, что всякое аффинное множество в R^n может быть представлено как пересечение конечного множества гиперплоскостей.

1.6. а) Показать, что для всякого шара $\mathcal{D}(a, r)$ существует такой прямоугольник $P(a-d, a+d)$, что $P(a-d, a+d) \subset \mathcal{D}(a, r)$.

б) Показать, что для всякого прямоугольника $P(a-d, a+d)$ существует такой шар $\mathcal{D}(a, r)$, что $\mathcal{D}(a, r) \subset P(a-d, a+d)$.

1.7. Показать, что $\mathcal{D}(a_1, r_1) + \mathcal{D}(a_2, r_2) = \mathcal{D}(a_1 + a_2, r_1 + r_2)$.

1.8. Показать, что $\partial M \supset \bar{\partial} M$.

1.9. См. с. 25.

1.10. Привести примеры таких множеств A и B , что $A \cap B = \emptyset$ и $\rho(A, B) = 0$.

1.11. Доказать непрерывность функций, приведенных в примерах п. 2.8.

1.12. Пусть функция $f: R^n \rightarrow R$ непрерывна на R^n . Доказать, что для любого α из R множество $L_f(\alpha) = \{x \in R^n: f(x) \leq \alpha\}$ замкнуто.

1.13. См. с. 17.

1.14. а) Является ли выпуклым множество $V \subset R^n$, если для любых точек a и b из V точка $1/2(a+b) \in V$?

б) Является ли выпуклым при выполнении предыдущего условия замкнутое множество V ?

1.15. См. с. 18. 1.16. См. с. 18. 1.17. См. с. 21.

1.18. См. с. 22. 1.19. См. с. 24.

1.20. См. с. 26. 1.21. См. с. 26. 1.22. См. с. 28.

1.23. См. с. 30. 1.24. См. с. 31. 1.25. См. с. 33.

1.26. См. с. 33. 1.27. См. с. 38. 1.28. См. с. 40.

1.29. См. с. 43. 1.30. См. с. 46. 1.31. См. с. 53.

1.32. Пусть $f: R^n \rightarrow R$ и $g: R \rightarrow R$ — выпуклые функции. Будет ли суперпозиция $g \circ f$ выпуклой функцией?

1.33. Пусть $f: R^n \rightarrow R$ — выпуклая функция, а $g: R \rightarrow R$ монотонно возрастающая. Будет ли их суперпозиция $g \circ f$ выпуклой функцией?

1.34. Пусть для функции f , выпуклой и дифференцируемой на R^n , существуют такие x_1 и x_2 из R^n , что $f(x_2) < f(x_1)$. Показать, что $\nabla f(x_1)(x_2 - x_1) < 0$.

1.35. Пусть $\nabla f(x_1)(x - x_1) \geq 0$ для любого $x \in R^n$ и функция f выпукла на R^n . Доказать, что f достигает минимального значения в точке x_1 .

1.36. Пусть функция f непрерывна на R^n и для каждой точки из R^n существует окрестность, в которой f выпукла. Является ли f выпуклой на R^n ?

1.37. См. с. 54. 1.38. См. с. 64. 1.39. См. с. 64.

Глава 2

ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ

Предметом математического программирования являются экстремальные задачи, их свойства и методы решения. Вообще говоря, исследование экстремальных задач составляет содержание широкого класса математических дисциплин. В настоящем курсе мы ограничимся рассмотрением лишь конечномерных экстремальных задач, которые уже традиционно относят к задачам математической экономики. Содержательное описание интересующих нас экстремальных задач будет приведено в данной главе.

§ 1. Определение экстремальной задачи

Под *экстремальной задачей* мы понимаем задачу нахождения наибольшего (наименьшего) значения числовой (вещественной) функции на некотором множестве и хотя бы одной точки этого множества, в которой функция принимает это наибольшее (наименьшее) значение. А именно экстремальной называем следующую задачу:

даны множество $\mathcal{D}$ и функция $f: \mathcal{D} \rightarrow R$; требуется найти такой элемент $\bar{x} \in \mathcal{D}$, что для любого $x \in \mathcal{D}$ выполняется неравенство

$$f(\bar{x}) \geq f(x) \quad (1)$$

— *задача максимизации (задача на max)* или неравенство

$$f(\bar{x}) \leq f(x) \quad (1')$$

— *задача минимизации (задача на min)*.

Экстремальную задачу с областью определения $\mathcal{D}$ и функцией f , которая называется *целевой функцией* задачи, будем обозначать через $(\mathcal{D}, f)$, указывая, если не было предварительно оговорено, какой задачей является $(\mathcal{D}, f)$: задачей макси-

зации или минимизации. Например, запись

$$(\mathcal{D}, f) : f(x) = x_1^2 + x_2^2 \rightarrow \min, \mathcal{D} = \{x \in R^2 : x_1 \geq 0, x_2 \geq 3/2, x_2 - \text{целое}\}$$

означает, что $(\mathcal{D}, f)$ есть задача минимизации функции $f(x) = x_1^2 + x_2^2$ на множестве $\mathcal{D}$ точек пространства R^2 , для которых выполнены условия, указанные в фигурных скобках.

В дальнейшем предполагается, что $\mathcal{D}$ — подмножество пространства R^n при некотором n . Точки пространства R^n будем также называть *планами*, элементы множества $\mathcal{D}$ — *допустимыми планами*. Допустимый план $\bar{x}$, удовлетворяющий соотношению (1) в задаче максимизации или (1') в задаче минимизации, называется *оптимальным планом (решением задачи)*.

Экстремальную задачу (R^n, f) называют *задачей безусловной оптимизации*.

Традиционный способ задания множества $\mathcal{D}$ допустимых планов состоит в его определении как множества решений некоторой системы уравнений и неравенств. Предполагается, что задана система функций $f_i : R^n \rightarrow R$, $i \in 1:m$, множество индексов $1:m$ разбито на подмножества индексов уравнений I_y и индексов неравенств I_n ($I_y \cup I_n = 1:n$, $I_y \cap I_n = \emptyset$) и

$$\mathcal{D} = \{x \in R^n : f_i(x) \leq 0, i \in I_n; f_i(x) = 0, i \in I_y\}.$$

Система уравнений и неравенств, определяющих множество $\mathcal{D}$, называется *системой ограничений (условий) задачи*, или просто *системой ограничений (условий)*, а каждое из уравнений $f_i(x) = 0$, $i \in I_y$, и неравенств $f_i(x) \leq 0$, $i \in I_n$, — *ограничением (условием) задачи*, или просто *ограничением (условием)*. Очевидно, что каждое ограничение $f_i(x) = 0$, $i \in I_y$, можно заменить равносильной ему системой неравенств $f_i(x) \leq 0$, $-f_i(x) \leq 0$. Однако иногда полезно особо выделить ограничения в форме уравнений, а также условия неотрицательности переменных: $x[j] \geq 0$, $j \in J$, где J — некоторое подмножество множества $1:n$. Запись « $x[j]$ — ц. ч.» означает, что j -я компонента вектора x должна быть целочисленной. Иногда j -я компонента вектора x может принимать только два значения, например, 0 или 1. Такие переменные называются *булевыми*.

Отметим также, что одно и то же множество может быть задано разными, но, естественно, равносильными системами уравнений и неравенств. Например, нетрудно проверить, что $\{x \in R^3 : x_1^2 + x_2^2 \leq 1, x_1^2 + x_2^2 - 4x_1 \leq -3, x_3 \geq 0\}$ и $\{x \in R^3 : x_1 + x_2 = 1, x_1 - x_2 = 1, x_3 \geq 0\}$ определяют в R^3 луч $\{x \in R^3 : x = (1, 0, 0) + \lambda(0, 0, 1), \lambda \geq 0\}$. — Упражнение 2.1.

В множестве $\mathcal{D}$ может не существовать оптимального плана, например, в случае неограниченности целевой функции на $\mathcal{D}$ или незамкнутости множества $\mathcal{D}$. Так, в экстремальных задачах $(\mathcal{D}, f) : f(x) = x \rightarrow \max, \mathcal{D} = \{x \in R, x > 0\}$ и $(\mathcal{D}, -f)$ нет ре-

шений именно по указанным причинам. Более того, при формальном описании множества $\mathcal{D}$ оно может оказаться пустым. Напротив, множество оптимальных планов задачи может состоять из более чем одного элемента. Поэтому предложение решить экстремальную задачу означает: указать по крайней мере один оптимальный план или установить, что задача не имеет решения.

§ 2. Некоторые классы экстремальных задач

Предметом исследования настоящего курса математического программирования являются следующие классы экстремальных задач.

2.1. Задачи выпуклого программирования. Экстремальные задачи максимизации вогнутой функции на выпуклом множестве и задачи минимизации выпуклой функции на выпуклом множестве называются *задачами выпуклого программирования*. В соответствии со сказанным в предыдущем параграфе выберем следующую форму записи задачи выпуклого программирования:

$$(\mathcal{D}, f): f(x) \rightarrow \max, \mathcal{D} = \{x \in R^n: f_i(x) \leq 0, i \in I_n;\}$$

$$f_i(x) = a_i x - b[i] = 0, i \in I_y; x[j] \geq 0, j \in J\},$$

где f — вогнутая функция; I_n и I_y — непересекающиеся множества индексов, объединение которых равно $1:m$; f_i — выпуклые функции при $i \in I_n$; $a_i \in R^n$, $b[i] \in R$ при $i \in I_y$; $J \subset 1:n$. Другая форма записи задачи выпуклого программирования: $(\mathcal{D}, f): f(x) \rightarrow \min, x \in \mathcal{D}$, где f — выпуклая функция, а множество $\mathcal{D}$ определяется так, как это сделано выше.

В указанных предположениях выпуклость множества $\mathcal{D}$ следует из выпуклости функций f_i для $i \in I_n$. — Упражнение 2.2 (см. п. 1.8.4). Если к тому же все функции f_i , $i \in I_n$, непрерывны, то $\mathcal{D}$ — замкнутое выпуклое множество. — Упражнение 2.3. (см. также п. 1.8.4). Напомним, что выпуклые функции f_i , $i \in I_n$, непрерывны на R^n , если $\text{dom } f_i = R^n$ (см. п. 1.8.5).

Если система уравнений

$$a_i x - b[i] = 0, i \in I_y, \quad (1)$$

совместна, то она определяет аффинное множество в R^n , в котором остальные ограничения «высекают» множество $\mathcal{D}$. В этом случае естественно предполагать, что система векторов $\{a_i, i \in I_y\}$ линейно независима, иначе, исключив уравнения, являющиеся линейными комбинациями остальных, получим систему, равносильную системе (1). В ряде случаев оказывается полезным выделить в системе ограничений $\{f_i(x) \leq 0, i \in I_n\}$ аффинные. Множество индексов ограничений, не являющихся аффинными, обозначим через I_v ($I_v \subset I_n$). Тогда $f_i(x) = a_i x -$

— $b[i]$, $a_i \in R^n$, $b[i] \in R$ для каждого $i \in I_n \setminus I_b$ и множество $\mathcal{D}$ записывается в виде

$$\mathcal{D} = \{x \in R^n : f_i(x) \leq 0, i \in I_b; f_i(x) = a_i x - b[i] \leq 0, i \in I_n \setminus I_b; \\ f_i(x) = a_i x - b[i] = 0, i \in I_y; x[j] \geq 0, j \in J\}.$$

Множество $\mathcal{D}$ может оказаться выпуклым, даже если функции f_i , $i \in I_n$, не являются выпуклыми, а функции f_i , $i \in I_y$, не являются аффинными. В частности, множество $\mathcal{D}$ всегда выпукло, если f_i , $i \in I_y$, — аффинные, а f_i , $i \in I_n$, — квазивыпуклые функции (функция $g: R^n \rightarrow R$ называется *квазивыпуклой*, если $g((1-\lambda)a + \lambda b) \leq \max\{g(a), g(b)\}$ для любых $a \in R^n$, $b \in R^n$ и $\lambda \in [0, 1]$).

2.2. Задачи линейного программирования. Экстремальные задачи максимизации или минимизации линейной функции на многогранном выпуклом множестве (см. п. 1.5.4) называются *задачами линейного программирования*. В соответствии с определением многогранного выпуклого множества оно может быть задано некоторой системой линейных уравнений и неравенств. Поэтому задача линейного программирования может быть записана в виде

$$(\mathcal{D}, f) : f(x) = cx \rightarrow \max(\min), \mathcal{D} = \{x \in R^n : a_i x \leq b[i], i \in I_n; \\ a_i x = b[i], i \in I_y; x[j] \geq 0, j \in J\},$$

где $c \in R^n$, $I_n \subset 1:m$, $I_y \subset 1:m$, $I_n \cup I_y = 1:m$, $I_n \cap I_y = \emptyset$, $J \subset 1:n$, $b[i] \in R$, $a_i \in R^n$.

Заметим, что при формальном задании множества $\mathcal{D}$ некоторой системой линейных уравнений и неравенств эта система может оказаться несовместной ($\mathcal{D} = \emptyset$). Однако независимо от того, совместна или несовместна эта система (напомним, что пустое множество не является многогранным выпуклым множеством), экстремальную задачу $(\mathcal{D}, f)$ указанного выше вида будем называть *общей задачей линейного программирования*.

Пусть $(\mathcal{D}, f)$ — общая задача линейного программирования и $A — m \times n$ матрица, строками которой являются векторы a_i , $i \in 1:m$. Обозначим через A_n и A_y матрицы, строками которых являются векторы a_i с $i \in I_n$ и $i \in I_y$ соответственно, через b_n и b_y — векторы, координаты которых $b[i]$ с $i \in I_n$ и $i \in I_y$ упорядочены так же, как строки матриц A_n и A_y . Тогда

$$\mathcal{D} = \{x \in R^{n+} : A_n x \leq b_n, A_y x = b_y\}$$

— матричная запись множества допустимых планов общей задачи линейного программирования (напомним, что через R^{n+} обозначается множество $\{x \in R^n : x[j] \geq 0, j \in J\}$, а $R^{n+} = \{x \in R^n : x \geq 0\}$, см. п. 1.2.1).

Если $I_y = 1:m$ ($I_n = \emptyset$) и $J = 1:n$, то задача линейного программирования называется *канонической*. Каноническая задача линейного программирования записывается в виде

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max(\min), \quad \mathcal{D} = \{x \in R^n : Ax = b, x \geq 0\}.$$

Очевидно, что задачи линейного программирования являются задачами выпуклого программирования.

2.3. Задачи квадратичного программирования. Экстремальные задачи вида

$$(\mathcal{D}, f): f(x) = cx + x^* D x \rightarrow \max(\min),$$

$$\mathcal{D} = \{x \in R^n : a_i x - b[i] \leq 0, i \in I_n; a_i x - b[i] = 0, i \in I_y; x[j] \geq 0, j \in J\},$$

где D — симметричная квадратная матрица порядка n , неположительно определенная в случае задачи максимизации и неотрицательно определенная в случае задачи минимизации, называются *задачами квадратичного программирования*. Читателю предлагается показать, что задачи квадратичного программирования являются задачами максимизации вогнутой квадратичной функции или минимизации выпуклой квадратичной функции на многогранном выпуклом множестве (если $\mathcal{D} \neq \emptyset$), а следовательно, задачами выпуклого программирования. — Упражнение 2.4 (см. п. 1.8.7).

2.4. Задачи дискретного программирования. Экстремальные задачи максимизации или минимизации некоторой функции на конечном или счетном множестве $\mathcal{D}$ называются *дискретными экстремальными задачами*, или *задачами дискретного программирования*.

Пусть $\mathcal{D} \subset R^n$ и $(\mathcal{D}, f)$ — некоторая экстремальная задача. Рассмотрим множество $\mathcal{D}_\mathbb{Z}$ тех точек множества $\mathcal{D}$, координаты которых целые числа. Экстремальную задачу $(\mathcal{D}_\mathbb{Z}, f)$ назовем *задачей целочисленного программирования*. При этом если $(\mathcal{D}, f)$ — задача выпуклого программирования, то $(\mathcal{D}_\mathbb{Z}, f)$ — *целочисленная задача выпуклого программирования*; если $(\mathcal{D}, f)$ — задача линейного программирования, то $(\mathcal{D}_\mathbb{Z}, f)$ — *целочисленная задача линейного программирования*, и т. п. Иногда условие целочисленности накладывают не на все компоненты вектора $x \in \mathcal{D}$, а только на те, индексы которых принадлежат некоторому множеству $J_\mathbb{Z} \subset 1:n$. Такие экстремальные задачи называют соответственно *задачами частично целочисленного программирования*, или *смешанно-целочисленного программирования*, *частично целочисленными задачами выпуклого программирования*, *частично целочисленными задачами линейного программирования* и т. п.

Если на координаты вектора $x \in \mathcal{D}$ наложено дополнительное ограничение $x[j] = 0$ или $x[j] = 1$, то соответствующие экстремальные задачи называются *задачами с булевыми переменными*.

§ 3. Экономико-математическое моделирование. Примеры экстремальных задач

3.1. Экономико-математическое моделирование. В настоящем параграфе будут приведены примеры экстремальных задач. Все они возникают на основе экономико-математических моделей некоторых экономических ситуаций и в дальнейшем будут изучаться с целью указания эффективных вычислительных методов и качественного анализа, результаты которого можно интерпретировать экономически.

Однако тот факт, что экстремальные задачи порождены экономическими ситуациями, нисколько не умаляет их общности. Можно привести много примеров из самых неожиданных областей, где возникают задачи, подобные приведенным в этом параграфе или сводимые к ним, так что сфера приложений теории математического программирования достаточно широка. В этой книге экономическим приложениям уделяется особое внимание, поскольку учебное пособие предназначается прежде всего для специалистов по экономической кибернетике.

Несколько слов о математическом моделировании в экономике. В настоящее время число экономико-математических моделей стремительно увеличивается. Все они аппроксимируют (с различной степенью точности) многовариантные ситуации, в которых необходимо указать наилучшее решение с точки зрения определенного критерия, выбор которого составляет специальную задачу.

Сложность экономики как объекта исследования (на различных уровнях) не вызывает сомнений. Поэтому моделированию должен предшествовать тщательный анализ объекта моделирования, выясняющий его характеристики, важнейшие функциональные зависимости, возможность получения необходимой информации и степень ее достоверности. В результате возникают различные предположения и упрощения, продиктованные прежде всего уровнем доступности информации и возможностями математического аппарата, привлекаемого для изучения модели, и, естественно, ухудшающие точность аппроксимации, что является превосходным поводом для критики. И если все же эти предположения признаются допустимыми, появляется возможность формализации рассматриваемой экономической ситуации, т. е. моделирования.

К сожалению, сам процесс моделирования еще плохо поддается формализации, и хотя подобные попытки имели место, их вряд ли можно признать успешными. Эту сторону моделирования мы не рассматриваем.

Далее везде приведены уже формализованные описания ситуаций (большой частью экономических) при «упрощающих» предположениях. Их анализ, о котором шла речь выше, опущен.

Понятно, что любое разумное приложение неизбежно повлечет за собой множество специфических вопросов. Мы не считаем, что их успешное решение легче (если такое сравнение уместно) чисто математических проблем, которые излагаются ниже.

Мы будем стараться разделять процесс составления модели и формулировку экстремальной задачи, возникающей на ее основе. Это не случайно, поскольку нам кажется неправомерным отождествление модели и экстремальной задачи, которое бытует в литературе. Ведь на основе модели, вообще говоря, можно формулировать (или не формулировать) различные экстремальные задачи.

Некоторые экстремальные задачи, вводимые в настоящем параграфе, будут подробно изучены в последующих главах книги. Предполагается, что их название и терминология соответствуют той ситуации, в которой они возникают: транспортная, станковая и т. п. Читателя не должно смущать, что он встретится с аналогичными задачами в таких конкретных приложениях, где о станках, транспорте и т. д. не будет сказано ни слова.

Отметим, что мы рассматриваем здесь только *статические модели* и соответствующие им экстремальные задачи, т. е. модели, все параметры и переменные которых не зависят от времени. Именно такие экономико-математические модели были исторически первыми и наиболее часто в настоящее время применяются в экономических исследованиях. Однако существуют экономические ситуации и процессы, описание которых статическими моделями выглядит неестественно, и учет фактора времени является необходимым по существу. В этом случае используются *динамические модели*, где либо переменные, либо некоторые параметры зависят от времени. Соответствующее обобщение некоторых статических моделей (например, из приводимых ниже) достаточно естественно. Динамические модели и соответствующие экстремальные задачи традиционно являются предметом исследования специального раздела математического программирования — *динамического программирования*.

3.2. Транспортная задача в матричной постановке. Рассмотрим m пунктов производства (или хранения) и n пунктов потребления некоторого однородного продукта (такими продуктами являются, например, нефть, уголь, цемент, песок, щебень и т. п.). Занумеруем пункты производства индексами $i \in 1:m$, а пункты потребления — индексами $j \in 1:n$. Обозначим через $a[i]$ объем производства (запас) продукта в i -м пункте производства и через $b[j]$ — потребность в продукте в j -м пункте потребления. Предполагается, что задана $m \times n$ матрица $C = (c[i, j])$; ее элемент $c[i, j]$ — стоимость перевозки единицы продукта из i -го пункта производства в j -й пункт потребления. Предполагается также выполнение следующих условий:

- 1) из каждого пункта производства должна быть вывезена вся имеющаяся в этом пункте продукция;
- 2) потребность в продукте в пунктах потребления должна быть удовлетворена;
- 3) затраты на объем перевозок во всей системе равны сумме затрат на отдельных ее участках, а на каждом участке затраты на перевозку пропорциональны объему перевозки;
- 4) суммарный объем производства равен общему объему потребления, т. е.

$$\sum_{i=1}^m a[i] = \sum_{j=1}^n b[j].$$

Приступим к построению математической модели. Она будет описывать все возможные варианты транспортировки (с указанием объемов перевозок, а также «закрепления» пунктов потребления за пунктами производства) и позволит для каждого варианта вычислить затраты на его осуществление.

Обозначим через $x[i, j]$ объем перевозок из i -го пункта производства в j -й пункт потребления. Тогда $\sum_{i=1}^m x[i, j]$ — количество продукта, которое поступает из всех пунктов производства в j -й пункт потребления, а $\sum_{j=1}^n x[i, j]$ — количество вывозимого груза из i -го пункта производства. Следовательно, каждый вариант транспортировки задается $m \times n$ матрицей $x = (x[i, j])$:

$$x[i, j] \geq 0, \quad i = 1:m, \quad j = 1:n, \quad (1)$$

$$\sum_{j=1}^n x[i, j] = a[i], \quad i = 1:m, \quad (2)$$

$$\sum_{i=1}^m x[i, j] = b[j], \quad j = 1:n. \quad (3)$$

Если $x = (x[i, j])$ — допустимый вариант транспортировки, т. е. удовлетворяет условиям (1) — (3), то

$$f(x) = \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j] \quad (4)$$

— стоимость его реализации. Условия (1) — (4) задают *транспортную модель*. На ее основе можно формулировать различные экстремальные задачи.

Чаще всего рассматривается следующая задача:

$$(\mathcal{D}, f) : f(x) = \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j] \rightarrow \min,$$

$$\mathcal{D} = \{x \in R^{mn} : \sum_{j=1}^n x[i, j] = a[i], \quad i = 1:m; \sum_{i=1}^m x[i, j] = b[j], \\ j = 1:n; x \geq 0\},$$

которую в дальнейшем будем называть *транспортной задачей в матричной постановке*. Легко видеть, что $(\mathcal{D}, f)$ является задачей линейного программирования.

3.3. Транспортная задача в сетевой постановке [19, 21, 41]. Имеется m пунктов (здесь пункты производства и потребления

не различаются), соединенных сетью дорог с односторонним движением, состоящей из n участков. Для каждого пункта $i \in 1:m$ задано число $b[i]$. Если $b[i] > 0$, то в i -й пункт необходимо привезти $b[i]$ единиц груза; если $b[i] < 0$, то из i -го пункта нужно вывезти $-b[i]$ единиц груза; наконец, если $b[i] = 0$, то в i -й пункт не надо доставлять груз, и запас его в данном пункте равен 0 (этот пункт можно использовать как транзитный). Предполагается однородность перевозимого груза.

Для каждого участка сети $j \in 1:n$ задана величина $c[j]$ — затраты на перевозку единицы груза по этому участку (напомним, что перевозки осуществляются только в одном направлении). План перевозок можно задавать вектором $x = (x[1], x[2], \dots, x[n])$, компонентами которого являются объемы перевозок по каждому из участков сети. Обозначим через $S^+(i)$ множество номеров дорог сети, ведущих в i -й пункт, а через $S^-(i)$ — выходящих из него. При сделанных выше предположениях сформулируем следующую задачу:

$$(\mathcal{D}', f') : f'(x) = \sum_{j=1}^n c[j] x[j] \rightarrow \min,$$

$$\mathcal{D}' = \{x \in R^n : \sum_{j \in S^+(i)} x[j] - \sum_{j \in S^-(i)} x[j] = b[i], i \in 1:m; x \geq 0\},$$

которую будем называть *транспортной задачей в сетевой постановке*. Она также является задачей линейного программирования. Задачу $(\mathcal{D}, f)$ можно свести к задаче $(\mathcal{D}', f')$. Подробнее об этом см. в п. 6.3.6.

3.4. Транспортные задачи с ограниченными сверху переменными. Напрашивается естественное усложнение моделей и экстремальных задач пп. 3.2 и 3.3. В модели п. 3.2 предположим, что объем перевозок между i -м пунктом производства и j -м пунктом потребления не превосходит заданной величины $h[i, j]$ — пропускной способности транспорта между этими пунктами. Тогда к системе ограничений, определяющих множество $\mathcal{D}$, следует добавить неравенства $0 \leq x[i, j] \leq h[i, j]$, $i \in 1:m$, $j \in 1:n$. Если для некоторых i и j допускается неограниченная по объему транспортировка груза, то $h[i, j] = +\infty$. Соответствующую экстремальную задачу будем называть *транспортной задачей в матричной постановке с ограниченными сверху переменными*.

Аналогично определяется *транспортная задача в сетевой постановке с ограниченными сверху переменными*: для каждого j -го участка, $j \in 1:n$, определяется верхняя граница $h[j]$ объема перевозок по этому участку и к системе ограничений задачи п. 3.3 добавляются ограничения $0 \leq x[j] \leq h[j]$, $j \in 1:n$.

Заметим, что в практических задачах проблема определения $h[i, j]$ ($h[j]$) оказывается очень сложной.

3.5. Станковая задача, или задача распределения производственной программы [18]. Детали n видов, из которых собирается некоторое изделие, можно обрабатывать на m станках. На

сборку одного изделия расходуется $k[j]$ единиц j -й детали, $k[j] > 0$ для $j \in 1:n$. Известна $A = (a[i, j]) - m \times n$ матрица производительностей станков. Ее элемент $a[i, j]$ — производительность i -го станка по j -й детали, т. е. число деталей j -го вида, которое можно произвести на i -м станке в единицу времени. Если $a[i, j] = 0$, то на i -м станке нельзя производить детали j -го вида.

Сделаем следующие предположения.

1. Плановый период равен единице (году, кварталу, месяцу и т. п.) и совпадает с плановым фондом времени каждого из рассматриваемых станков.

2. Количество деталей каждого вида, производимых на любом станке, линейно зависит от времени, в течение которого на этом станке обрабатываются детали данного вида.

3. Можно не учитывать различные организационные мероприятия, направленные на выполнение производственной программы.

Математическая модель, к построению которой мы приступаем, будет описывать все возможные варианты распределения производственной программы, т. е. распределение станков в течение планового периода для обработки деталей определенного вида с указанием времени их работы. Она позволит для каждого варианта распределения производственной программы определять общее число изделий, выпускаемых за плановый период.

Обозначим через $x[i, j]$ время, в течение которого i -й станок будет занят обработкой деталей j -го вида. Тогда $\sum_{j=1}^n x[i, j]$ — время работы i -го станка в течение планируемого периода и, как в транспортной задаче, вариант распределения станков для обработки деталей на планируемый период задается $m \times n$ матрицей $x = (x[i, j])$ или вектором из R^{mn} , если строки этой матрицы выписать последовательно в одну строку. Множество $\mathcal{D}$ допустимых вариантов $x \in R^{mn}$ закрепления определяется, как нетрудно видеть, следующей системой ограничений:

$$x[i, j] \geq 0, \quad i \in 1:m, \quad j \in 1:n, \quad (5)$$

$$\sum_{j=1}^n x[i, j] \leq 1, \quad i \in 1:m. \quad (6)$$

При работе по варианту $x = (x[i, j])$ за планируемый период будет произведено $\sum_{i=1}^m a[i, j] x[i, j]$ единиц деталей j -го вида, $j \in 1:n$. Из этих деталей можно собрать не больше $\frac{1}{k[j]} \sum_{i=1}^m a[i, j] x[i, j]$ готовых изделий. Следовательно, всего можно изготовить (без учета целочисленности)

$$f(x) = \min_{j \in 1:n} \frac{1}{k[j]} \sum_{i=1}^m a[i, j] x[i, j] \quad (7)$$

изделий.

Условия (5)–(7) задают *станковую (распределительную) модель*. На ее основе обычно рассматривают экстремальную задачу $(\mathcal{D}, f): f(x) \rightarrow \max, x \in \mathcal{D}$, которая называется *станковой, или распределительной*.

Легко проверить, что f — *кусочно-линейная вогнутая функция*, и, следовательно, $(\mathcal{D}, f)$ — *задача выпуклого программирования*. — Упражнение 2.5, а). Однако, как будет показано в п. 4.5, она достаточно просто сводится к задаче линейного программирования.

3.6. Основная задача производственного планирования (задача Л. В. Канторовича [20, с. 281]). Для некоторой производственной системы (участка, цеха, завода, народного хозяйства и т. п.) известны L *технологических способов организации производства*, в которых участвуют N факторов (ингредиентов). Они разбиты на две группы. К одной из них отнесены первые m ингредиентов: это разные виды сырья, трудовые ресурсы и другие факторы. Они могут производиться или потребляться. Условно будем считать, что i -й ингредиент этой группы ($i \in 1:m$) должен «производиться» в количестве, не меньшем, чем $b[i]$ единиц (если $b[i] > 0$, то он действительно производится, если же $b[i] < 0$, то он может потребляться в количестве, не большем $|b[i]|$).

К другой группе отнесены последние n ингредиентов (т. е. $N = m + n$). Они также могут производиться или потребляться в технологических способах, но их итоговый выпуск за планируемый период (без учета собственных потребностей системы) должен удовлетворять отношению $k[1]:k[2]:\dots:k[n]$ (т. е. производиться в заданном ассортименте), где $k[j] > 0$ для $j \in 1:n$.

Технологический способ с номером $l \in 1:L$ задается N -мерным вектором $a^l = (a[1, l], a[2, l], \dots, a[m, l], a[m+1, l], \dots, a[m+n, l])$. Его компонентами являются коэффициенты затрат — выпуска каждого из N ингредиентов при использовании l -го способа в единицу времени. Точнее, если $a[i, l] > 0$, то i -й ингредиент производится в l -м технологическом способе; если же $a[i, l] < 0$, то он расходуется. Возможен случай, когда $a[i, l] = 0$, т. е. i -й ингредиент не производится и не потребляется. Таким образом, каждый технологический способ организации производства рассматривается как поведение «черного ящика». Так принято называть объект, внутренняя структура которого неизвестна, но по одним каналам (входам) на него можно оказывать воздействие, а по другим (выходам) фиксировать его результаты.

Приступим к составлению математической модели. Она будет описывать все допустимые варианты использования технологических способов организации производства (на планируемый период) и позволит вычислять для каждого такого варианта число производимых ассортиментных наборов. При этом будем

считать, что объем производства (потребления) каждого ингредиента линейно зависит от времени использования технологических способов.

Обозначим через $x[l]$ время, в течение которого в планируемый период используется l -й технологический способ. Тогда вариант использования технологических способов можно задавать вектором $x = (x[1], x[2], \dots, x[L])$. При его использовании будет «произведено» за планируемый период $\sum_{l=1}^L a[i, l] x[l]$ единиц i -го ингредиента, $i \in 1:N$. Следовательно, множество $\mathcal{D}$ допустимых вариантов организации производства $x \in R^L$ задается следующим образом:

$$x[l] \geq 0, \quad l \in 1:L, \quad (8)$$

$$\sum_{l=1}^L a[i, l] x[l] \geq b[i], \quad i \in 1:m. \quad (9)$$

Так же, как и в станковой модели п. 3.5,

$$f(x) = \min_{j \in 1:n} \frac{1}{k[j]} \sum_{l=1}^L a[m+j, l] x[l] \quad (10)$$

— число ассортиментных наборов, производимых по варианту x .

Условия (8)–(10) задают модель Канторовича. Она является одной из самых популярных и полезных моделей в отечественной литературе, посвященной применению математических методов в экономических исследованиях. На ее основе обычно рассматривают экстремальную задачу $(\mathcal{D}, f): f(x) \rightarrow \max, x \in \mathcal{D}$, которая называется *основной задачей производственного планирования*, или *задачей Канторовича*. Нетрудно видеть, что она, как и станковая задача, является задачей выпуклого программирования. — Упражнение 2.5, 6). В п. 4.6 будет показано, что задача Канторовича сводится к задаче линейного программирования.

3.7. Задача о «диете». Рассмотрим следующую условную ситуацию. В рацион некоторого вида скота может входить n продуктов. Их стоимости равны $c[j]$, $j \in 1:n$, рублей за единицу, например, массы, а суточная потребность одного животного в продукте j -го наименования не превышает $h[j]$ единиц, $j \in 1:n$. Известно также, что животному необходимо m видов питательных веществ. Возможность усвоения вещества i -го вида составляет не менее $b[i]$ единиц ($i \in 1:m$), а в единице продукта j -го наименования содержится $a[i, j]$ единиц питательного вещества i -го вида ($j \in 1:n$). Наконец, суммарный прием продуктов за сутки не превосходит K единиц.

Обозначим через $x[j]$ количество продукта j -го наименования в суточном рационе. Тогда $\sum_{j=1}^n x[j]$ — суммарное потребление продуктов за сутки, а $\sum_{j=1}^n a[i, j] x[j]$ — общее количество i -го питательного вещества в суточном рационе, который задается вектором $x = (x[1], x[2], \dots, x[n])$. Очевидно, что множество $\mathcal{D}$

«допустимых» рационов $x \in R^n$ определяется следующей системой ограничений:

$$0 \leq x[j] \leq h[j], \quad j \in 1:n, \quad (11)$$

$$\sum_{j=1}^n x[j] \leq K, \quad (12)$$

$$\sum_{j=1}^n a[i, j] x[j] \geq b[i], \quad i \in 1:m. \quad (13)$$

Стоимость рациона $x = (x[1], x[2], \dots, x[n])$ составляет

$$f(x) = \sum_{j=1}^n c[j] x[j] \quad (14)$$

рублей. На основе модели (11)–(14) естественно возникает экстремальная задача $(\mathcal{D}, f): f(x) \rightarrow \min, x \in \mathcal{D}$, которая и называется *задачей о диете*. Она является задачей линейного программирования с ограниченными сверху переменными (см. (11)).

3.8. Задача о назначениях (задача выбора). Ситуация, на основе которой обычно ставится задача о назначениях, выглядит достаточно искусственной. Поэтому целесообразность рассмотрения этой задачи объясняется в первую очередь обилием приложений, т. е. можно привести множество примеров, где возникают задачи, которые можно интерпретировать как задачи о назначениях.

Пусть n видов работ могут выполнять n кандидатов, при этом каждого кандидата можно назначать только на одну работу и каждая работа должна выполняться только одним кандидатом. Известна $n \times n$ матрица $c = (c[i, j])$, где $c[i, j]$ — «эффективность» (в некоторых выбранных единицах) i -го претендента на j -й работе. Каждый вариант распределения кандидатов по работам задается $n \times n$ матрицей $x = (x[i, j])$ (или вектором из R^{n^2}). Множество $\mathcal{D}$ допустимых вариантов распределения $x \in R^{n^2}$, как нетрудно видеть, определяется следующей системой ограничений:

$$x[i, j] = 1 \quad \text{или} \quad x[i, j] = 0, \quad i \in 1:n, \quad j \in 1:n, \quad (15)$$

$$\sum_{j=1}^n x[i, j] = 1, \quad i \in 1:n, \quad (16)$$

$$\sum_{i=1}^n x[i, j] = 1, \quad j \in 1:n, \quad (17)$$

где $x[i, j] = 1$, если i -й кандидат назначается на j -ю работу, и $x[i, j] = 0$ в противном случае. Общая «эффективность» распределения кандидатов по варианту $x = (x[i, j])$ равна

$$f(x) = \sum_{i=1}^n \sum_{j=1}^n c[i, j] x[i, j]. \quad (18)$$

На основе модели (15)–(18) возникает экстремальная задача $(\mathcal{D}, f): f(x) \rightarrow \max, x \in \mathcal{D}$, которая называется *задачей о назначениях* (или *задачей выбора*). Нетрудно видеть, что $(\mathcal{D}, f)$ является задачей дискретного программирования. Действительно, $|\mathcal{D}| = n!$, и, вообще говоря, можно привести правило, позволяю-

щее при любом n перечислять все элементы множества $\mathcal{D}$. Понятно, что подобную «деятельность» могут лимитировать только сроки.

3.9. Задача коммивояжера. Эту экстремальную задачу принято связывать со следующей достаточно искусственной проблемой. Бродячему торговцу (коммивояжеру) необходимо, начиная с города, в котором он находится, обойти, не побывав нигде дважды, n городов и вернуться в исходный. Известна $n \times n$ матрица $c = (c[i, j])$, где $c[i, j]$ — «стоимость» путешествия из i -го города в j -й. Поскольку коммивояжера интересуют замкнутые маршруты (циклы), то, как легко видеть, безразлично, какой из городов считать начальным. Поэтому начальным будем считать город с индексом 1.

Пусть $x[i, j] = 1$, если коммивояжер переезжает из города i в город j , и $x[i, j] = 0$ в противном случае. Тогда очевидно любой маршрут путешествия коммивояжера можно задать $n \times n$ матрицей $x = (x[i, j])$ или множеством

$$N(x) = \{(i, j) : x[i, j] = 1\} = \{(1, i_1), (i_1, i_2), \dots, (i_{n-1}, 1)\}$$

пар смежных городов маршрута. Так, например, при $n = 3$ матрица

$$x^0 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \text{ или множество } N(x^0) = \{(1, 2), (2, 3), (3, 1)\}$$

задают маршрут: из первого города коммивояжер «идет» во второй, из второго — в третий, из третьего возвращается в первый, что также можно изобразить диаграммой $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$.

Множество $\mathcal{D} \subset R^n$ допустимых маршрутов определяется системой условий

$$x[i, j] = 1 \text{ или } x[i, j] = 0, \quad i \in 1:n, \quad j \in 1:n, \quad (19)$$

$$\sum_{j=1}^n x[i, j] = 1, \quad i \in 1:n, \quad (20)$$

$$\sum_{i=1}^n x[i, j] = 1, \quad j \in 1:n, \quad (21)$$

$$i_1, i_2, \dots, i_{n-1} \text{ (в } N(x)) \text{ — перестановка чисел } 2, 3, \dots, n. \quad (22)$$

Стоимость путешествия по маршруту x равна

$$f(x) = \sum_{i=1}^n \sum_{j=1}^n c[i, j] x[i, j] = \sum_{(i, j) \in N(x)} c[i, j]. \quad (23)$$

На основе модели (19) — (23) естественно возникает экстремальная задача $(\mathcal{D}, f) : f(x) \rightarrow \min, x \in \mathcal{D}$, которая и называется *задачей коммивояжера*. Форма описания множества $\mathcal{D}$ (далеко не самая рациональная) выбрана так, чтобы подчеркнуть его связь с множеством допустимых планов задачи о назначениях (см. п. 3.8): $\mathcal{D}$ является подмножеством допустимых планов этой задачи, не распадающихся в произведение двух и более

циклов. Так, например, при $n = 3$ план $x' = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ является

допустимым планом соответствующей задачи о назначениях, но $x' \notin \mathcal{D}$, поскольку распадается на два цикла: $1 \rightarrow 2 \rightarrow 1$ и $3 \rightarrow 3$.

Ясно, что $(\mathcal{D}, f)$ является задачей дискретного программирования и $|\mathcal{D}| = (n-1)!$

Чтобы подчеркнуть комбинаторный характер задачи коммивояжера, приведем еще одну формулировку:

$$(\mathcal{D}, f): f(\pi) = f(i_1, i_2, \dots, i_{n-1}) = \\ = c[1, i_1] + \sum_{k=1}^{n-2} c[i_k, i_{k+1}] + c[i_{n-1}, 1] \rightarrow \min,$$

$\mathcal{D}$ — множество перестановок $\pi = (i_1, i_2, \dots, i_{n-1})$ чисел $2, 3, \dots, n$.

Задача коммивояжера, пожалуй, одна из самых популярных задач дискретного программирования, что обусловлено простотой и изящностью постановки, с одной стороны, и значительными математическими трудностями ее анализа — с другой. Первый достаточно удовлетворительный точный метод ее решения был предложен только в 1963 г. [52].

3.10. Задача о переналадке оборудования. Опишем следующую производственно-экономическую проблему, которая сводится к решению задачи коммивояжера. Пусть на поточной линии предусмотрена обработка n изделий. При обработке каждого вида изделия на поточной линии используются разные приспособления и инструмент. Поэтому при переходе от обработки изделия одного вида к обработке изделия другого вида необходима переналадка оборудования: снятие и установка приспособлений и инструмента, что естественно требует затрат времени. Известна $n \times n$ матрица $c = (c[i, j])$, где $c[i, j]$ — время переналадки оборудования при переходе от обработки i -го изделия к обработке j -го. Если обработка всех n изделий на поточной линии осуществляется последовательно, партиями, то возникает задача: в какой последовательности запускать партии изделий на линию, чтобы общее время на переналадку оборудования было минимальным? Решение этой задачи очевидно сводится к решению задачи коммивояжера с матрицей c .

3.11. Задача о раскрое. Одна из первых производственно-экономических проблем, которая была успешно решена Л. В. Канторовичем (1939 г.) и практически послужила толчком к развитию нового направления прикладной математики — линейного программирования, сводилась к следующему.

В заготовительный цех предприятия поступает материал в виде листов проката, фанеры и т. п. определенного размера. Из него вырезаются заготовки m видов. Размеры заготовок известны.

Например, из листов проката 3×7 нужно вырезать заготовки размерами 2×3 и 2×2 (в метрах). На рис. 13 схематически изображены исходный лист проката и вырезаемые из него заготовки.

Предполагается, что заданы n вариантов раскроя и с каждым вариантом связан $(m+1)$ -мерный вектор $\bar{a}^i = (c[j], a[1, j], \dots, a[m, j])$, $j \in 1:n$, где $a[i, j]$ — число заготовок вида i

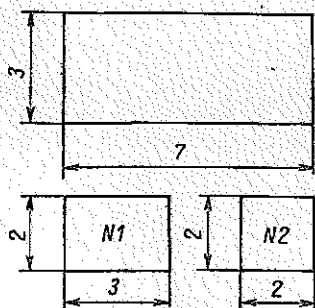


Рис. 13.

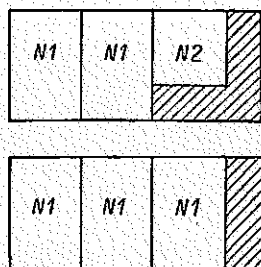


Рис. 14.

и $c[j]$ — величина отходов, которые получаются при раскрое по варианту с номером j .

Если в описанном выше случае заданы варианты раскроя, приведенные на рис. 14 (неиспользуемые части листа заштрихованы), то первому отвечает вектор $\bar{a}^1 = (5, 2, 1)$, а второму — вектор $\bar{a}^2 = (3, 3, 0)$ (по первому варианту раскроя вырезаются две заготовки № 1, одна заготовка № 2, и отходы равны 5 м^2 , а по второму — три заготовки № 1, и отходы равны 3 м^2).

Известно также, что необходимо получить не меньше $b[i]$ заготовок i -го вида для каждого $i \in 1:m$. Возникает проблема рационального раскроя поступившего материала.

Данную ситуацию нетрудно формализовать. Пусть $x[j]$ — число листов, раскраиваемых по варианту j . Тогда $\sum_{j=1}^n a[i, j] x[j]$ — число заготовок i -го вида, соответствующее плану $x = (x[1], x[2], \dots, x[n])$ использования вариантов раскроя. Допустимые планы использования вариантов раскроя описываются системой ограничений

$$\sum_{j=1}^n a[i, j] x[j] \geq b[i], \quad i \in 1:m, \quad (24)$$

$$x[j] \geq 0, \quad x[j] - \text{ц.ч.}, \quad j \in 1:n. \quad (25)$$

Число разрезанных листов равно

$$f_1(x) = \sum_{j=1}^n x[j], \quad (26)$$

а общая площадь отходов

$$\bar{f}_2(x) = \sum_{j=1}^n c[j] x[j]. \quad (27)$$

На основе модели (24) — (27) можно сформулировать две экстремальные задачи минимизации:

$$(\mathcal{D}, f_1) : f_1(x) = \sum_{j=1}^n x[j] \rightarrow \min, \mathcal{D} = \{x \in R^n :$$

$\sum_{j=1}^n a[i, j] x[j] \geq b[i], i \in 1 : m; x \geq 0, x[j] - \text{ц.ч.}, j \in 1 : n\}$
и $(\mathcal{D}, f_2)$, которые и назовем *задачами о раскрое*. Каждая из них является целочисленной задачей линейного программирования.

Существуют различные варианты и обобщения задачи о раскрое [23]. Но мы обращаем внимание лишь на один момент. Для того чтобы сформулировать в практической ситуации любую задачу о раскрое, необходимо предварительно перечислить все возможные технологически допустимые варианты раскроя.

На рис. 14 приведены лишь два варианта раскроя листа 3×7 на заготовки двух видов. Читателю предлагается перечислить все остальные возможные варианты раскроя. —

Упражнение 2.6.

3.12. Задача сглаживания. Цех может производить r видов продуктов. Для их изготовления используется q видов сырья. Работу цеха рассматриваем в p периодах. Известны следующие величины: $a[i, j]$ — потребность в продукте вида i в j -м периоде, $i \in 1 : r, j \in 1 : p$; $t[i]$ — трудоемкость изготовления (единицы) i -го продукта; $c[i, k]$ — расход сырья вида k на производство единицы i -го продукта, $k \in 1 : q$; $C[k, j]$ — планируемая поставка k -го вида сырья в j -м периоде; $z_0[i]$ — начальный запас i -го продукта; $T[j]$ — общий фонд рабочего времени в j -м периоде.

В указанных условиях предлагается определить план работ цеха, обеспечивающий равномерную загрузку по периодам.

Обозначим через $x[i, j]$ количество продукта i , изготавливаемое в периоде j , через Z — «среднюю» загрузку цеха за период. Тогда, как легко видеть, множество

$$\mathcal{D} = \{(x, Z) \in R^{rp+1} : \sum_{i=1}^r c[i, k] x[i, j] \leq C[k, j], k \in 1 : q, j \in 1 : p;$$

$$\sum_{i=1}^r t[i] x[i, j] \leq T[j], j \in 1 : p;$$

$$z_0[i] + \sum_{j=1}^m x[i, j] \geq \sum_{j=1}^m a[i, j], m \in 1 : p-1, i \in 1 : r; \quad (28)$$

$$\sum_{j=1}^p x[i, j] \geq \sum_{j=1}^p a[i, j]; x \geq 0, Z \geq 0\}$$

содержит все допустимые варианты организации производства. Предполагается, что характеристикой равномерности загрузки по периодам является величина

$$f(x, Z) = \sum_{j=1}^p (\sum_{i=1}^r t[i] x[i, j] - Z)^2. \quad (29)$$

На основе модели (28), (29) сформулируем экстремальную задачу $(\mathcal{D}, f): f(x, Z) \rightarrow \min, (x, Z) \in \mathcal{D}$, которую назовем *задачей сглаживания*. Она является задачей квадратичного программирования. — Упражнение 2.7.

§ 4. Эквивалентные экстремальные задачи

4.1. Обсуждение проблемы. Пусть $\mathcal{E}$ — множество экстремальных задач и $(\mathcal{D}, f)$ — некоторая экстремальная задача из $\mathcal{E}$. Мы уже условились, что $\mathcal{D}$ — подмножество пространства R^n при некотором n . Описание множества $\mathcal{D}$ может производиться различными способами, но вряд ли в множестве экстремальных задач следует считать разными задачи максимизации (минимизации) $(\mathcal{D}, f)$ и $(\mathcal{D}', f')$, если $\mathcal{D} = \mathcal{D}'$ и $f(x) = f'(x)$ для любого $x \in \mathcal{D}$. Не будем считать разными задачу максимизации с областью допустимых планов $\mathcal{D}$ и функцией f и задачу минимизации с областью допустимых планов $\mathcal{D}'$ и функцией f' , если $\mathcal{D} = \mathcal{D}'$ и $f(x) = -f'(x)$ для всякого $x \in \mathcal{D}$. Поэтому можно ограничиться рассмотрением множества задач максимизации (или минимизации).

Пусть $\mathcal{E}$ — множество задач на тах. Для наших целей будет полезным отождествление в $\mathcal{E}$ экстремальных задач $(\mathcal{D}, f)$ и $(\mathcal{D}', f')$ с различными множествами $\mathcal{D}$ и $\mathcal{D}'$, быть может, лежащими даже в различных пространствах R^n и $R^{n'}$ ($n \neq n'$) и различными функциями f и f' (когда $\mathcal{D} = \mathcal{D}'$), если эти задачи обладают рядом общих свойств.

Естественно, что эквивалентные задачи $(\mathcal{D}, f)$ и $(\mathcal{D}', f')$ должны обладать следующим свойством: информации о решении любой из них должно быть достаточно для получения всей необходимой информации о решении другой. Так, например, если нас интересуют в первую очередь вычислительные аспекты, то мы должны уметь по оптимальному плану задачи $(\mathcal{D}, f)$ однозначно (и за разумное время) вычислять оптимальный план эквивалентной ей задачи $(\mathcal{D}', f')$, т. е. необходим алгоритм, позволяющий за конечное число шагов осуществлять такие вычисления. Если же задача $(\mathcal{D}, f)$ ($(\mathcal{D}', f')$) не имеет решения, то, вероятно, не должна иметь решения и эквивалентная ей задача $(\mathcal{D}', f')$ ($(\mathcal{D}, f)$). Только в этом случае можно отождествлять задачи $(\mathcal{D}, f)$ и $(\mathcal{D}', f')$.

В следующем пункте будет определено отношение эквивалентности на $\mathcal{E}$, которое приведет к разбиению множества $\mathcal{E}$ на классы, отвечающему поставленным целям.

4.2. Определение эквивалентных экстремальных задач. Отображение $\mu: A \rightarrow B$, $A \subset R^m$, $B \subset R^n$, назовем *алгоритмическим*, если для любого $x \in A$ каждая компонента вектора $\mu(x) \in B$ может быть вычислена с любой наперед заданной точностью

по компонентам вектора x в результате применения к ним конечного числа *арифметических операций*: сложения, вычитания, умножения, деления, сравнения.

Экстремальные задачи $(\mathcal{D}, f)$ и $(\mathcal{D}', f')$ из $\mathfrak{E}$ будем называть *эквивалентными*, если существуют взаимно однозначное монотонно возрастающее алгоритмичное отображение $g: R \rightarrow R$ (так, что g^{-1} также алгоритмичное монотонно возрастающее отображение) и такие алгоритмичные отображения $\varphi: \mathcal{D} \rightarrow \mathcal{D}'$ и $\psi: \mathcal{D}' \rightarrow \mathcal{D}$, что $f(x) \leq g(f'(\varphi(x)))$ для любого $x \in \mathcal{D}$, $f'(x') \leq g^{-1}(f(\psi(x')))$ для любого $x' \in \mathcal{D}'$.

Нетрудно проверить, что отношение, определенное нами на $\mathfrak{E}$, рефлексивно, симметрично и транзитивно, т. е. является отношением эквивалентности. — Упражнение 2.8.

Отображение φ будем называть *эквивалентным преобразованием задачи* $(\mathcal{D}, f)$ в задачу $(\mathcal{D}', f')$ (ψ — эквивалентное преобразование задачи $(\mathcal{D}', f')$ в задачу $(\mathcal{D}, f)$), отображение g — *корректирующим преобразованием задач* $(\mathcal{D}, f)$ и $(\mathcal{D}', f')$. В дальнейшем, когда это не вызовет недоразумений, φ и ψ будем называть просто *эквивалентными преобразованиями*, а g — *корректирующим преобразованием*.

4.3. Некоторые свойства эквивалентных экстремальных задач. Если экстремальные задачи $(\mathcal{D}, f)$ и $(\mathcal{D}', f')$ эквивалентны, то:

а) если множество $\mathcal{D}(\mathcal{D}')$ непусто, то и $\mathcal{D}'(\mathcal{D})$ непусто;
 б) если φ — эквивалентное преобразование задачи $(\mathcal{D}, f)$ в $(\mathcal{D}', f')$ и $\bar{x} \in \mathcal{D}$ — оптимальный план задачи $(\mathcal{D}, f)$, то $\varphi(\bar{x})$ — оптимальный план задачи $(\mathcal{D}', f')$ и $g^{-1}(f(\bar{x})) = f'(\varphi(\bar{x}))$ ($f(\bar{x}) = g(f'(\varphi(\bar{x})))$);

в) если существует оптимальный план в экстремальной задаче, то существует оптимальный план и в эквивалентной ей экстремальной задаче;

г) если $\sup_{x \in \mathcal{D}} f(x) = \alpha < +\infty$, то $\sup_{x' \in \mathcal{D}'} f'(x') = \alpha' < +\infty$ и $g^{-1}(\alpha) = \alpha'$, $\alpha = g(\alpha')$;

д) если α (и α') конечно и $x_1, x_2, \dots$ — такая последовательность планов из $\mathcal{D}$, что $\lim_{k \rightarrow \infty} f(x_k) = \alpha$, то $\lim_{k \rightarrow \infty} f'(\varphi(x_k)) = \alpha'$; в частности, если $\bar{x}$ — оптимальный план задачи $(\mathcal{D}, f)$ и $\lim_{k \rightarrow \infty} f(x_k) = f(\bar{x})$, то $\lim_{k \rightarrow \infty} f'(\varphi(x_k)) = f'(\varphi(\bar{x}))$.

Доказательство. Свойство а) следует из существования эквивалентных преобразований φ и ψ .

б) Пусть $\bar{x}$ — оптимальный план задачи $(\mathcal{D}, f)$, φ и ψ — эквивалентные преобразования задачи $(\mathcal{D}, f)$ в $(\mathcal{D}', f')$ и задачи $(\mathcal{D}', f')$ в $(\mathcal{D}, f)$ соответственно, а g — корректирующее преобразование. Тогда для любого $x' \in \mathcal{D}'$

$$f'(x') \leq g^{-1}(f(\psi(x'))) \leq g^{-1}(f(\bar{x})) \leq g^{-1}(g(f'(\varphi(\bar{x})))) = f'(\varphi(\bar{x})),$$

так что $\varphi(\bar{x})$ — решение задачи $(\mathcal{D}', f')$ и $g^{-1}(f(\bar{x})) \leq f'(\varphi(\bar{x}))$. Но $f'(\varphi(\bar{x})) \leq g^{-1}(f(\psi(\varphi(\bar{x})))) \leq g^{-1}(f(\bar{x}))$. Следовательно, $g^{-1}(f(\bar{x})) = f'(\varphi(\bar{x}))$ или $f(\bar{x}) = g(f'(\varphi(\bar{x})))$.

Свойство в) следует из свойства б).

г) Для любого $x' \in \mathcal{D}$

$$f'(x') \leq g^{-1}(f(\psi(x'))) \leq g^{-1}(\alpha),$$

и, следовательно,

$$\sup_{x' \in \mathcal{D}} f'(x') = \alpha' \leq g^{-1}(\alpha) < +\infty.$$

Аналогично получаем $\alpha \leq g(\alpha')$, $g^{-1}(\alpha) \leq \alpha'$, так что $g^{-1}(\alpha) = \alpha'$, $\alpha = g(\alpha')$.

д) Пусть $x_1, x_2, \dots$ — последовательность планов из $\mathcal{D}$ и $\lim_{k \rightarrow \infty} f(x_k) = \alpha$. Поскольку в силу свойства г) $g^{-1}(\alpha) = \alpha'$ и $f(x_k) \leq g(f'(\varphi(x_k)))$, т. е. $g^{-1}(f(x_k)) \leq f'(\varphi(x_k))$, то $0 \leq \alpha' - f'(\varphi(x_k)) \leq g^{-1}(\alpha) - g^{-1}(f(x_k))$, откуда, переходя к пределу при $k \rightarrow \infty$, и получаем свойство д), так как из определения g и g^{-1} следует их непрерывность. ■

Таким образом, если задачи $(\mathcal{D}, f)$ и $(\mathcal{D}', f')$ эквивалентны, т. е. известны соответствующие эквивалентные и корректирующее преобразования φ, ψ и g , то, как следует из свойств а)–д), информация о некоторых свойствах одной из этих задач может быть «преобразована» в информацию о соответствующих свойствах другой задачи, что вполне отвечает интуитивному представлению об эквивалентных задачах.

Приведем несколько примеров эквивалентных экстремальных задач.

4.4. Эквивалентные задачи с нетривиальным корректирующим преобразованием. Рассмотрим произвольную задачу максимизации $(\mathcal{D}, f)$ и три задачи на максимум: $(\mathcal{D}, f_1)$, $(\mathcal{D}, f_2)$ и $(\mathcal{D}, f_3)$, где $f_1(x) = f(x) + c$, $c = \text{const}$; $f_2(x) = cf(x)$, $c > 0$; $f_3(x) = (f(x))^3$. Покажем, что каждая из этих трех задач эквивалентна задаче $(\mathcal{D}, f)$ и тем самым в силу транзитивности отношения эквивалентности (см. упражнение 2.8) все перечисленные задачи эквивалентны.

Действительно, непосредственно видно, что функции $\varphi(x) = x$, $\psi(x) = x$, $g(z) = z - c$, $g^{-1}(z) = z + c$ алгоритмичны, g, g^{-1} — монотонно возрастающие и

$$g(f_1(\varphi(x))) = f_1(\varphi(x)) - c = f(x) + c - c = f(x), \quad g^{-1}(f(\psi(x))) = f(x) + c = f_1(x),$$

т. е. φ, ψ, g являются соответственно эквивалентными и корректирующим преобразованиями задач $(\mathcal{D}, f)$ и $(\mathcal{D}, f_1)$; аналогично эквивалентными и корректирующим преобразованиями для задач $(\mathcal{D}, f)$ и $(\mathcal{D}, f_2)$ являются функции $\varphi(x) = x$, $\psi(x) = x$,

$g(z) = (1/c)z$ и $\varphi(x) = x$, $\psi(x) = x$, $g(z) = \sqrt[n]{z}$ для задач $(\mathcal{D}, f)$ и $(\mathcal{D}, f_3)$.

4.5. Станковая задача и ее линейная форма. В п. 3.5 станковой была названа экстремальная задача

$$(\mathcal{D}, f): f(x) = \min_{j \in 1:n} \frac{1}{k[j]} \sum_{i=1}^m a[i, j] x[i, j] \rightarrow \max,$$

$$\mathcal{D} = \{x \in R^{mn}: \sum_{j=1}^n x[i, j] \leq 1, i \in 1:m; x \geq 0\},$$

которая, как было установлено, является задачей выпуклого программирования. Рассмотрим следующую задачу линейного программирования:

$$(\mathcal{D}', f'): f'(x', z') = z' \rightarrow \max,$$

$$\mathcal{D}' = \{(x', z') \in R^{mn+1}: x' \in R^{mn}, z' \in R;$$

$$\sum_{j=1}^n x'[i, j] \leq 1, i \in 1:m; x' \geq 0;$$

$$\frac{1}{k[j]} \sum_{i=1}^m a[i, j] x'[i, j] \geq z', j \in 1:n\}.$$

Проверим эквивалентность задач $(\mathcal{D}, f)$ и $(\mathcal{D}', f')$: Это сразу следует из того, что функции

$$\varphi: \mathcal{D} \rightarrow \mathcal{D}', \varphi(x) = (x, f(x)), \psi: \mathcal{D}' \rightarrow \mathcal{D},$$

$$\psi(x', z') = x', g(z) = g^{-1}(z) = z$$

алгоритмичны и

$$g(f'(\varphi(x))) = f'(\varphi(x)) = f(x), g^{-1}(f(\psi(x', z'))) = f(\psi(x', z')) \geq f'(x', z').$$

Задача линейного программирования $(\mathcal{D}', f')$ (ее естественно назвать *линейной формой станковой задачи*) формально более простая, чем задача $(\mathcal{D}, f)$.

4.6. Основная задача производственного планирования (задача Канторовича) и ее линейная форма. В п. 3.6 основной задачей производственного планирования была названа задача

$$(\mathcal{D}, f): f(x) = \min_{j \in 1:n} \frac{1}{k[j]} \sum_{l=1}^L a[m+j, l] x[l] \rightarrow \max,$$

$$\mathcal{D} = \{x \in R^L: \sum_{l=1}^L a[i, l] x[l] \geq b[i], i \in 1:m; x \geq 0\}.$$

Она является задачей выпуклого программирования. Рассмотрим следующую задачу линейного программирования:

$$(\mathcal{D}', f'): f'(x', z') = z' \rightarrow \max, \mathcal{D}' = \{(x', z') \in R^{L+1}:$$

$$x' \in R^L, z' \in R; \sum_{l=1}^L a[i, l] x'[l] \geq b[i], i = 1:m;$$

$$x' \geq 0; \frac{1}{k[j]} \sum_{l=1}^L a[m+j, l] x'[l] \geq z', j \in 1:n\}.$$

Задача $(\mathcal{D}', f')$ эквивалентна задаче $(\mathcal{D}, f)$ и называется ее *линейной формой*. Эквивалентными преобразованиями могут служить функции

$$\varphi: \mathcal{D} \rightarrow \mathcal{D}', \quad \varphi(x) = (x, f(x)) \quad \text{и} \quad \psi: \mathcal{D}' \rightarrow \mathcal{D}, \quad \psi(x', z') = x'.$$

Корректирующим преобразованием является $g(z) = z$.

4.7. Общая задача линейного программирования и ее каноническая форма. В п. 2.2 общей задачей линейного программирования была названа задача

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \quad \mathcal{D} = \{x \in R^n: a_i x \leq b[i],$$

$$i \in I_n; a_i x = b[i], i \in I_y; x[j] \geq 0, j \in J\}.$$

Если $I_y = 1:m$ ($I_n = \emptyset$), $J = 1:n$, то $(\mathcal{D}, f)$ — каноническая задача линейного программирования.

Каноническая задача, эквивалентная данной задаче линейного программирования, называется ее *канонической формой*. Справедливо следующее утверждение: для любой задачи линейного программирования существует каноническая форма.

Доказательство. Пусть $(\mathcal{D}, f)$ — задача линейного программирования, где $\mathcal{D}$ — множество таких $x \in R^n$, что

$$\sum_{j=1}^n a[i, j] x[j] \leq b[i], \quad i \in I_n,$$

$$\sum_{j=1}^n a[i, j] x[j] = b[i], \quad i \in I_y,$$

$$x[j] \geq 0, \quad j \in J,$$

$$\text{а } f(x) = \sum_{j=1}^n c[j] x[j] \quad (I_n \cap I_y = \emptyset, I_n \cup I_y = 1:m, J \subset 1:n).$$

Обозначим через n' число элементов множества $1:n \setminus J$, через m' — число элементов множества I_n . В пространстве $R^{n+m'+n'}$ рассмотрим подмножество $\mathcal{D}'$ векторов

$x' = \{x'[j], j \in J; \bar{x}'[j], j \in 1:n \setminus J; x'[n+i], i \in I_n\}$, удовлетворяющих условиям

$$x' \geq 0,$$

$$\sum_{j \in J} a[i, j] x'[j] + \sum_{j \in 1:n \setminus J} a[i, j] \bar{x}'[j] -$$

$$- \sum_{j \in 1:n \setminus J} a[i, j] \bar{x}'[j] + x'[n+i] = b[i], \quad i \in I_n,$$

$$\sum_{j \in J} a[i, j] x'[j] + \sum_{j \in 1:n \setminus J} a[i, j] \bar{x}'[j] -$$

$$- \sum_{j \in 1:n \setminus J} a[i, j] \bar{x}'[j] = b[i], \quad i \in I_y,$$

и линейную функцию f' на $\mathcal{D}'$:

$$f'(x') = \sum_{j \in J} c[j] x'[j] + \sum_{j \in 1:n \setminus J} c[j] \bar{x}'[j] - \sum_{j \in 1:n \setminus J} c[j] \bar{x}'[j].$$

Задача $(\mathcal{D}', f')$ есть каноническая задача линейного программирования. Покажем, что она является канонической формой задачи $(\mathcal{D}, f)$. Для этого определим эквивалентные преоб-

разования $\varphi: \mathcal{D} \rightarrow \mathcal{D}'$ и $\psi: \mathcal{D}' \rightarrow \mathcal{D}$. Для $x \in \mathcal{D}$ положим $\varphi(x) = x'$, где $x'[j] = x[j]$, $j \in J$,

$$\bar{x}'[j] = \max\{0, x[j]\}, \quad \bar{x}'[j] = \max\{0, -x[j]\}, \quad j \in 1:n \setminus J;$$

$$x'[n+i] = b[i] - \sum_{j=1}^n a[i, j] x[j], \quad i \in I_n.$$

Для $x' \in \mathcal{D}'$ положим $\psi(x') = x$, где $x[j] = x'[j]$, $j \in J$; $x[j] = \bar{x}'[j] - \bar{x}'[j]$, $j \in 1:n \setminus J$. Корректирующее преобразование определим как тождественное: $g(z) = z$. Тогда очевидно, что $f(x) = f'(\varphi(x))$ и $f'(x') = f(\psi(x'))$. ■

Переход от задачи линейного программирования к ее канонической форме назовем *приведением к каноническому виду*.

Мы доказали, что *любую задачу линейного программирования можно привести к каноническому виду*.

4.8. Еще один способ построения канонической формы задачи линейного программирования. В п. 4.7 было доказано, что любую задачу линейного программирования можно привести к каноническому виду, и приведен способ построения канонической формы задачи линейного программирования. В настоящем пункте будет указан другой способ построения канонической формы.

Пусть $(\mathcal{D}, f): f(x) = \sum_{j=1}^n c[j] x[j] \rightarrow \max$, $\mathcal{D} = \{x \in R^n: \sum_{j=1}^n a[i, j] x[j] = b[i], i \in I_y; \sum_{j=1}^n a[i, j] x[j] \leq b[i], i \in I_n; x[j] \geq 0, j \in J\}$ — общая задача линейного программирования. Обозначим через m' число элементов множества I_n . В пространстве $R^{n+m'+1}$ рассмотрим подмножество $\mathcal{D}''$ векторов

$$x'' = \{x''[j], j \in 0:n, x''[n+i], i \in I_n\},$$

удовлетворяющих условиям

$$x'' \geq 0,$$

$$\sum_{j=1}^n a[i, j] x''[j] - x''[0] \sum_{j \in J} a[i, j] = b[i], \quad i \in I_y, \quad (1)$$

$$\sum_{j=1}^n a[i, j] x''[j] - x''[0] \sum_{j \in J} a[i, j] + x''[n+i] = b[i], \quad i \in I_n, \quad (2)$$

и линейную функцию $f''(x'') = \sum_{j=1}^n c[j] x''[j] - x''[0] \sum_{j \in J} c[j]$.

Покажем, что задача $(\mathcal{D}'', f'')$ является канонической формой задачи $(\mathcal{D}, f)$. Отметим при этом, что размерность $n + m' + 1$ пространства, содержащего $\mathcal{D}''$, вообще говоря, меньше размерности $n + m' + n'$ пространства $R^{n+m'+n'}$, в котором содержится множество $\mathcal{D}'$ допустимых планов канонической формы $(\mathcal{D}', f')$, построенной в п. 4.7.

Построим отображение $\varphi: \mathcal{D} \rightarrow R^{n+m'+1}$. Пусть $x \in \mathcal{D}$. Определим $x'' = \varphi(x)$ следующим образом. Если $\min_{j \in J} x[j] < 0$, то положим $x''[0] = |\min_{j \in J} x[j]|$, в противном случае $x''[0] = 0$. Далее положим $x''[j] = x[j]$ для $j \in J$, $x''[j] = x[j] + x''[0]$ для $j \in J$, $x''[n+i] = b[i] - \sum_{j=1}^n a[i, j] x[j]$, $i \in I_n$.

Покажем, что $\varphi: \mathcal{D} \rightarrow \mathcal{D}''$. Очевидно, что $x'' \geq 0$. Поскольку $\sum_{i=1}^n a[i, j] x[j] = b[i]$, $i \in I_y$, то

$$\begin{aligned} \sum_{i=1}^n a[i, j] x''[j] - x''[0] \sum_{i \in J} a[i, j] &= \sum_{i \in J} a[i, j] x''[j] + \\ &+ \sum_{i \in J} a[i, j] (x''[j] - x''[0]) = \sum_{i \in J} a[i, j] x[j] + \\ &+ \sum_{i \in J} a[i, j] x[j] = b[i], \end{aligned}$$

так что равенства (1) выполняются.

Аналогично для $i \in I_n$

$$\begin{aligned} \sum_{i=1}^n a[i, j] x''[j] - x''[0] \sum_{i \in J} a[i, j] + x''[n+i] &= \\ = \sum_{i=1}^n a[i, j] x[j] + b[i] - \sum_{i=1}^n a[i, j] x[j] &= b[i], \end{aligned}$$

т. е. выполняются равенства (2).

Построим отображение $\psi: \mathcal{D}'' \rightarrow R^n$. Для $x'' \in \mathcal{D}''$ определим $x = \psi(x'')$ следующим образом: $x[j] = x''[j] - x''[0]$, $j \in J$; $x[j] = x''[j]$, $j \in J$. Тогда $\psi: \mathcal{D}'' \rightarrow \mathcal{D}$. Действительно, $x[j] \geq 0$, $j \in J$,

$$\begin{aligned} \sum_{i=1}^n a[i, j] x[j] &= \sum_{i \in J} a[i, j] x''[j] + \sum_{i \in J} a[i, j] (x''[j] - x''[0]) = \\ &= \sum_{i=1}^n a[i, j] x''[j] - x''[0] \sum_{i \in J} a[i, j]. \end{aligned}$$

Последнее выражение равно $b[i]$ для $i \in I_y$ и не превосходит $b[i]$ для $i \in I_n$, поскольку $x''[n+i] \geq 0$.

Определим корректирующее преобразование как тождественное. Тогда очевидно

$$\begin{aligned} f(x) &= f''(\varphi(x)) \quad \text{для любого } x \in \mathcal{D}, \\ f''(x'') &= f(\psi(x'')) \quad \text{для любого } x'' \in \mathcal{D}'', \end{aligned}$$

так что задачи $(\mathcal{D}, f)$ и $(\mathcal{D}'', f'')$ эквивалентны.

4.9. Примеры построения канонической формы. Приведем несколько примеров приведения задач линейного программирования к каноническому виду.

1. Для задачи максимизации $(\mathcal{D}, f)$, где $f(x) = x_1 - 2x_2 + x_3 + 4x_5$, а множество $\mathcal{D}$ определено следующей системой уравнений и неравенств

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 7, \\ -2x_2 + 3x_3 - x_4 - 5x_5 &\leq 2, \\ 2x_1 - 2x_3 + 2x_4 - 3x_5 &= 5, \\ x_1 \geq 0, \quad x_3 \geq 0, \end{aligned}$$

канонической формой, построение которой осуществлено по правилам, указанным в п. 4.7, является задача максимизации $(\mathcal{D}', f')$, где $f'(x') = x'_1 - 2x'_2 + 2x'_3 + x'_4 + 4x'_5 - 4x'_5$, а множе-

ство $\mathcal{D}'$ определено так:

$$\begin{aligned} x'_1 + 3\bar{x}'_2 - 3\bar{x}'_2 + x'_3 &= 7, \\ -2\bar{x}'_2 + 2\bar{x}'_2 + 3x'_3 - \bar{x}'_4 + \bar{x}'_4 - 5\bar{x}'_5 + 5\bar{x}'_5 + x'_7 &= 2, \\ 2x'_1 - 2x'_3 + 2\bar{x}'_4 - 2\bar{x}'_4 - 3\bar{x}'_5 + 3\bar{x}'_5 &= 5, \\ x'_1 \geq 0, \bar{x}'_2 \geq 0, \bar{x}'_2 \geq 0, x'_3 \geq 0, \bar{x}'_4 \geq 0, \bar{x}'_4 \geq 0, \bar{x}'_5 \geq 0, \bar{x}'_5 \geq 0, \\ x'_7 \geq 0. \end{aligned}$$

Непосредственно проверяется, что

$$\begin{aligned} x = (1, 2, 0, 0, -1) \in \mathcal{D}, \quad \varphi(x) = (1, 2, 0, 0, 0, 0, 0, 1, 1) \in \mathcal{D}', \\ \psi(\varphi(x)) = (1, 2, 0, 0, -1) \quad \text{и} \quad f(x) = f'(\varphi(x)) = -7. \end{aligned}$$

2. Рассмотрим теперь задачу

$$\begin{aligned} (\mathcal{D}, f): f(x) = x_1 + x_2 - x_3 + 5x_4 \rightarrow \max, \\ \mathcal{D} = \{x \in R^4: 5x_1 - 6x_2 + 2x_3 = -28, 3x_1 + 4x_2 - 5x_3 + 3x_4 = 3\}. \end{aligned}$$

Следующие задачи $(\mathcal{D}', f')$ (см. п. 4.7) и $(\mathcal{D}'', f'')$ (см. п. 4.8) являются ее каноническими формами:

$$\begin{aligned} (\mathcal{D}', f'): f'(x') = \bar{x}'_1 - \bar{x}'_1 + \bar{x}'_2 - \bar{x}'_2 - \bar{x}'_3 + \bar{x}'_3 + 5\bar{x}'_4 - 5\bar{x}'_4 \rightarrow \max, \\ \mathcal{D}' = \{x' \in R^8: \end{aligned}$$

$$\begin{aligned} 5\bar{x}'_1 - 5\bar{x}'_1 - 6\bar{x}'_2 + 6\bar{x}'_2 + 2\bar{x}'_3 - 2\bar{x}'_3 &= -28, \\ 3\bar{x}'_1 - 3\bar{x}'_1 + 4\bar{x}'_2 - 4\bar{x}'_2 - 5\bar{x}'_3 + 5\bar{x}'_3 + 3\bar{x}'_4 - 3\bar{x}'_4 &= 3, \\ \bar{x}'_1 \geq 0, \bar{x}'_1 \geq 0, \bar{x}'_2 \geq 0, \bar{x}'_2 \geq 0, \bar{x}'_3 \geq 0, \bar{x}'_3 \geq 0, \bar{x}'_4 \geq 0, \bar{x}'_4 \geq 0 \}, \end{aligned}$$

$$(\mathcal{D}'', f''): f''(x'', x_0) = x''_1 + x''_2 - x''_3 + 5x''_4 - 6x_0 \rightarrow \max,$$

$$\begin{aligned} \mathcal{D}'' = \{(x'', x_0) \in R^5: 5x''_1 - 6x''_2 + 2x''_3 - x_0 = -28, \\ 3x''_1 + 4x''_2 - 5x''_3 + 3x''_4 - 5x_0 = 3, x''_1 \geq 0, x''_2 \geq 0, x''_3 \geq 0, \\ x''_4 \geq 0, x_0 \geq 0\}. \end{aligned}$$

4.10. Задача линейного программирования с ограниченными сверху переменными и ее каноническая форма. Задачу линейного программирования

$$\begin{aligned} (\mathcal{D}, f): f(x) = cx \rightarrow \max(\min), \\ \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0, x[j] \leq h[j], j \in J'\} \end{aligned}$$

назовем задачей линейного программирования с ограниченными сверху переменными. Если $J' = 1:n$, то такая задача запи-

сывается в виде

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max(\min),$$

$$\mathcal{D} = \{x \in R^n: Ax = b, 0 \leq x \leq h\}.$$

Этой записью иногда пользуются и в случае, когда $J' \neq 1:n$, полагая $h[j] = +\infty$ для $j \in 1:n \setminus J'$. Примерами задач линейного программирования с ограниченными сверху переменными могут служить задачи, описанные в п. 3.4.

Канонической формой задачи $(\mathcal{D}, f)$ является задача

$$(\mathcal{D}', f'): f'(x', y) = cx' \rightarrow \max(\min),$$

$$\mathcal{D}' = \{(x', y) \in R^{n+n'}: x' \in R^n, y \in R^{n'};$$

$$Ax' = b; x'[j] + y[j] = h[j], j \in J'; x' \geq 0;$$

$$y[j] \geq 0, j \in J'\}, \text{ где } n' = |J'|.$$

Действительно, легко показать, что функции $\varphi: \mathcal{D} \rightarrow \mathcal{D}'$, $\varphi(x) = (x, y(x))$, где $y(x)[j] = h[j] - x[j]$, $j \in J'$, и $\psi: \mathcal{D}' \rightarrow \mathcal{D}$, $\psi(x', y) = x'$ являются эквивалентными преобразованиями.

УПРАЖНЕНИЯ

2.1. См. с. 67. 2.2. См. с. 68. 2.3. См. с. 68.

2.4. См. с. 70. 2.5, а) См. с. 76. 2.5, б) См. с. 77. 2.6. См. с. 82.

2.7. См. с. 83. 2.8. См. с. 84.

2.9. Пусть на множестве $\mathcal{D} \subset R^n$ определены функции f и g . Проверить эквивалентность задач $(\mathcal{D}, f)$ и $(\mathcal{D}, g)$, если $\text{sign}[f(x) - f(y)] = \text{sign}[g(x) - g(y)]$ для любых x и y из $\mathcal{D}$.

2.10. Проверить эквивалентность задач $(\mathcal{D}, f): f(x) = \sum_{j=1}^n x[j]^p \rightarrow \min$, $\mathcal{D} = \{x \in R^n: \sum_{j=1}^n x[j] \geq a, x \geq 0\}$, где $a > 0$, $p > 0$, и $(\mathcal{D}', f'): f'(x) = \sum_{j=1}^n x'[j]^p \rightarrow \min$, $\mathcal{D}' = \{x' \in R^n: \sum_{j=1}^n x'[j] = a, x' \geq 0\}$.

2.11. Пусть для задачи квадратичного программирования (см. п. 2.3) каноническая задача и каноническая форма определяются аналогично тому, как это делается для задачи линейного программирования (см. пп. 2.2 и 4.7). Для всякой ли задачи квадратичного программирования существует каноническая форма?

2.12. Пусть $(\mathcal{D}, f)$ — задача выпуклого программирования. Построить эквивалентную ей задачу максимизации линейной функции на выпуклом множестве.

2.13. Проверить, являются ли эквивалентными задачи $(\mathcal{D}, f): f(x) = \sum_{i=1}^n x[i] \rightarrow \min$, $\mathcal{D} = \{x \in R^{n+}: x[i] + x[i+1] \geq a[i], i \in 1:n-1; x[n] \geq a[n]\}$ и $(\mathcal{D}', f'): f'(x') = \sum_{i=1}^n x'[i] \rightarrow \min$, $\mathcal{D}' = \{x' \in R^{n+}: x'[i] + x'[i+1] = a[i], i \in 1:n-1; x'[n] = a[n]\}$ ($a[i] \geq 0, i \in 1:n$).

Глава 3

КРИТЕРИИ ОПТИМАЛЬНОСТИ

В этой главе выясняются условия, позволяющие установить оптимальность или неоптимальность допустимого плана экстремальной задачи $(\mathcal{D}, f)$. Эти условия называются *критериями оптимальности*.

Среди критериев оптимальности различают необходимые и достаточные условия. Достаточные условия позволяют установить оптимальность или, если они являются и необходимыми, неоптимальность плана, построенного на некотором этапе процесса решения задачи. Необходимые условия также служат для выделения из области допустимых планов более узкого подмножества, в котором содержится оптимальный план.

Особый интерес представляют такие критерии оптимальности, проверка которых осуществляется эффективно. Так, например, «рабочий» критерий оптимальности (см. п. 1.7) сводит проверку оптимальности допустимого плана $\bar{x}$ задачи выпуклого программирования к проверке совместности некоторой системы линейных уравнений и неравенств. Однако и «неэффективные» критерии часто оказываются полезными для качественного анализа и конструирования алгоритмов решения определенного класса задач.

Мы будем рассматривать критерии оптимальности для задач выпуклого, и в частности линейного, программирования. Теоремы этой главы являются теоретической основой вычислительных методов, составляющих содержание ряда следующих глав.

Задачей выпуклого программирования была названа (см. п. 2.2.2) экстремальная задача $(\mathcal{D}, f)$ максимизации (минимизации выпуклой) функции f на выпуклом множестве $\mathcal{D} \subset R^n$. В дальнейшем предполагается, что f определена на R^n и множество допустимых планов $\mathcal{D}$ задается следующей си-

стемой уравнений и неравенств:

$$\mathcal{D} = \{x \in R^{n+} : f_i(x) \leq 0, i \in I_n; f_i(x) = 0, i \in I_y\}, \quad (1)$$

где

$$R^{n+} = \{x \in R^n : x[j] \geq 0, j \in J\}; \quad (2)$$

$J \subset 1:n, I_n \cup I_y = 1:m, I_n \cap I_y = \emptyset; f_i(x) = a_i x - b[i] (a_i \in R^n, b[i] \in R)$, если $i \in I_y$; f_i для $i \in I_n$ — выпуклые функции, определенные на R^n . При этом не исключается, что для некоторых $i \in I_n$ функции f_i могут быть аффинными, т. е. иметь вид $f_i(x) = a_i x - b[i]$. Множество индексов неаффинных функций f_i будем обозначать через I_n ($I_n \subset I_n$).

§ 1. Критерии оптимальности при условиях дифференцируемости

1.1. Допустимые и прогрессивные направления. Вектор $s \in R^n$ называется *допустимым* (возможным) *направлением* в точке $x \in \mathcal{D}$, если существует такое $\lambda' > 0, \lambda' \in R$, что $x + \lambda's \in \mathcal{D}$.

Вектор $s \in R^n$ называется *прогрессивным* (подходящим) *направлением* в точке $x \in \mathcal{D}$, если существует такое $\lambda'' > 0$, что $f(x + \lambda''s) > f(x)$ в задаче максимизации и $f(x + \lambda''s) < f(x)$, в задаче минимизации.

Если множество $\mathcal{D}$ не выпукло, то на отрезке, соединяющем точки x и $x + \lambda's$, могут оказаться точки, не принадлежащие $\mathcal{D}$ (рис. 15); если функция f не вогнута и $f(x + \lambda''s) >$

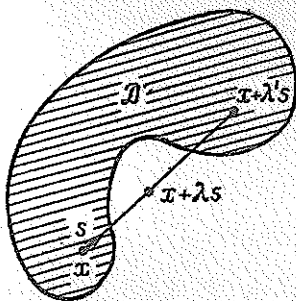


Рис. 15.

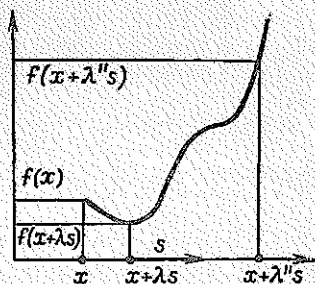


Рис. 16.

$> f(x)$ (рис. 16), то на отрезке, соединяющем точки x и $x + \lambda''s$, могут найтись точки, в которых значение функции меньше, чем $f(x)$. Однако для задачи выпуклого программирования ($\mathcal{D}, f$) такие случаи не имеют места и справедливы следующие свойства допустимых и прогрессивных направлений.

Если s — допустимое направление в точке $x \in \mathcal{D}$ (т. е. $x + \lambda's \in \mathcal{D}, \lambda' > 0$), то $x + \lambda s \in \mathcal{D}$ при любом $\lambda \in [0, \lambda']$.

Если s — прогрессивное направление в точке $x \in \mathcal{D}$ ($f(x + \lambda''s) > f(x)$, $\lambda'' > 0$, в задаче максимизации, $f(x + \lambda''s) < f(x)$, $\lambda'' > 0$, в задаче минимизации), то $f(x + \lambda s) > f(x)$ в задаче максимизации, $f(x + \lambda s) < f(x)$ в задаче минимизации при любом $\lambda \in (0, \lambda'']$.

Доказательство. Пусть $\mathcal{D}$ — выпуклое множество, $x \in \mathcal{D}$, $x + \lambda's \in \mathcal{D}$, $\lambda' > 0$ и $\lambda \in (0, \lambda')$. Тогда $x + \lambda s = (1 - \lambda/\lambda')x + \lambda/\lambda'(x + \lambda's) \in \mathcal{D}$, поскольку $\lambda/\lambda' \in (0, 1)$.

Если f — вогнутая функция, $f(x + \lambda''s) > f(x)$ и $\lambda \in (0, \lambda'')$, то $\lambda/\lambda'' \in (0, 1)$ и $f(x + \lambda s) = f[(1 - \lambda/\lambda'')x + \lambda/\lambda''(x + \lambda''s)] \geq (1 - \lambda/\lambda'')f(x) + \lambda/\lambda''f(x + \lambda''s) > (1 - \lambda/\lambda'')f(x) + \lambda/\lambda''f(x) = f(x)$.

Для выпуклой функции доказательство проводится аналогично. ■

1.2. Критерий оптимальности (в терминах допустимых и прогрессивных направлений). Для оптимальности допустимого плана $\bar{x}$ задачи выпуклого программирования $(\mathcal{D}, f)$ необходимо и достаточно, чтобы никакое допустимое направление в точке $\bar{x}$ не было прогрессивным.

Доказательство. Если $\bar{x}$ — решение задачи $(\mathcal{D}, f)$ и s — допустимое и прогрессивное направление в точке $\bar{x}$, то в силу сформулированных в п. 1.1 свойств допустимых и прогрессивных направлений найдется такое $\lambda > 0$, для которого и $\bar{x} + \lambda s \in \mathcal{D}$, и $f(\bar{x} + \lambda s) > f(\bar{x})$ (без ограничения общности предполагаем, что $(\mathcal{D}, f)$ — задача максимизации) — противоречие.

Обратно, если $\bar{x}$ не является решением задачи, то существует такое $x^0 \in \mathcal{D}$, что $f(x^0) > f(\bar{x})$. Тогда $s = x^0 - \bar{x}$ — допустимое и прогрессивное направление в точке $\bar{x}$. ■

Во всех последующих пунктах настоящего параграфа предполагается непрерывная дифференцируемость функций f и f_i , $i \in 1:m$.

1.3. Условия допустимости направления. Пусть $x \in \mathcal{D}$. Положим

$$I(x) = \{i \in I_n : f_i(x) = 0\}, \quad (3)$$

$$J(x) = \{j \in J : x[j] = 0\}. \quad (4)$$

Для допустимости направления $s \in R^n$ в точке $x \in \mathcal{D}$ необходимо выполнение условий

$$\nabla f_i(x)s \leq 0, \quad i \in I(x), \quad (5)$$

$$\nabla f_i(x)s = a_i s = 0, \quad i \in I_y, \quad (6)$$

$$s[j] \geq 0, \quad j \in J(x). \quad (7)$$

Для допустимости направления $s \in R^n$ в точке $x \in \mathcal{D}$ достаточно выполнения условий (5) — (7) и

$$\nabla f_i(x)s < 0, \quad i \in I(x) \cap I_b. \quad (8)$$

Доказательство. Необходимость. Предположим, что направление $s \in R^n$ допустимо в точке $x \in \mathcal{D}$, так что $x + \lambda s \in \mathcal{D}$ для любого $\lambda \in [0, \lambda']$, $\lambda' > 0$. Поэтому

$$f_i(x + \lambda s) - f_i(x) = f_i(x + \lambda s) \leq 0 \quad \text{для } i \in I(x),$$

а в силу выпуклости функции f_i (см. условие 2) п. 1.8.7)

$$f_i(x + \lambda s) - f_i(x) \geq \lambda \nabla f_i(x) s,$$

и, следовательно, $\nabla f_i(x) s \leq 0$ при $i \in I(x)$.

Для $i \in I_y$ и любого $\lambda \in [0, \lambda']$ имеем $a_i(x + \lambda s) - b[i] = a_i x - b[i] + \lambda a_i s = 0$, и так как $a_i x - b[i] = 0$, то $a_i s = 0$ при $i \in I_y$.

Наконец, для $j \in J(x)$ и $\lambda \in [0, \lambda']$ имеем $0 \leq (x + \lambda s)[j] = x[j] + \lambda s[j] = \lambda s[j]$, а значит, $s[j] \geq 0$ для $j \in J(x)$. ■

Достаточность. Докажем допустимость направления s в точке $x \in \mathcal{D}$, т. е. установим существование такого $\lambda' > 0$, что при любом $\bar{\lambda} \in [0, \bar{\lambda}]$ выполняются соотношения

$$f_i(x + \lambda s) \leq 0, \quad i \in I_{\bar{\lambda}}, \quad (9)$$

$$a_i(x + \lambda s) - b[i] = 0, \quad i \in I_y, \quad (10)$$

$$(x + \lambda s)[j] \geq 0, \quad j \in J. \quad (11)$$

Установим сначала существование такого $\bar{\lambda}_1$, что неравенства (9) выполняются для каждого $i \in I(x) \cap I_{\bar{\lambda}}$ при любом $\lambda \in [0, \bar{\lambda}_1]$. Воспользуемся тем, что для таких i справедливы неравенства (8). Поэтому

$$\lim_{\lambda \downarrow 0} \frac{f_i(x + \lambda s) - f_i(x)}{\lambda} = \lim_{\lambda \downarrow 0} \frac{1}{\lambda} [f_i(x + \lambda s) - f_i(x)] = \nabla f_i(x) s < 0,$$

и, следовательно, для каждого $i \in I(x) \cap I_{\bar{\lambda}}$ существует такое $\lambda_i > 0$, что $f_i(x + \lambda s) < 0$ при $\lambda \in (0, \lambda_i]$. Положим $\bar{\lambda}_1 = \min_{i \in I(x) \cap I_{\bar{\lambda}}} \{\lambda_i\}$. Тогда $\bar{\lambda} > 0$ и $f_i(x + \lambda s) \leq 0$ для $i \in I(x) \cap I_{\bar{\lambda}}$ и $\lambda \in [0, \bar{\lambda}_1]$.

Если $i \in I(x) \setminus I_{\bar{\lambda}}$, то $f_i(x) = a_i x - b[i] = 0$, $\nabla f_i(x) s = a_i s \leq 0$ (см. (5)), и, следовательно, $a_i(x + \lambda s) - b[i] = \lambda a_i s \leq 0$ при любом $\lambda \geq 0$, и в частности при $\lambda \in [0, \bar{\lambda}_1]$.

Для $i \in I_{\bar{\lambda}} \setminus I(x)$ имеем $f_i(x) < 0$, и в силу непрерывности функции f_i найдется такое λ'_i , что $f_i(x + \lambda s) < 0$ при $\lambda \in [0, \lambda'_i]$. Положим $\bar{\lambda}_2 = \min_{i \in I_{\bar{\lambda}} \setminus I(x)} \{\lambda'_i\}$. Тогда $\bar{\lambda}_2 > 0$ и $f_i(x + \lambda s) < 0$ для любых $i \in I_{\bar{\lambda}} \setminus I(x)$ и $\lambda \in [0, \bar{\lambda}_2]$.

* Напоминаем, что в принятых обозначениях $(x + \lambda s)[j]$ — j -я компонента вектора $x + \lambda s$.

Так как $x[j] > 0$, $j \in J \setminus J(x)$, то можно выбрать такое $\bar{\lambda}_3 > 0$, что неравенства $(x + \lambda s)[j] = x[j] + \lambda s[j] \geq 0$ выполняются при любом $\lambda \in [0, \bar{\lambda}_3]$ и любом $j \in J \setminus J(x)$.

Заметим теперь, что в силу (6) равенства (10) выполняются при любом $\lambda \in R$. В силу (7) $(x + \lambda s)[j] \geq 0$ для $j \in J(x)$ и $\lambda \geq 0$.

Таким образом, установлено, что соотношения (9)–(11) выполняются при любом $\lambda \in [0, \bar{\lambda}]$, где $\bar{\lambda} = \min\{\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3\} > 0$ (т. е. $x + \lambda s \in \mathcal{D}$ при любом $\lambda \in [0, \bar{\lambda}]$). ■

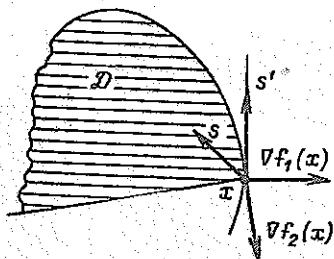


Рис. 17.

Доказанные выше условия допустимости направления имеют простой геометрический смысл: допустимое направление должно образовывать неострые углы с внешними нормальными к граничным гиперповерхностям, содержащим точку x (на рис. 17 направление s является допустимым в точке x). Необходимые условия (5)–(7) пришлось усилить условием (8), так как на-

правление, касательное к выпуклой гиперповерхности в точке x , может оказаться недопустимым. Например, для направления s' (см. рис. 17) $\nabla f_i(x)s' \leq 0$, $i \in 1:2$, но оно не является допустимым.

1.4. Условия прогрессивности направления. Для прогрессивности направления s в точке $x \in \mathcal{D}$ необходимо и достаточно выполнение неравенства

$$\nabla f(x)s > 0 \quad (12)$$

в задаче максимизации и

$$\nabla f(x)s < 0 \quad (12')$$

в задаче минимизации.

Доказательство мы проведем для случая задачи максимизации. Если s прогрессивно, т. е. найдется такое $\lambda > 0$, что $f(x + \lambda s) > f(x)$, то неравенство (12) в силу вогнутости f следует из соотношения $0 < f(x + \lambda s) - f(x) \leq \lambda \nabla f(x)s$ (см. замечание 2 п. 1.8.7).

Обратно, поскольку (см. п. 1.2.10)

$$\lim_{\lambda \downarrow 0} \frac{1}{\lambda} [f(x + \lambda s) - f(x)] = \nabla f(x)s > 0,$$

то очевидно найдется такое $\lambda' > 0$, что неравенство $f(x + \lambda s) - f(x) > 0$ выполняется при любом $\lambda \in (0, \lambda']$. ■

1.5. Условие регулярности. Рассматриваемые ниже необходимые условия оптимальности имеют место лишь при выполнении определенных условий, которые будем называть усло-

виями регулярности. Эти условия могут различаться и по содержанию, и по форме. Для наших целей наиболее удобными оказываются предположения о «телесности» множества $\mathcal{D}$ относительно выпуклых ограничений.

Пусть множество $\mathcal{D}$ задано системой ограничений (1) и $i_0 \in I_n$ ($j_0 \in J$). Будем говорить, что множество $\mathcal{D}$ телесно относительно ограничения i_0 (j_0), или, что то же, выполняется условие С (С — условие Слейтера) относительно i_0 (j_0), если существует такой план $\hat{x} \in \mathcal{D}$, что $f_{i_0}(\hat{x}) < 0$ (соответственно $\hat{x}[j_0] > 0$).

Если условие С выполняется для каждого $i \in I \subset I_n$ и каждого $j \in J' \subset J$ (т. е. найдутся такие векторы $\hat{x}_i \in \mathcal{D}$, $i \in I$, и $\hat{x}'_j \in \mathcal{D}$, $j \in J'$, что $f_i(\hat{x}_i) < 0$ и $\hat{x}'_j[j] > 0$), то существует такой $\hat{x} \in \mathcal{D}$, что $f_i(\hat{x}) < 0$, $\hat{x}[j] > 0$ для всех $i \in I$, $j \in J'$ соответственно.

В самом деле, положим

$$\hat{x} = \frac{1}{|I| + |J'|} (\sum_{i \in I} \hat{x}_i + \sum_{j \in J'} \hat{x}'_j).$$

Тогда $\hat{x} \in \mathcal{D}$, $\hat{x}[j] > 0$, $j \in J'$, и для каждого $i \in I$ в силу выпуклости функций f_i выполняются неравенства

$$\begin{aligned} f_i(\hat{x}) &\leq \frac{1}{|I| + |J'|} (\sum_{i \in I} f_i(\hat{x}_i) + \sum_{j \in J'} f_i(\hat{x}'_j)) \leq \\ &\leq \frac{1}{|I| + |J'|} f_i(\hat{x}_i) < 0. \blacksquare \end{aligned}$$

Доказанное утверждение позволяет говорить о выполнении условия С относительно системы ограничений $I \subset I_n$ и $J' \subset J$.

1.6. Критерий оптимальности. Предположим, что для множества ограничений I_n выполняется условие С. Тогда для оптимальности допустимого плана $\bar{x}$ задачи $(\mathcal{D}, f)$ необходимо и достаточно, чтобы для любого $s \in R^n$, удовлетворяющего соотношениям (5) — (7) при $x = \bar{x}$:

$$\nabla f_i(\bar{x})s \leq 0, \quad i \in I(\bar{x}), \quad (5')$$

$$a_i s = 0, \quad i \in I_y, \quad (6')$$

$$s[j] \geq 0, \quad j \in J(\bar{x}), \quad (7')$$

имело место неравенство $\nabla f(\bar{x})s \leq 0$ в задаче максимизации ($\nabla f(\bar{x})s \geq 0$ в задаче минимизации).

Доказательство.* Достаточность. Пусть s — допустимое направление в точке $\bar{x}$. Тогда s удовлетворяет соотношениям (5') — (7') (см. п. 1.3) и, поскольку $\nabla f(\bar{x})s \leq 0$, не является прогрессивным (см. п. 1.4), а значит (см. п. 1.2), $\bar{x}$ — решение задачи $(\mathcal{D}, f)$. $\blacksquare$

* Проводится для задачи максимизации.

Необходимость. Пусть $\bar{x}$ — решение задачи $(\mathcal{D}, f)$ и s — вектор из R^n , удовлетворяющий условиям $(5') - (7')$. Если направление s допустимо в точке $\bar{x}$, то (см. п. 1.2) оно не может быть прогрессивным, и, следовательно (см. п. 1.4), $\nabla f(\bar{x})s \leq 0$. Но, как уже отмечалось в п. 1.3 (см. рис. 3), направление s , удовлетворяющее только условиям $(5') - (7')$, может не быть допустимым в точке $\bar{x}$. Поэтому для доказательства теоремы построим некоторое множество допустимых в точке $\bar{x}$ направлений $\{s(\lambda)\}$, «близких» к направлению s , так что неравенство $\nabla f(\bar{x})s \leq 0$ можно будет получить из неравенств $\nabla f(\bar{x})s(\lambda) \leq 0$ предельным переходом.

Для построения векторов $s(\lambda)$ воспользуемся условием С: пусть $\hat{x}$ — такой вектор из $\mathcal{D}$, что $f_i(\hat{x}) < 0$, $i \in I_{\bar{x}}$; положим $s(\lambda) = s + \lambda(\hat{x} - \bar{x})$. Покажем, что $s(\lambda)$ при любом $\lambda > 0$ — допустимое направление в точке $\bar{x}$. Для этого в силу п. 1.3 достаточно проверить справедливость соотношений

$$\nabla f_i(\bar{x})s(\lambda) \leq 0, \quad i \in I(\bar{x}), \quad (5'')$$

$$a_i s(\lambda) = 0, \quad i \in I_y, \quad (6'')$$

$$s(\lambda)[j] \geq 0, \quad j \in J(\bar{x}), \quad (7'')$$

$$\nabla f_i(\bar{x})s(\lambda) < 0, \quad i \in I(\bar{x}) \cap I_{\bar{x}}. \quad (8'')$$

Покажем, что выполняются соотношения $(5'')$ и $(8'')$. Пусть $i \in I(\bar{x})$, т. е. $i \in I_{\bar{x}}$ и $f_i(\bar{x}) = 0$. Тогда в силу выпуклости функции f_i

$$\begin{aligned} \nabla f_i(\bar{x})s(\lambda) &= \nabla f_i(\bar{x})s + \lambda \nabla f_i(\bar{x})(\hat{x} - \bar{x}) \leq \\ &\leq \nabla f_i(\bar{x})s + \lambda [f_i(\hat{x}) - f_i(\bar{x})] = \nabla f_i(\bar{x})s + \lambda f_i(\hat{x}). \end{aligned}$$

Но $\nabla f_i(\bar{x})s \leq 0$ для $i \in I(\bar{x})$ (условие $(5')$), $f_i(\hat{x}) \leq 0$ при $i \in I_{\bar{x}}(\hat{x} \in \mathcal{D})$, $f_i(\hat{x}) < 0$ при $i \in I_{\bar{x}}$, $\lambda > 0$. Следовательно, $\nabla f_i(\bar{x})s(\lambda) \leq 0$ при $i \in I(\bar{x})$ и $\nabla f_i(\bar{x})s(\lambda) < 0$ при $i \in I(\bar{x}) \cap I_{\bar{x}}$.

Условия $(6'')$ и $(7'')$ также непосредственно следуют из определения $s(\lambda)$: для $i \in I_y$

$$\begin{aligned} a_i s(\lambda) &= a_i s + \lambda(a_i \hat{x} - a_i \bar{x}) = a_i s + \lambda(a_i \hat{x} - b[i]) - \\ &- \lambda(a_i \bar{x} - b[i]) = 0, \end{aligned}$$

так как $a_i s = 0$ при $i \in I_y$ (условие $(6')$), $a_i \hat{x} - b[i] = 0$, $a_i \bar{x} - b[i] = 0$ (векторы $\hat{x}$ и $\bar{x}$ из $\mathcal{D}$); для $j \in J(\bar{x})$, $\lambda > 0$

$$s(\lambda)[j] = s[j] + \lambda(\hat{x}[j] - \bar{x}[j]) \geq 0,$$

так как $s[j] \geq 0$ при $j \in J(\bar{x})$ (условие $(7')$), $\bar{x}[j] = 0$ для $j \in J(\bar{x})$ и $\hat{x}[j] \geq 0$ при $j \in J(\hat{x} \in \mathcal{D})$.

Итак, $s(\lambda)$ — допустимое направление в точке $\bar{x}$ при любом $\lambda > 0$. Так как $\bar{x}$ — решение задачи $(\mathcal{D}, f)$, то ни при каком $\lambda > 0$ направление $s(\lambda)$ не является прогрессивным, и, следовательно, при любом $\lambda > 0$

$$\nabla f(\bar{x})s(\lambda) = \nabla f(\bar{x})s + \lambda \nabla f(\bar{x})(\hat{x} - \bar{x}) \leq 0.$$

Переходя в последнем неравенстве к пределу при $\lambda \rightarrow 0$, получаем $\nabla f(\bar{x})s \leq 0$, что и требовалось доказать. ■

Замечание 1. Если множество $I_{\bar{x}} = \emptyset$, как, например, в задачах линейного и квадратичного программирования (см. пп. 2.2.2 и 2.2.3), то направление s , удовлетворяющее условиям (5')—(7'), допустимо в точке $\bar{x} \in \mathcal{D}$ (см. п. 1.3) и неравенство $\nabla f(\bar{x})s \leq 0$ в задаче максимизации ($\nabla f(\bar{x})s \geq 0$ в задаче минимизации) очевидно необходимо для оптимальности $\bar{x}$.

Замечание 2. При доказательстве теоремы было установлено, что при любом векторе $s \in R^n$, удовлетворяющем условиям (5')—(7'), и любом $\lambda > 0$ для вектора $s(\lambda) = s + \lambda(\bar{x} - \bar{x})$ выполняются соотношения (5'')—(8''). Таким образом, если s — допустимое направление в точке $\bar{x}$, то при любом $\lambda > 0$ вектор $s(\lambda)$ является допустимым направлением в точке $\bar{x}$, удовлетворяющим условию (8''). Этим утверждением воспользуемся в § 4 гл. 9.

1.7. «Рабочий» критерий оптимальности — разложение градиента. Установленные в п. 1.6 условия оптимальности допустимого плана $\bar{x}$ задачи выпуклого программирования $(\mathcal{D}, f)$ в случае выполнения для множества ограничений $I_{\bar{x}}$ условия регулярности C позволяют свести проверку оптимальности допустимого плана $\bar{x}$ и поиск решения задачи к следующим процедурам: проверить, следует ли из неравенств (5')—(7') неравенство $\nabla f(\bar{x})s \leq 0$ в задаче максимизации ($\nabla f(\bar{x})s \geq 0$ в задаче минимизации); в противном случае построить допустимое прогрессивное направление $\bar{s}$ в точке $\bar{x}$ и перейти в допустимую точку $\bar{x} + \bar{\lambda}\bar{s}$ при некотором $\bar{\lambda} > 0$.

Реализация такого подхода к решению некоторых классов задач выпуклого программирования будет осуществлена в ряде следующих глав. В настоящем пункте воспользуемся теоремой Фаркаша (см. п. 1.7.4) для приведения условий оптимальности к виду, допускающему более эффективное их применение: задача проверки оптимальности допустимого плана сводится к проверке совместности некоторой системы линейных уравнений и неравенств. Название «рабочий критерий» подчеркивает возможность и целесообразность использования такого подхода как в теоретических исследованиях, так и при конструировании эффективных вычислительных алгоритмов решения экстремальных задач выпуклого программирования.

Заменим каждое из равенств (6') равносильной ему системой неравенств $a_i s \leq 0$, $(-a_i)s \leq 0$ и рассмотрим матрицу A , строками которой являются следующие векторы из R^n :

$$\begin{aligned} \nabla f_i(\bar{x}), i \in I(\bar{x}); a_i \text{ и } -a_i, i \in I_j; \\ -e_j = (0, \dots, 0, \underbrace{-1}_{j\text{-я}}, 0, \dots, 0), j \in J(\bar{x}). \end{aligned}$$

Выполнение условий критерия п. 1.6 означает, что при любом $s \in R^n$, для которого $As \leq 0$, выполняется неравенство $\nabla f(\bar{x})s \leq 0$ ($-\nabla f(\bar{x})s \leq 0$).^{*} Но в силу теоремы Фаркаша это имеет место в том и только в том случае, когда вектор $\nabla f(\bar{x})$ ($-\nabla f(\bar{x})$) может быть представлен в виде неотрицательной линейной комбинации строк матрицы A . Следовательно, возможность представления вектора $\nabla f(\bar{x})$ ($-\nabla f(\bar{x})$) в виде неотрицательной линейной комбинации строк матрицы A необходима и достаточна для оптимальности плана $\bar{x} \in \mathcal{D}$.

Итак, если для множества ограничений I_α выполняется условие C , то для оптимальности допустимого плана $\bar{x}$ задачи выпуклого программирования необходимо и достаточно существование таких чисел $u[i] \geq 0, i \in I(\bar{x}); u'[i] \geq 0, u''[i] \geq 0, i \in I_y; v[j] \geq 0, j \in J(\bar{x})$, что

$$\pm \nabla f(\bar{x}) = \sum_{i \in I(\bar{x})} u[i] \nabla f_i(\bar{x}) + \sum_{i \in I_y} (u'[i] a_i + u''[i] (-a_i)) + \sum_{j \in J(\bar{x})} v[j] (-e_j) \quad (13)$$

(знак «+» в левой части (13) относится к случаю задачи максимизации, а знак «-» — к случаю задачи минимизации).

Полагая $u[i] = 0$ при $i \in I_\alpha \setminus I(\bar{x})$, $v[j] = 0$ при $j \in 1:n \setminus J(\bar{x})$, $u[i] = u'[i] - u''[i]$ при $i \in I_y$ и имея в виду, что любое вещественное число может быть представлено (естественно, не единственным способом) в виде разности двух неотрицательных чисел, приходим к следующей формулировке *рабочего критерия оптимальности* допустимого плана $\bar{x}$: для оптимальности вектора $\bar{x} \in \mathcal{D}$ необходимо и достаточно, чтобы существовали такие $u \in R^m, v \in R^n$, что

$$\pm \nabla f(\bar{x}) = \sum_{i=1}^m u[i] \nabla f_i(\bar{x}) - \sum_{j=1}^n v[j] e_j, \quad (14)$$

$$u[i] \geq 0, \quad i \in I_\alpha; \quad v[j] \geq 0, \quad j \in J, \quad (15)$$

$$u[i] = 0, \quad i \in I_\alpha \setminus I(\bar{x}), \quad (16)$$

$$v[j] = 0, \quad j \in J(\bar{x}) \quad (17)$$

(знак «+» в левой части (14) относится к задаче максимизации, а знак «-» — к задаче минимизации).

Еще раз отметим, что необходимые условия критериев оптимальности доказаны в предположении выполнения условия C относительно множества I_α . Существование этого предположения будет показана в п. 1.10 (см. пример 4).

Геометрическая интерпретация этого критерия такова: при выполнении условия C для множества I_α задачи выпуклого программирования $(\mathcal{D}, f)$ план $\bar{x} \in \mathcal{D}$ оптимален тогда и только тогда, когда вектор $\nabla f(\bar{x})$ в задаче максимизации ($-\nabla f(\bar{x})$)

* Для задачи минимизации

в задаче минимизации) принадлежит многогранному выпуклому конусу, натянутому на внешние нормали к граничным гиперповерхностям множества $\mathcal{D}$, содержащим точку $\bar{x}$.

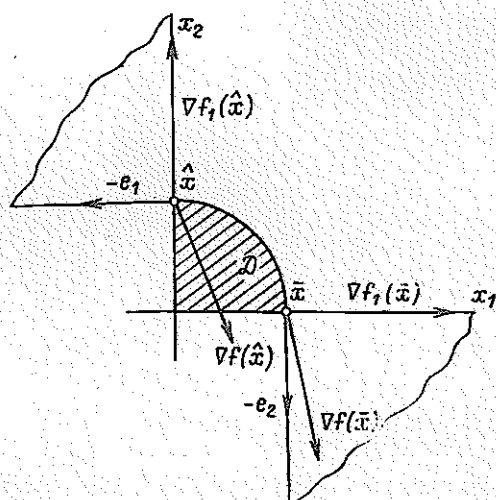


Рис. 18.

На рис. 18 изображено множество $\mathcal{D}$ допустимых планов задачи выпуклого программирования

$$(\mathcal{D}, f): f(x) = \frac{1}{2}x_1 - \frac{1}{8}x_1^2 - \frac{3}{2}x_2 \rightarrow \max,$$

$$\mathcal{D} = \{x \in R^2: f_1(x) = x_1^2 + x_2^2 - 1 \leq 0, x_1 \geq 0, x_2 \geq 0\}.$$

Проверим оптимальность допустимых планов $\bar{x} = (1, 0)$ и $\hat{x} = (0, 1)$. Имеем $\nabla f(x) = (\frac{1}{2} - \frac{1}{4}x_1, -\frac{3}{2})$, $\nabla f_1(x) = (2x_1, 2x_2)$. Внешними нормальными к линиям, ограничивающим $\mathcal{D}$ и проходящим через $\bar{x}$, являются векторы $\nabla f_1(\bar{x}) = (2, 0)$ и $-e_2 = (0, -1)$. Так как вектор $\nabla f(\bar{x}) = (\frac{1}{4}, -\frac{3}{2})$ содержится в конусе, натянутом на векторы $\nabla f_1(\bar{x})$ и $-e_2$, то $\bar{x}$ — оптимальный план. План $\hat{x}$ неоптимален, поскольку вектор $\nabla f(\hat{x}) = (\frac{1}{2}, -\frac{3}{2})$ не принадлежит конусу, натянутому на внешние нормали $\nabla f_1(\hat{x}) = (0, 2)$ и $-e_1 = (-1, 0)$ к линиям, ограничивающим $\mathcal{D}$ и проходящим через точку $\hat{x}$.

Рассмотрим теперь применение рабочего критерия к некоторым частным случаям задач выпуклого программирования.

1.8. Задача квадратичного программирования. В п. 2.2.3 задачей квадратичного программирования была названа задача

$$(\mathcal{D}, f): f(x) = cx + x^T Dx \rightarrow \max,$$

$$\mathcal{D} = \{x \in R^n: a_i x \leq b[i], i \in I_R; a_i x = b[i], i \in I_Y\}.$$

где $c^T = (c[1], c[2], \dots, c[n])$; a_i — i -я строка $m \times n$ матрицы A ; D — неположительно определенная матрица порядка n ; $I_n \cup I_y = 1:m$, $I_n \cap I_y = \emptyset$; $J \subset 1:n$.

Критерий п. 1.7 может быть применен к задаче квадратичного программирования, так как условия, при которых он был установлен, выполняются: f — дифференцируемая вогнутая (см. п. 2.2.3) функция; f_i , $i \in 1:m$, — аффинные функции, а $I_s = \emptyset$.

Легко проверить, что $\nabla f(x) = 2Dx + c$, $\nabla f_i(x) = a_i$, $i \in 1:m$. Тогда, используя (14) — (17), получаем:

для оптимальности допустимого плана $\bar{x}$ задачи квадратичного программирования $(\mathcal{D}, f)$ необходимо и достаточно существование таких $u \in R^m$, $v \in R^n$, что

$$2D\bar{x} + c = A^T u - I v, \quad (18)$$

$$u[i] \geq 0, \quad i \in I_n, \quad v[j] \geq 0, \quad j \in J, \quad (19)$$

$$u[i] = 0, \quad i \in I_n \setminus I(\bar{x}), \quad (20)$$

$$v[j] = 0, \quad j \in J(\bar{x}), \quad (21)$$

где через I обозначена единичная матрица порядка n .

Особый интерес представляет случай $I_y = 1:m$, $J = 1:n$. Задача

$$(\mathcal{D}, f): f(x) = cx + x^T D x \rightarrow \max, \quad \mathcal{D} = \{x \in R^{n+}: Ax = b\} \quad (22)$$

называется канонической задачей квадратичного программирования.

Для оптимальности допустимого плана $\bar{x}$ задачи (22) необходимо и достаточно существование таких $u \in R^m$ и $v \in R^n$, что

$$2D\bar{x} + c = A^T u - I v, \quad (23)$$

$$v[j] \geq 0, \quad j \in 1:n, \quad (24)$$

$$v[j] = 0, \quad j \in J(\bar{x}),$$

т. е.

$$v[j] = 0, \quad \text{если } \bar{x}[j] > 0. \quad (25)$$

Эти условия, как нетрудно видеть, могут быть переформулированы следующим образом: для оптимальности $\bar{x} \in \mathcal{D}$ необходимо и достаточно существование таких $u \in R^m$ и $v \in R^n$, что

$$2D\bar{x} - A^T u + I v = -c, \quad (23')$$

$$v \geq 0, \quad (24')$$

$$v\bar{x} = 0. \quad (25')$$

1.9. Задача линейного программирования. В п. 2.2.2 общей задачей линейного программирования была названа задача

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max,$$

$$\mathcal{D} = \{x \in R^n: a_i x \leq b[i], i \in I_n; a_i x = b[i], i \in I_y\},$$

где $c = (c[1], c[2], \dots, c[n])$, a_i — i -я строка $m \times n$ матрицы A , $I_n \cup I_y = 1:m$, $I_n \cap I_y = \emptyset$, $J \subset 1:n$.

Задача линейного программирования является задачей выпуклого программирования, к которой очевидно может быть применен критерий п. 1.7. Имеем $\nabla f(x) = c$, $\nabla f_i(x) = a_i$, $i \in 1:m$, так что для оптимальности допустимого плана $\bar{x}$ задачи $(\mathcal{D}, f)$ необходимо и достаточно существование таких $u \in R^m$, $v \in R^n$, что

$$c = \sum_{i=1}^m u[i] a_i - \sum_{j=1}^n v[j] e_j, \quad (26)$$

$$u[i] \geq 0, \quad i \in I_n, \quad (27)$$

$$v[j] \geq 0, \quad j \in J, \quad (28)$$

$$u[i] = 0, \text{ если } a_i \bar{x} < b[i], \quad (29)$$

$$v[j] = 0, \text{ если } j \notin J \text{ или } j \in J \text{ и } \bar{x}[j] > 0. \quad (30)$$

Равенство (26) можно записать в матричной форме:

$$c = uA - vI, \quad (26')$$

где I — единичная матрица порядка n , а условия (26'), (27) — (30) переформулировать следующим образом: для оптимальности $\bar{x} \in \mathcal{D}$ необходимо и достаточно существование такого $u \in R^m$, что

$$c[j] \leq ua^j, \quad j \in J, \quad (31)$$

$$c[j] = ua^j, \quad j \notin J \text{ или } j \in J \text{ и } \bar{x}[j] > 0, \quad (32)$$

$$u[i] \geq 0, \quad i \in I_n, \quad (33)$$

$$u[i] = 0, \text{ если } a_i \bar{x} < b[i] \quad (34)$$

(напомним, что через a^j обозначен j -й столбец матрицы A).

Действительно, из (26') и (28) следует (31), из (26') и (30) следует (32), а (33) и (34) совпадают с условиями (27) и (29).

Обратно, пусть u — такой вектор из R^m , для которого справедливы (31) — (34). Положим $v[j] = ua^j - c[j]$, $j \in 1:n$, $v = (v[1], v[2], \dots, v[n])$ и для векторов $u \in R^m$, $v \in R^n$ условия (26'), (27) — (30) очевидно выполняются. ■

Особый интерес представляет случай канонической задачи линейного программирования (см. п. 2.4.7) $(\mathcal{D}, f): f(x) = cx \rightarrow \max$, $\mathcal{D} = \{x \in R^n: Ax = b\}$, где множества $I_n = \emptyset$, $I_y = 1:m$, $J = 1:n$. Тогда в силу (31) и (32) имеем: для оптимальности допустимого плана $\bar{x}$ задачи $(\mathcal{D}, f)$ необходимо и до-

статочно, чтобы существовал такой вектор $u \in R^m$, что

$$c \leq uA, \quad (35)$$

$$c[j] = ua^j, \text{ если } x[j] > 0. \quad (36)$$

1.10. Примеры применения рабочего критерия оптимальности. Критерии оптимальности, сформулированные выше, сводят проверку оптимальности или неоптимальности допустимого плана к проверке совместности некоторой системы линейных уравнений и неравенств. Приведем несколько примеров.

1. Легко проверить, что экстремальная задача

$$(\mathcal{D}, f): f(x) = -3x_1^2 - 2x_2^2 - 6x_3^2 + 2x_1x_2 + 4x_2x_3 + 13x_1 + 22x_3 + 10 \rightarrow \max, \quad \mathcal{D} = \{x \in R^3: x_1^2 + x_2^2 + x_3^2 \leq 2, \\ x_1^4 + 2x_2 + x_3 \leq 3, 4x_1 + 6x_2 + x_3 \leq 5, x_1 - 2x_3 = -1, x_2 \geq 0, x_3 \geq 0\}$$

является задачей выпуклого программирования; функции f, f_1, f_2, f_3, f_4 дифференцируемы,

$$\nabla f(x) = (-6x_1 + 2x_2 + 13, 2x_1 - 4x_2 + 4x_3, 4x_2 - 12x_3 + 22), \\ \nabla f_1(x) = (2x_1, 2x_2, 2x_3), \nabla f_2(x) = (4x_1^3, 2, 1), \nabla f_3(x) = (4, 6, 1) = a_3, \\ \nabla f_4(x) = (1, 0, -2) = a_4, \quad I_u = \{1, 2, 3\}, \quad I_v = \{4\}, \quad I_w = \{1, 2\}, \\ J = \{2, 3\}.$$

Поскольку для $\hat{x} = (0, 0, 1/2)$ из $\mathcal{D}$ имеем $f_1(\hat{x}) = 0 + 0 + 1/4 - 2 < 0$, $f_2(\hat{x}) = 0 + 0 + 1/2 - 3 < 0$, то выполняется условие С, а следовательно, для проверки оптимальности допустимого плана этой задачи, например $\bar{x} = (1, 0, 1)$, можно применить рабочий критерий оптимальности. Установив, что $I(\bar{x}) = \{1, 3\}$, $J(\bar{x}) = \{2\}$, заключаем, что вектор оптимален в том и только в том случае, если совместна следующая система уравнений и неравенств:

$$\nabla f(\bar{x}) = u_1 \nabla f_1(\bar{x}) + u_3 a_3 + u_4 a_4 - v_2 e_2,$$

$$u_1 \geq 0, \quad u_3 \geq 0, \quad v_2 \geq 0,$$

или

$$7 = 2u_1 + 4u_3 + u_4,$$

$$6 = 6u_3 - v_2,$$

$$10 = 2u_1 + u_3 - 2u_4,$$

$$u_1 \geq 0, \quad u_3 \geq 0, \quad v_2 \geq 0.$$

Эта система совместна и имеет решение $u_1 = 1, u_3 = 2, u_4 = -3, v_2 = 6$. Следовательно, вектор $\bar{x}$ является оптимальным планом задачи $(\mathcal{D}, f)$.

В свою очередь план $\hat{x}$ не является решением рассматриваемой задачи, поскольку система

$$\begin{aligned} 13 &= u_4, \\ 2 &= -v_2, \\ 16 &= -2u_4, \quad v_2 \geq 0, \end{aligned}$$

совместность которой равносильна оптимальности $\hat{x}$, не имеет решения.

2. Рассмотрим задачу квадратичного программирования

$$(\mathcal{D}, f) = 6x_1^2 + x_2^2 - 2x_1x_2 + 9x_1 - 3x_2 \rightarrow \min,$$

$$\mathcal{D} = \{x \in R^2: -x_1 + x_2 \leq 3, x_1 - 5x_2 \leq -15, 0 \leq x_1 \leq 5\}.$$

Для проверки оптимальности ее допустимого плана можно использовать условия оптимальности (23)–(25), полученные в п. 1.8. Так как

$$\begin{aligned} D &= \begin{pmatrix} 6 & -1 \\ -1 & 1 \end{pmatrix}, \quad c = \begin{pmatrix} 9 \\ -3 \end{pmatrix}, \quad \text{то } \nabla f(x) = 2Dx + c = \\ &= \begin{pmatrix} 12x_1 - 2x_2 + 9 \\ -2x_1 + 2x_2 - 3 \end{pmatrix}. \end{aligned}$$

Векторы $x^0 = (5, 4)$ и $\bar{x} = (0, 3)$ являются допустимыми планами. Проверим их оптимальность. Для оптимальности x^0 необходима и достаточна (см. (14), (23)–(25)) совместность системы

$$\begin{aligned} \nabla f(x^0) &= \begin{pmatrix} -61 \\ 5 \end{pmatrix} = u_2 \begin{pmatrix} 1 \\ -5 \end{pmatrix} + u_3 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ u_2 &\geq 0, \quad u_3 \geq 0. \end{aligned}$$

Поскольку она несовместна, вектор x^0 не является решением задачи.

Аналогично для проверки оптимальности $\bar{x}$ достаточно установить совместность системы

$$\begin{aligned} -3 &= -u_1 + u_2 - v_1, \\ -3 &= u_1 - 5u_2, \\ u_1 &\geq 0, \quad u_2 \geq 0, \quad v_1 \geq 0. \end{aligned}$$

Так как $u_1 = 2, u_2 = 1, v_1 = 2$ — ее решение, то $\bar{x}$ — оптимальный план.

3. Используем теперь условия (31)–(34), полученные в п. 1.9, для проверки оптимальности планов $x^0 = (0, 1, 4, 0, 0)$ и

$\bar{x} = (-1, 0, 3, 2, 1)$ задачи линейного программирования

$$(\mathcal{D}, f): f(x) = x_1 + 2x_2 - 5x_3 - 7x_4 + 4x_5 \rightarrow \min,$$

$$\mathcal{D} = \{x \in R^5: x_1 - x_2 + 4x_3 + 6x_4 + x_5 \leq 26, 2x_1 + x_2 - x_4 + 5x_5 = 1, \\ x_1 - x_2 + 6x_3 + 6x_4 \leq 29, 2x_2 - x_3 + x_5 = -2, x_2 \geq 0, x_4 \geq 0\}.$$

Подставляя векторы $x^0, \bar{x}$ в систему ограничений, задающих множество $\mathcal{D}$, устанавливаем их допустимость. Так как $f(x^0) = \emptyset$, то $u_1 = u_3 = 0$, и для проверки оптимальности или неоптимальности плана x^0 достаточно установить совместность или несовместность системы (см. (31)–(33) и (14))

$$\begin{aligned} -1 &= 2u_2, \\ -2 &= u_2 + 2u_4, \\ 5 &= -u_4, \\ 7 &= -u_2, \\ -4 &= 5u_2 + u_4. \end{aligned}$$

Поскольку система несовместна, план x^0 не является решением задачи $(\mathcal{D}, f)$.

Аналогично для оптимальности плана $\bar{x}$ достаточно совместности системы

$$\begin{aligned} -1 &= 2u_2 + u_3, \\ -2 &\leq u_2 - u_3 + 2u_4, \\ 5 &= 6u_3 - u_4, \\ 7 &= -u_2 + 6u_3, \\ -4 &= 5u_2 + u_3, \\ u_3 &\geq 0. \end{aligned}$$

Она совместна ($u_2 = -1, u_3 = 1, u_4 = 1$). Следовательно, $\bar{x}$ — решение задачи $(\mathcal{D}, f)$.

1.11. Критерии оптимальности для некоторых частных случаев задач линейного программирования. Рассмотрим примеры применения условий оптимальности (31)–(34) и (35), (36) в некоторых задачах линейного программирования.

1°. Задача Канторовича

$$(\mathcal{D}, f): f(x) = \min_{j \in 1:n} \frac{1}{k[j]} \sum_{l=1}^L a[m+j, l] x[l] \rightarrow \max,$$

$$\mathcal{D} = \{x \in R^L: \sum_{l=1}^L a[i, l] x[l] \geq b[i], i \in 1:m; x \geq 0\}$$

рассматривалась в пп. 2.3.6 и 2.4.6. В силу свойства б) эквивалентных экстремальных задач (см. п. 2.4.3) допустимый план $\bar{x}$ задачи $(\mathcal{D}, f)$ оптимален тогда и только тогда, когда вектор

$\Phi(\bar{x}) = (\bar{x}, f(\bar{x}))$ является решением линейной формы этой задачи

$$(\mathcal{D}', f') : f'(x', z) = z \rightarrow \max,$$

$$\mathcal{D}' = \{(x', z) \in R^{L+1} : x' \in R^L, z \in R;$$

$$\sum_{l=1}^L a[l, l] x'[l] \geq b[l], l \in 1:m;$$

$$\frac{1}{k[j]} \sum_{l=1}^L a[m+j, l] x'[l] \geq z, j \in 1:n; x' \geq 0\},$$

или, что то же:

$$\mathcal{D}' = \{(x', z) \in R^{L+1} : x' \in R^L, z \in R,$$

$$\sum_{l=1}^L (-a[l, l]) x'[l] \leq -b[l], l \in 1:m;$$

$$\sum_{l=1}^L (-a[m+j, l]) x'[l] + k[j] z \leq 0, j \in 1:n, x' \geq 0\}.$$

Задача $(\mathcal{D}', f')$ есть задача линейного программирования, в которой $I_N = 1:N, N = m+n, I_y = \emptyset, J = 1:L,$

$$a_l = (-a[l, 1], \dots, -a[l, L], 0), l \in 1:m,$$

$$a_{m+j} = (-a[m+j, 1], \dots, -a[m+j, L], k[j]), j \in 1:n,$$

$$c = (\underbrace{0, 0, \dots, 0}_L, 1), b^* = (-b[1], \dots, -b[m], \underbrace{0, \dots, 0}_n).$$

Поэтому в силу (31) — (34) для оптимальности $\bar{x} \in \mathcal{D}$ в задаче $(\mathcal{D}, f)$ (или, что то же, для оптимальности $(\bar{x}, f(\bar{x}))$ в задаче $(\mathcal{D}', f')$) необходимо и достаточно существование такого вектора $u \in R^N$, что

$$\sum_{l=1}^N u[l] a[l, l] \leq 0, l \in 1:L, \quad (37)$$

$$\sum_{j=1}^n u[m+j] k[j] = 1, \quad (38)$$

$$\sum_{l=1}^N u[l] a[l, l] = 0, \text{ если } \bar{x}[l] > 0, \quad (39)$$

$$u[l] \geq 0, l \in 1:N, \quad (40)$$

$$u[l] = 0, \text{ если } i \in 1:m \text{ и } \sum_{l=1}^L a[l, l] \bar{x}[l] > b[l], \quad (41)$$

$$u[m+j] = 0, \text{ если } j \in 1:n \text{ и}$$

$$\frac{1}{k[j]} \sum_{l=1}^L a[m+j, l] \bar{x}[l] > f(\bar{x}). \quad (42)$$

Отметим следующее. Задача $(\mathcal{D}, f)$ является задачей выпуклого программирования. Однако для получения условий оптимальности (37) — (42) нельзя непосредственно использовать критерий п. 1.7, так как кусочно-линейная функция f недифференцируема.

Обратим также внимание на возможность интересной экономической интерпретации условий (37) — (42). С оптимальным планом задачи связывается система оценок ингредиентов технологических способов (система объективно обусловленных оценок, как их назвал Л. В. Канторович). В статической ситуации, когда фиксированы технологические коэффициенты $a[i, l]$ и целевые параметры производства $k[j]$, эти оценки выступают как функции ресурсов и функционально обладают свойствами «цен идеализированной экономики» отдельного предприятия, отрасли, народного хозяйства и т. п. (т. е. в зависимости от того, что аппроксимирует данная линейная модель).

Интересно, что соответствующие оценки-цены связываются как с ингредиентами первой группы (сырье, трудовые и другие ресурсы производства), так и с ингредиентами второй группы — продуктами производства и не только порождаются оптимальным планом, но и могут стимулировать оптимальное планирование. В этом свете условия (37) — (42) получают естественное толкование.

Условия (40) и (38) — соответственно условия неотрицательности и нормировки (фиксированного масштаба) цен. В частности, поскольку все $k[j] > 0$, $j \in 1:n$, из (38) следует, что по крайней мере один выпускаемый продукт имеет положительную оценку-цену.

Условия (37) и (39) переписываем в следующем виде:

$$\sum_{l=m+1}^N u[l] a[i, l] \leq \sum_{l=1}^m u[l] (-a[i, l]), \quad i \in 1:L,$$

$$\sum_{l=m+1}^N u[l] a[i, l] = \sum_{l=1}^m u[l] (-a[i, l]), \quad \text{если } \bar{x}[i] > 0,$$

и ясно, что суммарная оценка продукции, производимой любым технологическим способом, не может превосходить суммарной оценки использованных в производстве ресурсов, причем в «рентабельных» технологических способах имеет место равенство, т. е. выполняется условие (39).

Естественна и интерпретация свойств (41), (42): не полностью используемые ресурсы и продукты, произведенные в количестве, превышающем заданные потребности, обесцениваются (их оценки-цены равны нулю).

Вопросам экономической интерпретации оценок оптимального плана посвящена обширная литература (см., например, [20, 22]), к которой мы и отсылаем читателя, интересующегося качественным анализом соответствующих проблем.

2°. Задача линейного программирования с ограниченными сверх переменными

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \quad \mathcal{D} = \{x \in R^n: Ax = b, 0 \leq x \leq h\}$$

рассматривалась в п. 2.4.10. Напоминаем, что $A — m \times n$ матрица, столбцы которой обозначаем через a^j , $j \in 1:n$. Предпо-

лагается, что $h[j] < +\infty$ для $j \in J' \subset 1:n$ и $h[j] = +\infty$, если $j \notin J'$, $|J'| = n'$.

В силу свойства б) эквивалентных экстремальных задач (см. п. 2.4.3) вектор $\bar{x} \in \mathcal{D}$ является решением задачи $(\mathcal{D}, f)$ тогда и только тогда, когда $\varphi(\bar{x}) = (\bar{x}, y(\bar{x}))$ — оптимальный план ее канонической формы:

$$(\mathcal{D}', f'): f'(x', y) = cx' \rightarrow \max, \mathcal{D}' = \{(x', y) \in R^{n+n'} : x' \in R^n, \\ y \in R^{n'}; Ax' = b; x'[j] + y[j] = h[j], j \in J'; \\ x' \geq 0; y[j] \geq 0, j \in J'\}.$$

Задача $(\mathcal{D}', f')$ является канонической задачей линейного программирования, поэтому в силу (35), (36) для оптимальности $(\bar{x}, \bar{y}) \in \mathcal{D}'$ необходимо и достаточно существование такого $(u, w) \in R^{m+n'}$ ($u \in R^m, w \in R^{n'}$), что

$$c[j] \leq ua^j + w[j], j \in J', \quad (43)$$

$$c[j] \leq ua^j, j \in J', \quad (44)$$

$$0 \leq w[j], j \in J', \quad (45)$$

$$c[j] = ua^j + w[j], \text{ если } \bar{x}[j] > 0, j \in J', \quad (46)$$

$$c[j] = ua^j, \text{ если } \bar{x}[j] > 0, j \in J', \quad (47)$$

$$w[j] = 0, \text{ если } \bar{y}[j] > 0, j \in J'. \quad (48)$$

По условию $\bar{y}[j] = h[j] - \bar{x}[j]$, $j \in J'$. Тогда, как нетрудно видеть, из (43) — (48) получаем: вектор $\bar{x} \in \mathcal{D}$ является решением задачи $(\mathcal{D}, f)$ тогда и только тогда, когда существует такое $u \in R^m$, что

$$c[j] \leq ua^j \text{ для } j \in J' \text{ и } j \in J', \text{ если } \bar{x}[j] = 0, \quad (49)$$

$$c[j] = ua^j, \text{ если } 0 < \bar{x}[j] < h[j], \quad (50)$$

$$c[j] \geq ua^j, \text{ если } \bar{x}[j] = h[j]. \quad (51)$$

3°. Транспортная задача

$$(\mathcal{D}, f): f(x) = \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j] \rightarrow \min,$$

$$\mathcal{D} = \{x \in R^{mn}: \sum_{j=1}^n x[i, j] = a[i], i \in 1:m,$$

$$\sum_{i=1}^m x[i, j] = b[j], j \in 1:n; x \geq 0\}$$

рассматривалась в п. 2.3.2. Если компоненты $m \times n$ матриц $x = (x[i, j])$ и $c = (c[i, j])$ занумеровать в последовательности

$$x = (x[1, 1], x[1, 2], \dots, x[1, n], x[2, 1], \dots, x[2, n], \dots$$

$$\dots, x[m, 1], \dots, x[m, n]),$$

$$c = (c[1, 1], c[1, 2], \dots, c[1, n], c[2, 1], \dots, c[2, n], \dots$$

$$\dots, c[m, 1], \dots, c[m, n]),$$

то задачу $(\mathcal{D}, f)$ можно записать в форме канонической задачи линейного программирования $(\mathcal{D}, -f) : -f(x) = (-c)x \rightarrow \max$, $\mathcal{D} = \{x \in R^{m+n} : Ax = b, x \geq 0\}$, где через A обозначена $(m+n) \times mn$ матрица

$$A = \left[\begin{array}{ccc|ccc|ccc|ccc} 1 & 1 & \dots & 1 & & & & & & & & \\ & & & 1 & 1 & \dots & 1 & & & & & \\ & & & & & & \dots & & & & & \\ & & & & & & & 1 & 1 & \dots & 1 & \\ \hline 1 & & & & 1 & & & & 1 & & & \\ & 1 & & & & 1 & & & & 1 & & \\ & & \cdot & & & & \cdot & & & & \cdot & \\ & & & \cdot & & & & \cdot & & & & \cdot \\ & & & & \cdot & & & & \cdot & & & \\ & & & & & \cdot & & & & \cdot & & \\ & & & & & & 1 & & & & & 1 \end{array} \right],$$

состоящая из нулей и единиц (нули не проставлены), а через b — вектор-столбец из R^{m+n} , составленный из правых частей уравнений, определяющих множество $\mathcal{D}$: $b^T = (a[1], \dots, a[m], b[1], \dots, b[n])$.

Заметим, что столбец $a^{(i,n)}$ матрицы A , отвечающий компоненте $x[i, j]$ вектора x , содержит единицы лишь на i -м и $(m+j)$ -м местах (остальные элементы этого столбца равны нулю).

В силу условий (35), (36) для оптимальности допустимого плана $\bar{x}$ задачи $(\mathcal{D}, f)$ необходимо и достаточно существование такого вектора $u \in R^{m+n}$, что $-c \leq uA$ и $-c[i, j] = ua^{(i,n)}$, если $\bar{x}[i, j] > 0$. Обозначим компоненты вектора u через $u[i]$, $i \in 1:m$, и $-v[j]$, $j \in 1:n$, так что $u = (u[1], \dots, u[m], -v[1], \dots, -v[n])$.

Принимая во внимание структуру столбцов $a^{(i,n)}$ матрицы A , заключаем, что для оптимальности вектора $\bar{x} \in \mathcal{D}$ необходимо и достаточно существование таких чисел $u[i]$, $i \in 1:m$, $v[j]$, $j \in 1:n$, что $-c[i, j] \leq u[i] - v[j]$, $i \in 1:m$, $j \in 1:n$, и $-c[i, j] = u[i] - v[j]$, если $\bar{x}[i, j] > 0$, или план $\bar{x} \in \mathcal{D}$ является решением транспортной задачи тогда и только тогда, когда существуют такие числа $u[1], u[2], \dots, u[m], v[1], v[2], \dots, v[n]$ (обычно их называют потенциалами плана $\bar{x}$), что

$$v[j] - u[i] \leq c[i, j], \quad i \in 1:m, \quad j \in 1:n; \quad (52)$$

$$v[j] - u[i] = c[i, j], \quad \text{если } \bar{x}[i, j] > 0. \quad (53)$$

Условия оптимальности (52), (53) допускают естественную экономическую интерпретацию: с каждым решением $\bar{x}$ транспортной задачи можно связать такую систему цен-оценок единицы продукта в пунктах хранения ($u[i]$) и пунктах потребления ($v[j]$), что цена единицы продукта в пункте j не может быть меньше его цены в пункте i , увеличенной на стоимость

перевозки, и «рентабельными» являются только такие перевозки, при которых цена возрастает лишь на стоимость транспортировки.

4°. *Транспортная задача с ограничениями на пропускные способности* отличается от обычной транспортной задачи тем, что в систему ограничений, задающих множество $\mathcal{D}$, добавляются условия вида $0 \leq x[i, j] \leq h[i, j]$ (допускается $h[i, j] = +\infty$). В силу (49)–(51) для оптимальности допустимого плана $\bar{x}$ транспортной задачи с ограничениями на пропускные способности необходимо и достаточно существование таких потенциалов $u[1], \dots, u[m], v[1], \dots, v[n]$, что

$$v[j] - u[i] \leq c[i, j], \text{ если } h[i, j] = +\infty \text{ или } \bar{x}[i, j] = 0, \quad (54)$$

$$v[j] - u[i] = c[i, j], \text{ если } 0 < \bar{x}[i, j] < h[i, j], \quad (55)$$

$$v[j] - u[i] \geq c[i, j], \text{ если } \bar{x}[i, j] = h[i, j]. \quad (56)$$

Экономическая интерпретация условий (54), (55) аналогична соответствующей интерпретации условий (52), (53). Условия (56) связывают с необходимостью выплаты дополнительных рентных платежей за эксплуатацию предельно нагруженных дорог.

§ 2. Теорема Куна—Таккера

2.1. *Функция Лагранжа задачи выпуклого программирования. Седловая точка.* Пусть $(\mathcal{D}, f)$ — задача выпуклого программирования, например задача максимизации*

$$(\mathcal{D}, f): f(x) \rightarrow \max, \quad \mathcal{D} = \{x \in R^n_+ : f_i(x) \leq 0, \quad (1)$$

$$i \in I_n; f_i(x) = a_i x - b[i] = 0, i \in I_y\},$$

где f — вогнутая функция; $f_i, i \in I_n$, — выпуклые функции. Свяжем с этой задачей числовую функцию Φ , определенную на $R^n \times R^m$:

$$\Phi(x, u) = f(x) - \sum_{i=1}^m u[i] f_i(x). \quad (2)$$

Функция Φ называется *функцией Лагранжа задачи* $(\mathcal{D}, f)$. Очевидно, что Φ — аффинная функция от u при каждом фиксированном x и вогнутая функция от x при каждом фиксированном $u \in R^m_+$.

Пусть числовая функция Ψ определена на $R^n \times R^m$ и $U \subset R^m, X \subset R^n$. Точка $(\bar{x}, \bar{u}) \in X \times U$ называется *седловой точкой*.

* Напоминаем, что через R^n_+ мы обозначили множество таких векторов из R^n , у которых $x[j] \geq 0$ для $j \in J$; аналогично $R^m_+ = \{u \in R^m : u[i] \geq 0, i \in I_n\}$.

кой функции Ψ на множестве $X \times U$, если

$$\Psi(x, \bar{u}) \leq \Psi(\bar{x}, \bar{u}) \quad \text{для любых } x \in X \quad (3)$$

и

$$\Psi(\bar{x}, \bar{u}) \leq \Psi(\bar{x}, u) \quad \text{для любых } u \in U. \quad (4)$$

Приведем два примера. Для функции $\Psi(x, u) = -x^2 + u^2$ точка $(0, 0)$ является седловой на множестве $R \times R$, поскольку для любых $x \in R$ и $u \in R$ выполняются неравенства $-x^2 \leq 0$, $0 \leq u^2$. График функции Ψ в окрестности точки $(0, 0)$ имеет вид седла (рис. 19, а), что и объясняет происхождение соответствующего термина.

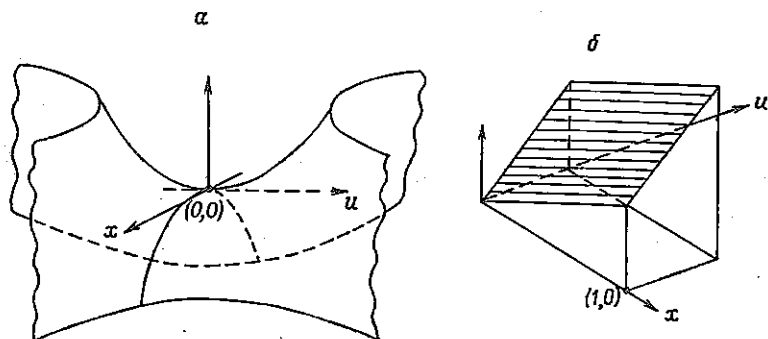


Рис. 19.

Аналогично для функции $\Lambda(x, u) = x + u$ точка $(1, 0)$ — седловая на $X \times U = [0, 1] \times [0, 1]$, так как $x \leq 1$, $1 \leq 1 + u$ для любых $x \in [0, 1]$ и $u \in [0, 1]$. Здесь она является граничной точкой множества $[0, 1] \times [0, 1]$. График функции Λ изображен на рис. 19, б.

Неравенства (3) и (4) называют неравенствами седловой точки.

2.2. Формулировка теоремы Куна—Таккера. Рассматривается задача выпуклого программирования $(\mathcal{D}, f)$. Пусть Φ — ее функция Лагранжа.

Теорема 1. Для оптимальности вектора $\bar{x} \in R^n^+$ достаточно существования такого вектора $\bar{u} \in R_l^{m+}$, что $(\bar{x}, \bar{u})$ — седловая точка функции Φ на $R^n^+ \times R_l^{m+}$.

Теорема 2. Если функции f_i , $i \in I_n$, и f дифференцируемы и множество $\mathcal{D}$ удовлетворяет условию С относительно множества ограничений $I_n \subset I_n$ (см. п. 1.5), то для оптимальности вектора $\bar{x} \in \mathcal{D}$ необходимо, чтобы существовал такой вектор $\bar{u} \in R_l^{m+}$, что $(\bar{x}, \bar{u})$ — седловая точка функции Φ на $R^n^+ \times R_l^{m+}$.

Теорема 3. Если множество $\mathcal{D}$ удовлетворяет условию С относительно системы ограничений I_n и J и векторы a_i , $i \in I_n$,

линейно независимы, то для оптимальности вектора $\bar{x} \in \mathcal{D}$ необходимо, чтобы существовал такой вектор $\bar{u} \in R_{I_n}^{m+}$, что $(\bar{x}, \bar{u})$ — седловая точка функции Φ на $R_I^{n+} \times R_{I_n}^{m+}$.

Следует обратить внимание на то, что теорема 1 — достаточность условий Куна—Таккера — формулируется без каких-либо дополнительных условий на f и f_i , $i \in 1:m$. Более того, как будет видно из доказательства, утверждение теоремы 1 верно и для экстремальных задач типа (1) без каких-либо предположений о свойствах функций f и f_i , $i \in 1:m$. Теорема 3 утверждает необходимость условия Куна—Таккера без предположений дифференцируемости целевой функции и функций f_i , $i \in I_n$. «Платой» за освобождение от этих ограничений является распространение требования о выполнении условия С на все неравенства. Как будет показано ниже (см. п. 3.8), без предположения выполнения условия С для выпуклых ограничений необходимость условий Куна—Таккера может не иметь места.

Отметим, что утверждение теоремы 3 верно и без предположения выполнения условия С для линейных неравенств. Однако его доказательство сложно и использует специальный аппарат [43].

2.3. Доказательство теоремы 1. Пусть $\bar{x} \in R_I^{n+}$, $\bar{u} \in R_{I_n}^{m+}$ и

$$\Phi(x, \bar{u}) \leq \Phi(\bar{x}, \bar{u}), \quad \Phi(\bar{x}, \bar{u}) \leq \Phi(\bar{x}, u) \quad (5)$$

для любых $x \in R_I^{n+}$, $u \in R_{I_n}^{m+}$. Второе неравенство в (5) эквивалентно неравенству

$$-\sum_{i=1}^m \bar{u}[i] f_i(\bar{x}) \leq -\sum_{i=1}^m u[i] f_i(\bar{x}), \quad u \in R_{I_n}^{m+} \quad (6)$$

и может иметь место для любого $u \in R_{I_n}^{m+}$ лишь тогда, когда $f_i(\bar{x}) = 0$, $i \in I_n$, и $f_i(\bar{x}) \leq 0$, $i \in I_n$, так как нарушение какого-либо из этих условий позволяет найти такое $u \in R_{I_n}^{m+}$, для которого (6) не выполняется. Поэтому $\bar{x} \in \mathcal{D}$, и остается проверить его оптимальность.

Из условия (6) следует, что $\sum_{i=1}^m \bar{u}[i] f_i(\bar{x}) = 0$. Действительно, полагая в (6) $u = 0 \in R_{I_n}^{m+}$, имеем $\sum_{i=1}^m \bar{u}[i] f_i(\bar{x}) \geq 0$. Вместе с тем поскольку $\bar{x} \in \mathcal{D}$ и $\bar{u} \in R_{I_n}^{m+}$, то

$$\sum_{i=1}^m \bar{u}[i] f_i(\bar{x}) = \sum_{i \in I_n} \bar{u}[i] f_i(\bar{x}) \leq 0. \quad (7)$$

Поэтому первое неравенство в (5) может быть переписано в виде

$$f(x) - \sum_{i=1}^m \bar{u}[i] f_i(x) \leq f(\bar{x}),$$

что выполняется для любого $x \in R_I^{n+}$. Если $x \in \mathcal{D}$, то $\sum_{i=1}^m \bar{u}[i] f_i(x) \leq 0$ (ср. с (7)), и, следовательно, $f(x) \leq f(\bar{x})$. ■

2.4. Доказательство теоремы 2. Пусть выполняются условия теоремы 2 и $\bar{x}$ — решение задачи выпуклого программирования $(\mathcal{D}, f)$ (см. (1)). Тогда для $\bar{x}$ в силу критерия оптимальности п. 1.7 существуют такие $\bar{u} \in R_{I_n}^{m+}$ и $\bar{v} \in R_J^{n+}$, что

$$\nabla f(\bar{x}) = \sum_{i=1}^m \bar{u}[i] \nabla f_i(\bar{x}) - \sum_{j=1}^n \bar{v}[j] e_j, \quad (8)$$

$$\bar{u}[i] = 0, \text{ если } f_i(\bar{x}) < 0, \quad (9)$$

$$\bar{v}[j] = 0, \text{ если } j \in J \text{ или } j \in J \text{ и } \bar{x}[j] > 0. \quad (10)$$

Покажем, что $(\bar{x}, \bar{u})$ удовлетворяет неравенствам (5). Так как $\bar{x} \in \mathcal{D}$, то из (9) следует

$$\sum_{i=1}^m \bar{u}[i] f_i(\bar{x}) = 0, \quad (11)$$

и для доказательства второго неравенства в (5) остается заметить, что оно сводится к неравенству $0 \leq -\sum_{i \in I_n} u[i] f_i(\bar{x})$, очевидно справедливому для любого $u \in R_{I_n}^{m+}$ в силу $f_i(\bar{x}) \leq 0$ для $i \in I_n$.

Далее,

$$\begin{aligned} \Phi(x, \bar{u}) - \Phi(\bar{x}, \bar{u}) &\leq \nabla_x \Phi(\bar{x}, \bar{u})(x - \bar{x}) = \\ &= (\nabla f(\bar{x}) - \sum_{i=1}^m \bar{u}[i] \nabla f_i(\bar{x}))(x - \bar{x}), \end{aligned}$$

поскольку $\Phi(x, \bar{u})$ — вогнутая функция от x при любом фиксированном $\bar{u} \in R_{I_n}^{m+}$ (см. п. 2.1), и в силу (8) и (10)

$$\begin{aligned} \Phi(x, \bar{u}) - \Phi(\bar{x}, \bar{u}) &\leq -\sum_{j=1}^n \bar{v}[j] e_j(x - \bar{x}) = \\ &= -\sum_{j \in J(\bar{x})} \bar{v}[j] x[j] \leq 0, \text{ если } x \in R_J^{n+}. \blacksquare \end{aligned}$$

2.5. Необходимость условий Куна—Таккера (доказательство теоремы 3). Пусть выполняются условия теоремы 3 и $\bar{x}$ — решение задачи $(\mathcal{D}, f)$. Для удобства доказательство теоремы разобьем на части.

1. Построение вектора $\bar{u}$. Рассмотрим в R^{m+n+1} множества A и B : вектор $a = (a[0], a[1], \dots, a[m+n]) \in R^{m+n+1}$ отнесем к множеству A , если существует такой $x \in R^n$, что

$$a[0] \leq f(x), \quad (12A)$$

$$a[i] \leq -f_i(x), \quad i \in I_n, \quad (13A)$$

$$a[i] = -f_i(x), \quad i \in I_y, \quad (14A)$$

$$a[m+j] \leq x[j], \quad j \in J; \quad (15A)$$

вектор $b = (b[0], b[1], \dots, b[m+n]) \in R^{m+n+1}$ отнесем к множеству B , если

$$b[0] > f(\bar{x}), \quad (12B)$$

$$b[i] \geq 0, \quad i \in I_n, \quad (13B)$$

$$b[i] = 0, \quad i \in I_y, \quad (14B)$$

$$b[m+j] \geq 0, \quad j \in J. \quad (15B)$$

Покажем, что A и B — непустые непересекающиеся выпуклые множества в R^{m+n+1} .

Действительно, множества A и B непустые, поскольку при любом $x \in R^n$ вектор

$$a(x) = (f(x), -f_1(x), \dots, -f_m(x), x[1], \dots, x[n]) \quad (16)$$

удовлетворяет неравенствам (12A) — (15A), а вектор $(f(\bar{x}) + 1, 0, 0, \dots, 0) \in R^{m+n+1}$ очевидно принадлежит B . Они непересекающиеся, так как если для $a \in A$, $x \in R^n$ выполняются соотношения (12A) — (15A), то для вектора a при $x \in \mathcal{D}$ неверно (12B), а при $x \notin \mathcal{D}$ не выполняется по крайней мере одно из неравенств (13B) — (15B), и, следовательно, $a \notin B$. Множество B выпукло, так как является пересечением полупространств и гиперплоскостей. Если $a_1, a_2 \in A$ и x_1, x_2 — соответствующие векторы из R^n , так что для a_1, x_1 и a_2, x_2 выполняются соотношения (12A) — (15A), то для $a = (1 - \lambda)a_1 + \lambda a_2$, $\lambda \in [0, 1]$, выполняются соотношения (12A) — (15A) при $x = (1 - \lambda)x_1 + \lambda x_2$. Это следует из вогнутости f и выпуклости $f_i, i \in I_n$.

К множествам A и B может быть применена теорема отделимости (см. п. 1.4.3): существуют такой ненулевой вектор $z = (z[0], z[1], \dots, z[m+n]) \in R^{m+n+1}$ и число α , что $za \leq \alpha$ для любого a из A и $\alpha \leq zb$ для любого b из B . Возьмем такие z и α . Очевидно, что неравенство $\alpha \leq zb$ выполняется и для любого b из замыкания $\bar{B}$ множества B . Таким образом, для любых a из A и b из $\bar{B}$ справедливо неравенство

$$za \leq zb. \quad (17)$$

Ниже будет показано, что $z[0] > 0$, $z[i] \geq 0, i \in I_n$. Поэтому можно определить $\bar{u} = (z[1]/z[0], z[2]/z[0], \dots, z[m]/z[0])$ и $\bar{u} \in R_n^{m+}$.

2. Свойства коэффициентов разделяющей гиперплоскости. Покажем сначала, что

$$\begin{aligned} z[0] &\geq 0, \quad z[i] \geq 0, \quad i \in I_n; \quad z[m+j] \geq 0, \quad j \in J, \\ z[m+j] &= 0, \quad j \in 1:n \setminus J. \end{aligned} \quad (18)$$

Для этого возьмем в $\bar{B}$ точку

$$\bar{b} = (f(\bar{x}), 0, \dots, 0) \quad (19)$$

и $x \in R^n$. Тогда из (17) следует, что неравенство

$$z[0] a[0] + \sum_{i=1}^{m+n} z[i] a[i] \leq z[0] f(\bar{x})$$

должно выполняться для любых чисел $a[i]$, $i \in 0:m+n$, удовлетворяющих соотношениям (12A)–(15A). Но это возможно лишь в том случае, если выполняются условия (18).

Действительно, предположим, что какое-нибудь из этих условий нарушается, например, $z[i_0] < 0$ для некоторого $i_0 \in I_n$, и пусть $\bar{a}[0], \bar{a}[1], \dots, \bar{a}[m+n]$ — последовательность чисел, удовлетворяющая условиям (12A)–(15A). Тогда при любом $\lambda > 0$ условия (12A)–(15A) выполняются и для последовательности чисел $\bar{a}[0], \bar{a}[1], \dots, \bar{a}[i_0] - \lambda, \dots, \bar{a}[m+n]$. Следовательно, при любом $\lambda > 0$ справедливо неравенство

$$-z[i_0] \cdot \lambda \leq z[0] f(\bar{x}) - z[0] \bar{a}[0] - \sum_{i=1}^{m+n} z[i] \bar{a}[i] = \text{const},$$

что невозможно при $z[i_0] < 0$. Аналогично устанавливается необходимость выполнения всех остальных условий (18).

Покажем теперь, что $z[0] > 0$. Выше было доказано неравенство $z[0] \geq 0$. Поэтому предположение $z[0] \leq 0$ влечет за собой равенство $z[0] = 0$. Установим, что из условия $z[0] = 0$ следует невозможное равенство $z = 0$.

Итак, пусть $z[0] = 0$. Так как множество $\mathcal{D}$ удовлетворяет условию С относительно системы ограничений I_n и J , то существует такой вектор $\hat{x} \in \mathcal{D}$, что

$$f_i(\hat{x}) < 0, \quad i \in I_n; \quad \hat{x}[j] > 0, \quad j \in J. \quad (20)$$

Подставим в (17) вместо a и b векторы $a(\hat{x}) \in A$ (см. (16) при $x = \hat{x}$) и $\bar{b} \in \bar{B}$ (см. (19)). Так как $z[0] = 0$, то

$$-\sum_{i \in I_n} z[i] f_i(\hat{x}) + \sum_{j \in J} z[m+j] \hat{x}[j] \leq 0.$$

Но это неравенство в силу (18) и (20) может иметь место лишь в случае, если $z[i] = 0$, $i \in I_n$, и $z[m+j] = 0$, $j \in J$. Ранее мы уже установили, что $z[m+j] = 0$, $j \in 1:n \setminus J$.

Возьмем теперь произвольный x из R^n и подставим в (17) вместо a вектор $a(x)$ (см. (16)), а вместо b — вектор $\bar{b}$. Получим

$$-\sum_{i \in I_y} z[i] f_i(x) = -\left(\sum_{i \in I_y} z[i] a_i\right)x + \sum_{i \in I_y} z[i] b[i] \leq 0,$$

поскольку $f_i(x) = a_i x - b[i]$ для $i \in I_y$. Последнее неравенство справедливо для любого x из R^n , что возможно лишь при условии $\sum_{i \in I_y} z[i] a_i = 0$, а значит, и $z[i] = 0$, $i \in I_y$, в силу предположенной линейной независимости векторов a_i , $i \in I_y$.

Итак, $z[i] = 0$, $i \in 0:m+n$, что невозможно. Следовательно, $z[0] > 0$ и $\bar{u} \in R_{n+1}^{m+}$.

Покажем, что

$$z[i] = 0 \text{ при } i \in I_n \setminus I(\bar{x}), \quad (21)$$

т. е. при $i \in I_n$ и $f_i(\bar{x}) < 0$. Для этого возьмем $\bar{b} \in \bar{B}$ (см. (19)) и $a(x)$ (см. (16) при $x = \bar{x}$). Подставляя $a(\bar{x})$ и $\bar{b}$ в (17), имеем

$$-\sum_{i \in I_n} z[i] f_i(\bar{x}) + \sum_{j \in J} z[m+j] \bar{x}[j] \leq 0,$$

откуда с учетом (18) и допустимости $\bar{x}$ следует (21).

3. Неравенства седловой точки. Покажем, что $(\bar{x}, \bar{u})$ — седловая точка функции Φ на $R_I^{n+} \times R_I^{m+}$. Из (21) и определения $\bar{u}$ следует, что

$$\sum_{i=1}^m \bar{u}[i] f_i(\bar{x}) = 0, \quad (22)$$

и второе неравенство в (5) очевидно, поскольку $\bar{x} \in \mathcal{D}$, а $u \in R_I^{m+}$. Пусть $x \in R_I^{n+}$ и вектор $a(x)$ определяется формулой (16). Тогда в силу (18) (напомним, что $z[0] > 0$)

$$\begin{aligned} \Phi(x, \bar{u}) &= f(x) - \sum_{i=1}^m \bar{u}[i] f_i(x) = \\ &= \frac{1}{z[0]} \left(z[0] f(x) - \sum_{i=1}^m z[i] f_i(x) \right) \leq \\ &\leq \frac{1}{z[0]} \left(z[0] f(x) - \sum_{i=1}^m z[i] f_i(x) + \sum_{j \in J} z[m+j] x[j] \right) = \\ &= \frac{1}{z[0]} z a(x), \end{aligned}$$

и так как $a(x) \in A$, $\bar{b} \in \bar{B}$, то из (17) и (22) получаем

$$\Phi(x, \bar{u}) \leq \frac{1}{z[0]} z a(x) \leq \frac{1}{z[0]} z \bar{b} = f(\bar{x}) = \Phi(\bar{x}, \bar{u}). \blacksquare$$

§ 3. Двойственность в математическом программировании

Как уже отмечалось, критерии оптимальности предыдущих параграфов сводят проверку оптимальности, а при выполнении определенных условий и неоптимальности допустимого плана задачи $(\mathcal{D}, f)$ к проверке совместности некоторой системы линейных уравнений и неравенств. В настоящем параграфе мы продолжаем исследования в этой области. Для задачи выпуклого программирования будет построена экстремальная задача, которую естественно назвать двойственной к исходной задаче. В линейном случае эта конструкция становится особенно интересной.

3.1. Сопряженные задачи. Пусть $\bar{R}$ — расширенная прямая, т. е. множество вещественных чисел, пополненное точками $-\infty$ и $+\infty$ (см. п. 1.8.1). Пусть на $R^n \times R^m$ определена функция G , принимающая конечные или бесконечные значения: $G: R^n \times R^m \rightarrow \bar{R}$, и пусть $X \subset R^n$, $U \subset R^m$. Положим

$$f(x) = \inf_{u \in U} G(x, u), \quad (1)$$

$$g(u) = \sup_{x \in X} G(x, u). \quad (2)$$

Функции f и g определены соответственно на R^n и R^m и принимают значения из $\bar{R}$. Рассмотрим экстремальные задачи

$$(X, f): f(x) \rightarrow \max, \quad X \subset R^n, \quad (3)$$

$$(U, g): g(u) \rightarrow \min, \quad U \subset R^m. \quad (4)$$

Задачи (3) и (4) будем называть *сопряженными относительно функции G* . Точнее, будем говорить, что задача (U, g) сопряжена с задачей (X, f) относительно функции G .

Ниже будут установлены некоторые свойства сопряженных задач.

1°. Задача минимизации $(X, -f)$ сопряжена с задачей максимизации $(U, -g)$ относительно функции $\bar{G}(u, x) = -G(x, u)$.

Действительно,

$$\inf_{x \in X} \bar{G}(u, x) = \inf_{x \in X} (-G(x, u)) = - \sup_{x \in X} G(x, u) = -g(u),$$

$$\sup_{u \in U} \bar{G}(u, x) = \sup_{u \in U} (-G(x, u)) = - \inf_{u \in U} G(x, u) = -f(x). \blacksquare$$

Следующее очевидное утверждение иногда называют первой теоремой двойственности.

2°. Если задача минимизации (U, g) сопряжена с задачей максимизации (X, f) относительно функции G , то

$$f(x) \leq g(u) \quad \text{для любых } x \in X, u \in U. \quad (5)$$

Неравенство (5) сразу следует из формул (1) и (2):

$$f(x) \leq G(x, u) \leq g(u) \quad \text{для любых } x \in X, u \in U. \blacksquare$$

Предположим, что задачи (3) и (4) имеют решения: существуют такие $\bar{x} \in X$ и $\bar{u} \in U$, что $f(x) \leq f(\bar{x})$, $g(u) \geq g(\bar{u})$ для любых $x \in X$, $u \in U$. Тогда из 2° следует, что

$$f(\bar{x}) \leq g(\bar{u}). \quad (6)$$

3°. Если $\bar{x} \in X$, $\bar{u} \in U$ и

$$f(\bar{x}) = g(\bar{u}), \quad (7)$$

то $\bar{x}$, $\bar{u}$ — решения сопряженных задач (3), (4).

Действительно, для любых $x \in X$ и $u \in U$ в силу 2° и (7) имеем $f(x) \leq g(\bar{u}) = f(\bar{x})$, $g(u) \geq f(\bar{x}) = g(\bar{u})$. $\blacksquare$

Обратное утверждение, вообще говоря, неверно: $\bar{x}$ и $\bar{u}$ могут быть решениями задач (3) и (4), но $f(\bar{x}) < g(\bar{u})$.

Рассмотрим на $R \times R$ функцию $G(x, u) = \sin(x + u)$, и пусть $X = U = R$. Тогда

$$f(x) = \inf_{u \in R} \sin(x + u) = -1 \quad \text{при любом } x \in R,$$

$$g(u) = \sup_{x \in R} \sin(x + u) = +1 \quad \text{при любом } u \in R.$$

Очевидно, что любая точка $\bar{x} \in R$ является решением задачи максимизации (R, f) и любая точка $\bar{u} \in R$ есть решение задачи минимизации (R, g) . Но $f(\bar{x}) = -1 < g(\bar{x}) = +1$.

4°. Выполнение (7) эквивалентно неравенству седловой точки, а именно если $\bar{x} \in X$, $\bar{u} \in U$, то равенство (7) имеет место тогда и только тогда, когда для любых $x \in X$, $u \in U$

$$G(x, \bar{u}) \leq G(\bar{x}, \bar{u}) \leq G(\bar{x}, u), \quad (8)$$

причем $G(\bar{x}, \bar{u}) = f(\bar{x}) = g(\bar{u}). \quad (9)$

Докажем это. Если имеет место (7), то, как следует из (1) и (2), для любых $x \in X$, $u \in U$

$$G(x, \bar{u}) \leq g(\bar{u}) = f(\bar{x}) \leq G(\bar{x}, \bar{u}),$$

$$G(\bar{x}, \bar{u}) \leq g(\bar{u}) = f(\bar{x}) \leq G(\bar{x}, u),$$

так что выполняются соотношения (8) и (9). Обратно, пусть при любых $x \in X$, $u \in U$ выполняются неравенства (8). Поскольку $G(x, \bar{u}) \leq G(\bar{x}, \bar{u})$ при любом $x \in X$, то и $g(\bar{u}) = \sup_{x \in X} G(x, \bar{u}) \leq G(\bar{x}, \bar{u})$. Аналогично, $f(\bar{x}) = \inf_{u \in U} G(\bar{x}, u) \geq G(\bar{x}, \bar{u})$. Поэтому $g(\bar{u}) \leq f(\bar{x})$. Обратное неравенство было установлено в 2°. ■

3.2. Двойственные задачи выпуклого программирования. Пусть дана задача выпуклого программирования (см. начало настоящей главы)

$$(\mathcal{D}, f) : f(x) \rightarrow \max, \quad \mathcal{D} = \{x \in R^n : f_i(x) \leq 0, i \in I_n\};$$

$$f_i(x) = a_i x + b[i] = 0, i \in I_n\}.$$

Рассмотрим $\Phi(x, u) = f(x) - \sum_{i=1}^m u[i] f_i(x)$ — функцию Лагранжа этой задачи. Построим экстремальные задачи $(R^n, \bar{f})$, $(R_n^{m+}, \bar{g})$, сопряженные относительно функции Φ :

$$(R^n, \bar{f}) : \bar{f}(x) = \inf_{u \in R_n^{m+}} \Phi(x, u) \rightarrow \max, \quad R_n^{m+} = \{x \in R^n : x[j] \geq 0, j \in I_n\}, \quad (10)$$

$$(R_n^{m+}, \bar{g}) : \bar{g}(u) = \sup_{x \in R_n^{m+}} \Phi(x, u) \rightarrow \min, \quad R_n^{m+} = \{u \in R^m : u[i] \geq 0, i \in I_n\}. \quad (11)$$

Задача $(\mathcal{D}, f)$ является «сужением» задачи $(R_I^{n+}, \bar{f})$ на множество тех $x \in R_I^{n+}$, для которых $\bar{f}(x) > -\infty$: покажем, что $\bar{f}(x) = f(x)$ для $x \in \mathcal{D}$ и $\bar{f}(x) = -\infty$, если $x \notin \mathcal{D}$ (по определению будем считать, что $\sup_{x \in \mathcal{D}} f(x) = -\infty$, если $\mathcal{D} = \emptyset$).

Действительно, если $x \in \mathcal{D}$, то $f_i(x) = 0$ для $i \in I_y$, $f_i(x) \leq 0$ для $i \in I_n$ и при любом $u \in R_{I_n}^{m+}$ ($u[i] \geq 0$, $i \in I_n$)

$$\Phi(x, u) = f(x) - \sum_{i \in I_n} u[i] f_i(x) \geq f(x),$$

а $\Phi(x, 0) = f(x)$. Поэтому, если $x \in \mathcal{D}$, то

$$\bar{f}(x) = \inf_{u \in R_{I_n}^{m+}} \Phi(x, u) = \Phi(x, 0) = f(x).$$

Равенство $\bar{f}(x) = -\infty$ при $x \notin \mathcal{D}$ очевидно. ■

Так как любое число из R больше $-\infty$, то при $\mathcal{D} \neq \emptyset$ точка $\bar{x} \in R_I^{n+}$ может быть решением задачи (R_I^{n+}, f) только в том случае, если она является решением задачи $(\mathcal{D}, f)$.

Положим $\mathcal{D}^* = \{u \in R_{I_n}^{m+} : \bar{g}(u) < +\infty\}$, и пусть $f^*(u) = \bar{g}(u)$ для $u \in \mathcal{D}^*$. Задача минимизации $(\mathcal{D}^*, f^*)$ называется двойственной к задаче $(\mathcal{D}, f)$. Таким образом, двойственной к экстремальной задаче максимизации $(\mathcal{D}, f)$ выпуклого программирования называется экстремальная задача минимизации $(\mathcal{D}^*, f^*)$, получаемая сужением задачи $(R_{I_n}^{m+}, \bar{g})$, сопряженной к задаче $(R_I^{n+}, \bar{f})$ относительно функции Лагранжа Φ , на множество $\mathcal{D}^*$ тех $u \in R_{I_n}^{m+}$, для которых $\bar{g}(u) < +\infty$.

Заметим, что мы, вообще говоря, не фиксируем определение функции f^* на $R^m \setminus \mathcal{D}^*$, поскольку все такие экстремальные задачи $(\mathcal{D}^*, f^*)$ тривиально эквивалентны. Обычно, если f^* определена на $\mathcal{D}^*$ некоторым аналитическим выражением, считаем, что это же аналитическое выражение задает f^* в тех точках $R^m \setminus \mathcal{D}^*$, в которых оно имеет смысл.

Покажем, что задача $(\mathcal{D}^*, f^*)$, двойственная к задаче выпуклого программирования $(\mathcal{D}, f)$, также является задачей выпуклого программирования.

Действительно, при каждом фиксированном $x \in R_I^{n+}$ функция $\Phi_x(u) = \Phi(x, u)$ есть аффинная функция от u на R^m . Поэтому в силу утверждения п. 1.8.4 функция $\bar{g}(u)$, являющаяся поточечным супремумом семейства аффинных функций, есть выпуклая функция на R^m . Следовательно, $\mathcal{D}^* = R_{I_n}^{m+} \cap \{u \in R^m : \bar{g}(u) < +\infty\}$ — выпуклое множество в R^m и $(\mathcal{D}^*, f^*)$ есть задача минимизации выпуклой функции на выпуклом множестве в R^m . ■

3.3. Примеры двойственных задач. 1. Построим задачу $(\mathcal{D}^*, f^*)$, двойственную к задаче выпуклого программирования $(\mathcal{D}, f): f(x) = 5x_1 + 3x_2 \rightarrow \max, \mathcal{D} = \{x \in R^2: x_1^2 + x_2^2 \leq 1, x_1 \geq 0\}$.

Имеем $\Phi(x, u) = 5x_1 + 3x_2 - u(x_1^2 + x_2^2 - 1)$, $R_f^{2+} = \{x \in R^2: x_1 \geq 0\}$, $R_{f_u}^{1+} = \{u \in R: u \geq 0\}$. Очевидно, что при $u \leq 0$

$$\bar{g}(u) = \sup_{x \in R_f^{2+}} \Phi(x, u) = +\infty.$$

Для нахождения наибольшего значения функции $\Phi(x, u)$ на R_f^{2+} при $u > 0$ воспользуемся стандартными методами нахождения экстремума функции нескольких переменных. Имеем $\partial\Phi(x, u)/\partial x_1 = 5 - 2x_1u$, $\partial\Phi(x, u)/\partial x_2 = 3 - 2x_2u$, и так как при $u > 0$ решение $\bar{x} = (5/(2u), 3/(2u))$ системы уравнений $2ux_1 = 5$, $2ux_2 = 3$ принадлежит R_f^{2+} , то при $u > 0$

$$\begin{aligned} \bar{g}(u) &= \max_{x \in R_f^{2+}} \Phi(x, u) = \Phi(\bar{x}, u) = 5 \cdot \frac{5}{2u} + 3 \cdot \frac{3}{2u} - \\ &\quad - u \left(\frac{25}{4u^2} + \frac{9}{4u^2} - 1 \right) = \frac{17 + 2u^2}{2u} \end{aligned}$$

(достаточные условия максимума выполняются в силу вогнутости $\Phi(x, u)$ по x при $u > 0$). Поэтому $(\mathcal{D}^*, f^*): f^*(u) = (17 + 2u^2)/(2u) \rightarrow \min, \mathcal{D}^* = \{u \in R^+: u > 0\}$.

Отметим, что в этом примере множество $\mathcal{D}^*$ допустимых планов задачи, двойственной к задаче выпуклого программирования $(\mathcal{D}, f)$, незамкнуто.

2. Для задачи выпуклого программирования

$$\begin{aligned} (\mathcal{D}, f): f(x) &= -\sum_{i=1}^n x_i^2 \rightarrow \max, \mathcal{D} = \{x \in R^n: \sum_{i=1}^n x_i^2 \leq 1\}, \\ \Phi(x, u) &= -\sum_{i=1}^n x_i^2 - u(\sum_{i=1}^n x_i^2 - 1) = -(1+u)\sum_{i=1}^n x_i^2 + u. \end{aligned}$$

Очевидно

$$\bar{g}(u) = \sup_{x \in R^n} \Phi(x, u) = \begin{cases} u, & \text{если } u \geq -1, \\ +\infty, & \text{если } u < -1, \end{cases}$$

так что $(\mathcal{D}^*, f^*): f^*(u) = u \rightarrow \min, \mathcal{D}^* = R^+$, и $(\mathcal{D}^*, f^*)$ — задача линейного программирования.

3. Следующие примеры иллюстрируют важные свойства двойственных задач линейного программирования: условия существования у них решений, допустимых планов и т. д.

а) Пусть $\mathcal{D} = \{x \in R^2: x_1 + 2x_2 \leq 4, x_2 \geq 0\}$, $f(x) = 3x_1 + x_2$ и $(\mathcal{D}, f)$ — задача максимизации. Тогда

$$\Phi(x, u) = 3x_1 + x_2 - u(x_1 + 2x_2 - 4) = 4u + (3 - u)x_1 + (1 - 2u)x_2,$$

$$\bar{g}(u) = \sup_{x_2 \geq 0} \Phi(x, u) = \begin{cases} 4u, & \text{если } 3 - u = 0, 1 - 2u \leq 0, \\ +\infty & \text{в противном случае.} \end{cases}$$

Поэтому $(\mathcal{D}^*, f^*) : f^*(u) = 4u \rightarrow \min$, $\mathcal{D}^* = \{u \in R^+ : 3 = u, 1 \leq 2u\} = \{3\}$, $\bar{u} = 3$ — решение задачи $(\mathcal{D}^*, f^*)$, $f^*(3) = 12$. Непосредственно проверяется (см. п. 1.9), что $\bar{x} = (4, 0)$ — оптимальный план задачи $(\mathcal{D}, f)$ и $f(\bar{x}) = 12$, т. е. для линейных задач $(\mathcal{D}, f)$ и $(\mathcal{D}^*, f^*)$ выполняется равенство $f(\bar{x}) = f^*(\bar{u})$.

б) Аналогично для задачи линейного программирования $(\mathcal{D}, f) : f(x) = -3x_1 + 2x_2 \rightarrow \max$, $\mathcal{D} = \{x \in R^{2+} : x_1 + x_2 \leq -2\}$ двойственной является задача

$$(\mathcal{D}^*, f^*) : f^*(u) = -2u \rightarrow \min, \quad \mathcal{D}^* = \{u \in R^+ : u \geq -3, u \geq 2\}.$$

Очевидно, что $\mathcal{D} = \emptyset$, $\mathcal{D}^* \neq \emptyset$ и функция f^* не ограничена снизу на множестве $\mathcal{D}^*$.

в) Двойственной к задаче

$$(\mathcal{D}, f) : f(x) = x_2 \rightarrow \max, \quad \mathcal{D} = \{x \in R^{2+} : -x_1 + x_2 = 2\}$$

является задача

$$(\mathcal{D}^*, f^*) : f^*(u) = 2u \rightarrow \min, \quad \mathcal{D}^* = \{u \in R : 0 \leq -u, 1 \leq u\}.$$

В отличие от примера б) здесь $\mathcal{D}^* = \emptyset$, $\mathcal{D} \neq \emptyset$ и функция f не ограничена сверху на $\mathcal{D}$, так как вектор $x(\lambda) = (\lambda, 2 + \lambda) \in \mathcal{D}$ при любом $\lambda \in R^+$.

г) Наконец, в задаче $(\mathcal{D}, f) : f(x) = x_1 + x_2 \rightarrow \max$,

$$\mathcal{D} = \{x \in R^3 : x_1 - x_2 - x_3 \leq -3, x_2 + x_3 \leq 1, x_1 \geq 0\}$$

и двойственной к ней $(\mathcal{D}^*, f^*) : f^*(u) = -3u_1 + u_2 \rightarrow \min$,

$$\mathcal{D}^* = \{u \in R^{2+} : 1 \leq u_1, 1 = -u_1 + u_2, 0 = -u_1 + u_2\}$$

множества $\mathcal{D}$ и $\mathcal{D}^*$ пусты.

3.4. Двойственные задачи линейного программирования. Рассмотрим задачу линейного программирования

$$(\mathcal{D}, f) : f(x) = cx \rightarrow \max, \quad \mathcal{D} = \{x \in R_I^{n+} : a_i x \leq b[i], \\ i \in I_B; a_i x = b[i], i \in I_T\}. \quad (12)$$

Будем обозначать через $A = (a[i, j])$ матрицу, строками которой являются векторы a_i , $i \in 1:m$, столбцами — векторы a^j , $j \in 1:n$. Матрицу A называют матрицей задачи $(\mathcal{D}, f)$, а вектор b с компонентами $b[i]$, $i \in 1:m$, — вектором ограничений.

Функция Лагранжа задачи (12) имеет вид

$$\begin{aligned}\Phi(x, u) &= cx - \sum_{i=1}^m u[i](a_i x - b[i]) \\ &= cx - u(Ax - b) = ub + (c - uA)x.\end{aligned}$$

Рассмотрим экстремальные задачи $(R_J^{n+}, \bar{f})$ и $(R_n^{m+}, \bar{g})$, сопряженные относительно функции Φ (см. (10) и (11)), и построим задачу $(\mathcal{D}^*, f^*)$, двойственную к задаче $(\mathcal{D}, f)$. Имеем

$$\begin{aligned}\bar{g}(u) &= \sup_{x \in R_J^{n+}} \Phi(x, u) = \sup_{x \in R_J^{n+}} (ub + (c - uA)x) = \\ &= ub + \sup_{x \in R_J^{n+}} \sum_{j=1}^n (c[j] - ua^j)x[j].\end{aligned}\quad (13)$$

Очевидно, что $\bar{g}(u) = +\infty$, если $c[j] \neq ua^j$ при каком-либо $j \in J$ или $c[j] > ua^j$ при некотором $j \in J$. Если $c[j] \leq ua^j$, $j \in J$, и $c[j] = ua^j$, $j \in J$, то (см. (13))

$$\bar{g}(u) = \Phi(0, u) = ub.$$

Поэтому

$$(\mathcal{D}^*, f^*) : f^*(u) = ub \rightarrow \min, \quad (14)$$

$$\mathcal{D}^* = \{u \in R_n^{m+} : c[j] \leq ua^j, j \in J; c[j] = ua^j, j \in J\} \quad (15)$$

— задача, двойственная к задаче $(\mathcal{D}, f)$.

Рекомендуем сравнить условия допустимости вектора u в двойственной задаче с условиями оптимальности (31)–(34) п. 1.9.

Запишем теперь задачу $(\mathcal{D}^*, f^*)$ как задачу максимизации:

$$(\mathcal{D}^*, -f^*) : -f^*(u) = -ub \rightarrow \max,$$

$$\mathcal{D}^* = \{u \in R_n^{m+} : c[j] - ua^j \leq 0, j \in J; c[j] - ua^j = 0, j \in J\}.$$

Ее функция Лагранжа Φ^* определена на $R^m \times R^n$ и

$$\Phi^*(u, x) = -ub - (c - uA)x = -cx + u(Ax - b) = -\Phi(x, u).$$

Следовательно (см. 1° п. 3.1), задача минимизации $(R_J^{n+}, -\bar{f})$ сопряжена к задаче максимизации $(R_n^{m+}, -\bar{g})$ относительно функции Φ^* . В п. 3.2 было установлено, что задача $(\mathcal{D}, f)$ является сужением задачи максимизации $(R_J^{n+}, \bar{f})$ на множество тех $x \in R_J^{n+}$, для которых $\bar{f}(x) > -\infty$. Таким образом, мы доказали, что задача минимизации $(\mathcal{D}, -f)$, или, что то же, задача максимизации $(\mathcal{D}, f)$, является двойственной к задаче $(\mathcal{D}^*, f^*)$, т. е. $(\mathcal{D}^{**}, f^{**}) = (\mathcal{D}, f)$. Другими словами, отношение двойственности симметрично в классе задач линейного программирования: $(\mathcal{D}, f)$ и $(\mathcal{D}^*, f^*)$ образуют пару двойственных задач.

Между формами записи задач двойственной пары наблюдается интересная зависимость. Так, каждому ограничению в

форме неравенства в одной из двойственных задач соответствует условие неотрицательности связанной с этим неравенством двойственной переменной. Если ограничение записано в форме уравнения, то на переменную двойственной задачи с номером этого ограничения условие неотрицательности не накладывается. Аналогично переменным задачи, на которые наложено условие неотрицательности, соответствуют ограничения двойственной задачи, записанные в форме неравенств, а переменным, на которые условие неотрицательности не наложено, — ограничения двойственной задачи в форме уравнений.

При построении двойственной задачи полезно использовать схему (16), в которой предполагается, что $I_1 = k + 1 : m$, $J = 1 : r$, $I_2 = 1 : k$. Система ограничений, задающая множество $\mathcal{D}$, и целевая функция f задачи $(\mathcal{D}, f)$ получаются в соответствии с этой схемой умножением вектора x на строки матрицы A и вектор c . Умножение вектора u на столбцы матрицы A и вектор b позволяет сформулировать условия двойственной задачи $(\mathcal{D}^*, f^*)$. Очевидно, что предположения специальной упорядоченности и формы ограничений задачи и упорядоченности координат вектора x не умаляют общности, так как любую задачу линейного программирования можно привести к виду, позволяющему использовать схему (16).

$$\begin{array}{c}
 \begin{array}{c} 0 \\ \wedge \end{array} \quad \begin{array}{c} 0 \\ \wedge \end{array} \\
 \begin{array}{|c|c|c|c|c|c|} \hline x[1] & \dots & x[r] & x[r+1] & \dots & x[n] \\ \hline \end{array}
 \end{array}$$

$$\begin{array}{c}
 0 \leq u[1] \\
 \vdots \\
 0 \leq u[k] \\
 u[k+1] \\
 \vdots \\
 u[m]
 \end{array}
 \begin{array}{|c|c|c|c|c|c|} \hline a[1,1] & \dots & a[1,r] & a[1,r+1] & \dots & a[1,n] \\ \hline \vdots & & \vdots & \vdots & & \vdots \\ \hline a[k,1] & \dots & a[k,r] & a[k,r+1] & \dots & a[k,n] \\ \hline a[k+1,1] & \dots & a[k+1,r] & a[k+1,r+1] & \dots & a[k+1,n] \\ \hline \vdots & & \vdots & \vdots & & \vdots \\ \hline a[m,1] & \dots & a[m,r] & a[m,r+1] & \dots & a[m,n] \\ \hline \end{array}
 \begin{array}{c}
 \leq b[1] \\
 \vdots \\
 \leq b[k] \\
 = b[k+1] \\
 \vdots \\
 = b[m]
 \end{array}
 \quad (16)$$

$$\begin{array}{c}
 \vee \quad \vee \\
 \begin{array}{|c|c|c|c|c|c|} \hline c[1] & \dots & c[r] & c[r+1] & \dots & c[n] \\ \hline \end{array}
 \end{array}
 \begin{array}{c}
 \min \\
 \rightarrow \max
 \end{array}$$

Предлагаем читателю проверить, пользуясь схемой (16), что задача, двойственная к задаче

$$(\mathcal{D}, f): f(x) = 2x_1 - 3x_2 + 4x_4 - x_5 \rightarrow \max,$$

$$\begin{aligned} \mathcal{D} = \{x \in R^5: & 3x_1 + 2x_2 - x_3 + x_4 \leq 2, \\ & x_2 + x_3 - 6x_4 + x_5 = 4, \\ & x_1 + 3x_2 + 2x_3 + x_4 + 2x_5 \leq 5, \\ & x_2 \geq 0, \quad x_5 \geq 0\}, \end{aligned}$$

$$\text{есть задача } (\mathcal{D}^*, f^*): f^*(u) = 2u_1 + 4u_2 + 5u_3 \rightarrow \min,$$

$$\begin{aligned} \mathcal{D}^* = \{u \in R^3: & 3u_1 + u_3 = 2, \\ & 2u_1 + u_2 + 3u_3 \geq -3, \\ & -u_1 + u_2 + 2u_3 = 0, \\ & u_1 - 6u_2 + u_3 = 4, \\ & u_2 + 2u_3 \geq -1, \\ & u_1 \geq 0, \quad u_3 \geq 0\}, \end{aligned}$$

и задача $(\mathcal{D}, f)$ — двойственная к задаче $(\mathcal{D}^*, f^*)$.

Особый интерес представляет случай канонической задачи линейного программирования $(\mathcal{D}, f): f(x) = cx \rightarrow \max, \mathcal{D} = \{x \in R^n: Ax = b\}$. Ее условия получаем из схемы (16) при $k = m$ и $r = n$. Задача $(\mathcal{D}^*, f^*)$, двойственная к задаче $(\mathcal{D}, f)$, имеет вид

$$(\mathcal{D}^*, f^*): f^*(u) = ub \rightarrow \min, \mathcal{D}^* = \{u \in R^m: uA \geq c\}.$$

В заключение отметим, что установленная в этом пункте симметричность отношения двойственности в классе задач линейного программирования не имеет места в классе задач выпуклого программирования: задача линейного программирования может быть двойственной к задаче выпуклого программирования, существенно нелинейной (см. пример 2 п. 3.3). Обратное, естественно, неверно.

3.5. Теоремы двойственности.

Теорема 1. Если $(\mathcal{D}, f)$ — задача выпуклого программирования, а $(\mathcal{D}^*, f^*)$ — двойственная к ней, то $f(x) \leq f^*(u)$ для любых $x \in \mathcal{D}$, $u \in \mathcal{D}^*$.

Доказательство следует из свойства 2° сопряженных задач (см. п. 3.1). ■

Теорема 2. Пусть $(\mathcal{D}, f)$ — задача выпуклого программирования, для которой выполняются условия теоремы 2 или теоремы 3 п. 2.2, и $(\mathcal{D}^*, f^*)$ — двойственная к ней. Тогда для оптимальности $\bar{x} \in \mathcal{D}$ необходимо и достаточно существование такого $\bar{u} \in \mathcal{D}^*$, что

$$f(\bar{x}) = f^*(\bar{u}). \quad (17)$$

Достаточность очевидна: если $\bar{x} \in \mathcal{D}$, $\bar{u} \in \mathcal{D}^*$ и $f(\bar{x}) = f^*(\bar{u})$, то $f(\bar{x}) = g(\bar{u})$ и оптимальность $\bar{x}$ в задаче (R_f^{n+}, f) , а значит,

и в задаче $(\mathcal{D}, f)$ следует из утверждения 3° п. 3.1. Для доказательства необходимости отметим, что поскольку предположено выполнение условий теоремы Куна — Таккера, то из оптимальности $\bar{x} \in \mathcal{D}$ следует существование такого $\bar{u} \in R_{I_n}^{m+}$, что $(\bar{x}, \bar{u})$ — седловая точка функции Лагранжа Φ на $R_{I_n}^{n+} \times R_{I_n}^{m+}$, из чего в силу утверждения 4° п. 3.1 следует равенство $\bar{f}(\bar{x}) = \bar{g}(\bar{u})$. Поскольку $\bar{x} \in \mathcal{D}$, то $\bar{f}(\bar{x}) = f(\bar{x}) < +\infty$, а значит, $\bar{g}(\bar{u}) < +\infty$, $\bar{u} \in \mathcal{D}^*$ и $f^*(\bar{u}) = \bar{g}(\bar{u}) = f(\bar{x})$. ■

Если $\bar{x} \in \mathcal{D}$, $\bar{u} \in \mathcal{D}^*$ и $f(\bar{x}) = f^*(\bar{u})$, то из свойства 3° п. 3.1 следует, что $\bar{u}$ — решение задачи $(\mathcal{D}^*, f^*)$. Таким образом, при выполнении условий теоремы 2 из существования решения в задаче $(\mathcal{D}, f)$ следуют существование решения в задаче $(\mathcal{D}^*, f^*)$ и выполнение равенства (17) для любых решений $\bar{x}$ и $\bar{u}$ этих задач. Обратное утверждение, вообще говоря, неверно: задача $(\mathcal{D}^*, f^*)$ может иметь решение и в том случае, когда в задаче $(\mathcal{D}, f)$ решения не существует (см. упражнение 3.7). Однако если $(\mathcal{D}, f)$ — задача линейного программирования, то задачи $(\mathcal{D}, f)$ и $(\mathcal{D}^*, f^*)$ одновременно имеют или не имеют решения. Это сразу следует из сказанного выше, поскольку, как было установлено в п. 3.4, задача $(\mathcal{D}, f)$ двойственна к задаче линейного программирования $(\mathcal{D}^*, f^*)$.

Некоторые условия существования решения в задачах линейного программирования будут установлены в п. 3.7.

3.6. Геометрическая интерпретация пары двойственных задач линейного программирования. В настоящем пункте будет построена одна из возможных интерпретаций канонической задачи линейного программирования и двойственной к ней. Это делается как для формирования у читателя некоторых геометрических представлений о задачах двойственной пары, так и для доказательства некоторых утверждений, необходимых в дальнейших исследованиях. Выбор для интерпретации канонической задачи связан, с одной стороны, с боль-

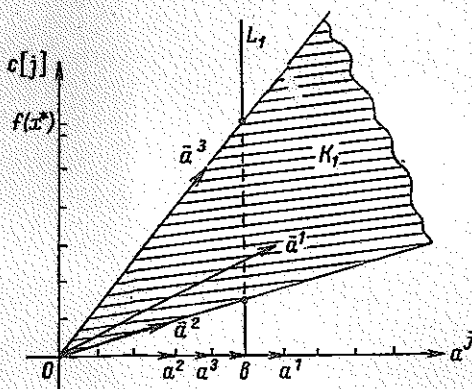


Рис. 20.

шей простотой возникающей здесь конструкции, с другой — с тем, что в силу существования канонических форм задач линейного программирования (см. п. 2.4.7) общность рассмотрений по существу не ограничивается.

Итак, пусть

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \quad \mathcal{D} = \{x \in R^{n+1} : Ax = b\}$$

— каноническая задача линейного программирования. Обозначим через $\bar{A}$ матрицу порядка $(m+1) \times n$, которая получается из матрицы A присоединением $(m+1)$ -й строки — вектора c ; столбцы матрицы $\bar{A}$ обозначим через $\bar{a}^j$, через $\bar{b}$ обозначим вектор из R^{m+1} , который получается присоединением к вектору b компоненты $b[m+1]$, равной нулю.

В пространстве R^{m+1} построим прямую $L = \{\bar{b} + \lambda e_{m+1}, \lambda \in R\}$, где $e_{m+1} = (0, \dots, 0, 1) \in R^{m+1}$, и многогранный выпуклый конус K с вершиной в нуле, натянутый на векторы $\bar{a}^j, j \in 1:n$ (см. п. 1.7.1).

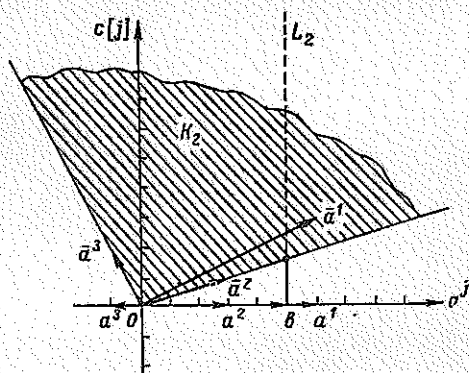


Рис. 21.

На рис. 20—22 изображены конусы K_1, K_2, K_3 и прямые L_1, L_2, L_3 , отвечающие соответственно задачам

$$(\mathcal{D}_1, f_1): f_1(x) = 3x_1 + x_2 + 5x_3 \rightarrow \max, \quad \mathcal{D}_1 = \{x \in R^3: 6x_1 + 3x_2 + 4x_3 = 5, x \geq 0\};$$

$$(\mathcal{D}_2, f_2): f_2(x) = 3x_1 + x_2 + 2x_3 \rightarrow \max, \quad \mathcal{D}_2 = \{x \in R^3: 6x_1 + 3x_2 - x_3 = 5, x \geq 0\}$$

$$\text{и } (\mathcal{D}_3, f_3): f_3(x) = 2x_1 - 2x_2 + x_3 \rightarrow \max, \quad \mathcal{D}_3 = \{x \in R^3: -2x_1 - 5x_2 - 3x_3 = 5, x \geq 0\}.$$

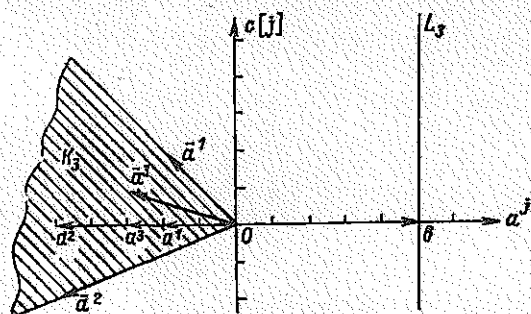


Рис. 22.

Пунктиром обозначены пересечения конуса K_1 с прямой L_1 и K_2 с L_2 ($K_3 \cap L_3 = \emptyset$).

Построим отображение φ множества $\mathcal{D}$ в прямую $L: \varphi(x) = \bar{b} + f(x)e_{m+1}$. Верны следующие утверждения.

1°. φ есть отображение $\mathcal{D}$ на $K \cap L$.

Пусть $x \in \mathcal{D}$. Тогда $Ax = b$, $cx = f(x)$, так что $\bar{A}x = \bar{b} + f(x)e_{m+1}$ и $x \in R^{n+}$. Следовательно, вектор $\bar{b} + f(x)e_{m+1} \in L$ есть неотрицательная линейная комбинация столбцов матрицы $\bar{A}$, так что $\bar{b} + f(x)e_{m+1} \in K$ (см. п. 1.7.2). Покажем теперь, что φ отображает $\mathcal{D}$ на $K \cap L$. Пусть $z \in K \cap L$. Поскольку $z \in K$ и $z \in L$, существуют такие $x \in R^{n+}$, $\lambda \in R$, что $\bar{A}x = z = \bar{b} + \lambda e_{m+1}$. Из последнего равенства следует $Ax = b$ и $cx = \lambda$. Тем самым установлено существование такого $x \in \mathcal{D}$ ($x \in R^{n+}$ и $Ax = b$), для которого $\varphi(x) = \bar{b} + f(x)e_{m+1} = z$. ■

2°. Условия $\mathcal{D} \neq \emptyset$ и $K \cap L \neq \emptyset$ эквивалентны.

Это утверждение очевидно следует из существования отображения из $\mathcal{D}$ на $K \cap L$. ■

Положим $\Lambda = \{\lambda \in R: \bar{b} + \lambda e_{m+1} \in K\}$.

3°. $\Lambda = \{f(x), x \in \mathcal{D}\}$.

Из определения отображения φ и множества Λ следует, что $f(x) \in \Lambda$ для любого $x \in \mathcal{D}$. Обратно, если $\lambda \in \Lambda$, то $\bar{b} + \lambda e_{m+1} \in K \cap L$ и существует такой $x \in \mathcal{D}$, что $\varphi(x) = \bar{b} + f(x)e_{m+1} = \bar{b} + \lambda e_{m+1}$, т. е. $f(x) = \lambda$. ■

4°. Если $\bar{x}$ — решение задачи $(\mathcal{D}, f)$, то $f(\bar{x}) = \max_{\lambda \in \Lambda} \lambda$. Если $\bar{\lambda} = \max_{\lambda \in \Lambda} \lambda$, $\bar{x} \in \mathcal{D}$ и $f(\bar{x}) = \bar{\lambda}$, то $\bar{x}$ — решение задачи $(\mathcal{D}, f)$.

Утверждение очевидно следует из 3°. ■

5°. Для существования оптимального плана в задаче $(\mathcal{D}, f)$ необходимо и достаточно выполнение условий: 1) $\mathcal{D} \neq \emptyset$; 2) множество $\Lambda = \{f(x), x \in \mathcal{D}\}$ ограничено сверху.

Если в задаче $(\mathcal{D}, f)$ существует оптимальный план $\bar{x}$, то $\mathcal{D} \neq \emptyset$ и $f(\bar{x}) = \max_{\lambda \in \Lambda} \lambda$ (см. 4°), так что множество Λ ограни-

чено сверху. Обратно, если $\mathcal{D} \neq \emptyset$, то $K \cap L \neq \emptyset$ (см. 2°). Поскольку конус K и прямая L — замкнутые выпуклые множества, то их пересечение $K \cap L$ — непустое замкнутое выпуклое подмножество прямой L , и если множество Λ ограничено сверху, то существует такое $\bar{\lambda} \in R$, что $\bar{\lambda} = \max_{\lambda \in \Lambda} \lambda$. Но тогда существует

такой $\bar{x} \in \mathcal{D}$, что $f(\bar{x}) = \bar{\lambda}$, и, следовательно, $\bar{x}$ — решение задачи $(\mathcal{D}, f)$. ■

Тем самым установлено, что решение канонической задачи линейного программирования сводится к нахождению параметра λ «последней» точки пересечения направленной прямой L с конусом K и неотрицательных коэффициентов разложения этой точки по образующим векторам $\bar{a}^i$ многогранного выпуклого конуса K .

Рассмотрим теперь множество U^* линейных функций u^* из R^m в R , удовлетворяющих условиям

$$u^*(a^j) \geq c[j], \quad j \in 1:n. \quad (18)$$

Каждая линейная функция $u^*: R^m \rightarrow R$ определяется вектором $u \in R^m$ (обозначаем его той же буквой, но без звездочки), так что $u^*(z) = uz$ для $z \in R^m$.

Из определения множества U^* следует, что $\{u \in R^m: u^* \in U^*\} = \{u \in R^m: ua^j \geq c[j], j \in 1:n\}$ есть множество $\mathcal{D}^*$ допустимых планов задачи, двойственной к задаче $(\mathcal{D}, f)$ (см. п. 3.4). Графиком функции u^* является проходящая через нуль невертикальная гиперплоскость

$$\Gamma_{u^*} = \{(z[1], \dots, z[m], z[m+1]) \in R^{m+1}:$$

$$\sum_{i=1}^m u[i]z[i] - z[m+1] = 0\},$$

или

$$\Gamma_{u^*} = \{z \in R^{m+1}: \hat{a}z = 0\},$$

где через $\hat{a}$ обозначен вектор $(u[1], \dots, u[m], -1)$, нормальный к гиперплоскости Γ_{u^*} . Если $u^* \in U^*$, то конус K находится в «нижнем» полупространстве $\bar{\Pi}^+ = \{z \in R^{m+1}: \hat{a}z \geq 0\}$, определяемом этой гиперплоскостью, поскольку из неравенств (18) следует, что в полупространстве $\bar{\Pi}^+$ находится каждый из векторов $\bar{a}^j$, $j \in 1:n$. Тем самым каждый допустимый план задачи $(\mathcal{D}^*, f^*)$ можно интерпретировать как такую линейную функцию из R^m в R , что конус K находится под ее графиком — проходящей через нуль невертикальной гиперплоскостью в R^{m+1} .

Очевидно, что график Γ_{u^*} линейной функции u^* пересекает прямую L в точке $\bar{b} + u^*(b)e_{m+1}$, т. е. параметр точки пересечения гиперплоскости Γ_{u^*} с прямой L равен $ub = f^*(u)$ — значению целевой функции задачи $(\mathcal{D}^*, f^*)$ в точке $u \in \mathcal{D}^*$. Поскольку множеству U^* принадлежат только такие линейные функции, графики которых расположены над конусом K , наименьшее возможное значение функции f^* на $\mathcal{D}^*$ равно $\bar{\lambda} = \max_{\lambda \in \Lambda} \lambda$.

Оно достигается на таком $\bar{u} \in \mathcal{D}^*$, для которого $\Gamma_{\bar{u}^*}$ является гиперплоскостью, опорной к конусу K в точке $\bar{b} + \bar{\lambda}e_{m+1}$. Тем самым установлено, что решение задачи $(\mathcal{D}^*, f^*)$, двойственной к канонической задаче линейного программирования $(\mathcal{D}, f)$, сводится к построению проходящей через нуль невертикальной гиперплоскости (она определяется нормалью $\hat{a}$), опорной к конусу K в «последней» точке пересечения K с прямой L .

На рис. 23 и 24 построены конусы K_1 , K_3 и прямые L_1 , L_3 , отвечающие задачам $(\mathcal{D}_1, f_1)$ и $(\mathcal{D}_3, f_3)$ (см. также рис. 20, 22). Гиперплоскости $\hat{a}_1z = 0$, $\hat{a}_3z = 0$, $\hat{u}z = 0$ (см. рис. 23) соответствуют допустимым планам задачи

$$(\mathcal{D}_1, f_1): f_1^*(u) = 5u \rightarrow \min,$$

$$\mathcal{D}^* = \{u \in R: 6u \geq 3, 3u \geq 1, 4u \geq 5\}.$$

В частности, вектор $\hat{u} = (1^{1/4}, -1)$ — нормальный вектор гиперплоскости, опорной к конусу K в «последней» точке пересечения K с прямой L (см. рис. 23), следовательно,

$\bar{u} = 1^{1/4}$ — решение задачи $(\mathcal{D}^*, f^*)$, $f^*(\bar{u}) = 6^{1/4}$.

Допустимым планам задачи $(\mathcal{D}_3^*, f_3^*)$ отвечают гиперплоскости $\hat{u}_1 z = 0$, $\hat{u}_2 z = 0$ (см. рис. 24). Задача не имеет решения: целевая функция f_3^* не ограничена снизу на множестве $\mathcal{D}_3^*$, поскольку множество точек пересечения прямой L_3 и гиперплоскостей, проходящих через нуль, не ограничено снизу. Не имеет решения и задача $(\mathcal{D}_3, f_3)$: множество $\mathcal{D}_3 = \emptyset$, прямая L_3 и конус K_3 (см. рис. 22) не имеют общих точек пересечения.

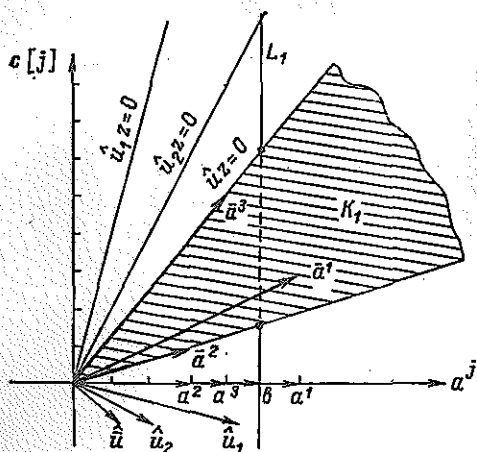


Рис. 23.

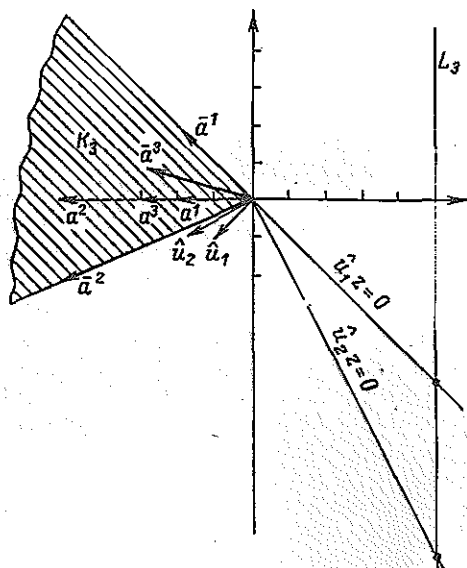


Рис. 24.

Аналогично не имеют решения задачи $(\mathcal{D}_2, f_2)$ и $(\mathcal{D}_2^*, f_2^*)$, поскольку множество $K_2 \cap L_2$ не ограничено сверху (см.

рис. 21) и тем самым не ограничено сверху множество Λ (см. утверждение 5°). Множество $\mathcal{D}_2^* = \emptyset$. Здесь, как легко видеть, не существует ни одной гиперплоскости, которая в соответствии с геометрической интерпретацией отвечала бы допустимому плану двойственной задачи.

3.7. Условие существования решения задачи линейного программирования.

Теорема. Пусть $(\mathcal{D}, f)$, $(\mathcal{D}^*, f^*)$ — пара двойственных задач линейного программирования. Для существования решения в одной из этих задач (а значит, и в обеих) необходимо и достаточно, чтобы $\mathcal{D} \neq \emptyset$, $\mathcal{D}^* \neq \emptyset$.

Доказательство. Необходимость приведенных условий очевидна. Она следует из теоремы 2 п. 3.5. Из симметричности отношений двойственности для линейной задачи заключаем, что при выполнении условий $\mathcal{D} \neq \emptyset$, $\mathcal{D}^* \neq \emptyset$ достаточно доказать существование решения в какой-либо из задач $(\mathcal{D}, f)$, $(\mathcal{D}^*, f^*)$, например в $(\mathcal{D}, f)$. В силу первой теоремы двойственности и условий $\mathcal{D} \neq \emptyset$, $\mathcal{D}^* \neq \emptyset$ целевая функция задачи $(\mathcal{D}, f)$ ограничена на непустом многогранном выпуклом множестве $\mathcal{D}$: для любого $u \in \mathcal{D}^*$ значение $f^*(u)$ служит верхней границей множества $\Lambda = \{f(x), x \in \mathcal{D}\}$. Остается проверить, что условие $\mathcal{D} \neq \emptyset$ и ограниченность сверху функции f (на множестве $\mathcal{D}$) задачи линейного программирования $(\mathcal{D}, f)$ достаточны для существования решения в задаче $(\mathcal{D}, f)$.

Для канонической задачи линейного программирования это утверждение было установлено в 5° п. 3.6. В общем случае построим каноническую форму $(\mathcal{D}', f')$ задачи $(\mathcal{D}, f)$ (см. п. 2.4.7). В силу свойств а) и г) эквивалентных экстремальных задач (см. п. 2.4.3) из непустоты множества $\mathcal{D}$ и ограниченности функции f на $\mathcal{D}$ следует непустота множества $\mathcal{D}'$ и ограниченность функции f' на $\mathcal{D}'$. Следовательно, задача $(\mathcal{D}', f')$, а вместе с ней и эквивалентная ей задача $(\mathcal{D}, f)$ (см. свойство в) п. 2.4.3) имеют решения. * ■

Аналогичное утверждение неверно для произвольной задачи выпуклого программирования. Например, в задаче

$$(\mathcal{D}, f) : f(x) = x_1 \rightarrow \max, \\ \mathcal{D} = \{x \in R^2 : x_1 \leq 1, -x_1 \leq 1, x_1^2 - x_2(1 - x_1^2) \leq 0\}$$

множества $\mathcal{D}$ и $\mathcal{D}^*$ непустые (см. упражнение 3.7), целевая функция f задачи ограничена сверху числом 1, однако оптимального плана в задаче $(\mathcal{D}, f)$ не существует.

3.8. Обсуждения.

1°. Об эквивалентности условий оптимальности. Рассмотренные в настоящей главе формально различные критерии опти-

* Другое доказательство см. в п. 1.8.2, 7°.

мальности (см. п. 1.7, теоремы 1 и 2 п. 2.2 и теорему 2 п. 3.5) равносильны. Остановимся на этом подробнее.

Пусть $(\mathcal{D}, f)$ — задача максимизации выпуклого программирования, функции f и f_i , $i \in 1:m$, непрерывно дифференцируемы и Φ — функция Лагранжа задачи $(\mathcal{D}, f)$. Предположим, что для векторов $\bar{x} \in R_J^{n+}$, $\bar{u} \in R_n^{m+}$ и $\bar{v} \in R^n$ выполняются соотношения (14) — (17) п. 1.7 при $u = \bar{u}$, $v = \bar{v}$. Тогда для любого $x \in R_J^{n+}$ имеет место неравенство $\Phi(x, \bar{u}) \leq \Phi(\bar{x}, \bar{u})$ (см. п. 2.4), а если $\bar{x} \in \mathcal{D}$, то и неравенство $\Phi(\bar{x}, \bar{u}) \leq \Phi(\bar{x}, u)$ верно для любого $u \in R_n^{m+}$. Итак, если выполнены соотношения (14) — (17) п. 1.7, то $(\bar{x}, \bar{u})$ — седловая точка функции Лагранжа задачи $(\mathcal{D}, f)$.

Пусть теперь $\bar{x} \in R_J^{n+}$, $\bar{u} \in R_n^{m+}$ и $(\bar{x}, \bar{u})$ — седловая точка функции Лагранжа задачи $(\mathcal{D}, f)$. Тогда (см. п. 2.3), $\bar{x} \in \mathcal{D}$ и $\sum_{i=1}^m \bar{u}[i] f_i(\bar{x}) = 0$. Следовательно, выполняются соотношения $\bar{u}[i] = 0$, $i \in I_n \setminus I(\bar{x})$. Так как $\Phi(x, \bar{u}) \leq \Phi(\bar{x}, \bar{u})$ при любом $x \in R_J^{n+}$, то $\bar{x}$ — решение задачи выпуклого программирования: $\Phi(x, \bar{u}) \rightarrow \max$, $x \in R_J^{n+}$. Поэтому в силу критерия оптимальности п. 1.7 должен существовать такой вектор $\bar{v} \in R^n$, что $\nabla_x \Phi(\bar{x}, \bar{u}) = \nabla f(\bar{x}) - \sum_{i=1}^m \bar{u}[i] \nabla f_i(\bar{x}) = -\bar{v}$ и $\bar{v}[j] \geq 0$, $j \in I(\bar{x})$, $\bar{v}[j] = 0$, $j \in I(\bar{x})$. Итак, если $(\bar{x}, \bar{u})$ — седловая точка функции Лагранжа задачи $(\mathcal{D}, f)$, то для векторов $\bar{x}$, $\bar{u}$ и $\bar{v} = -\nabla_x \Phi(\bar{x}, \bar{u})$ выполняются соотношения (14) — (17) п. 1.7.

Равносильность теорем 1 и 2 п. 2.2 и теоремы 2 п. 3.5 фактически была выяснена при доказательстве этой теоремы в п. 3.5.

2°. Об экономической интерпретации двойственных оценок. Если $(\bar{x}, \bar{u})$ — седловая точка функции Лагранжа задачи выпуклого программирования $(\mathcal{D}, f)$ на $R_J^{n+} \times R_n^{m+}$, то $\bar{x}$ — решение задачи $(\mathcal{D}, f)$ и, как установлено в п. 3.5, вектор $\bar{u}$ является решением задачи $(\mathcal{D}^*, f^*)$, двойственной к задаче $(\mathcal{D}, f)$ (конечно, при выполнении условий теоремы 2 п. 3.5). Вектор $\bar{u}$ называют также вектором *двойственных оценок* (или о. о. оценок [20]). Для некоторых специальных формулировок задачи $(\mathcal{D}, f)$ может быть дана интересная экономическая его интерпретация.

Рассмотрим задачу выпуклого программирования вида $(\mathcal{D}, f)$: $f(x) \rightarrow \max$,

$$\mathcal{D} = \{x \in R^n : f_i(x) = g_i(x) - b[i] \leq 0, i \in 1:m\}.$$

Вектор x будем истолковывать как план производства; $g_i(x)$ — затраты i -го ресурса, необходимые для выполнения плана x ; $b[i]$ — количество i -го ресурса, $i \in 1:m$; $f(x)$ — стоимость произведенной продукции при реализации плана x . Естественно

считать оптимальный план $\bar{x}$ задачи $(\mathcal{D}, f)$ функцией $\bar{x}(b)$ вектора ресурсов b , определенной на некотором множестве $\mathcal{B}$ допустимых значений вектора ресурсов. Предположим, что функции f , $\bar{x}$ и g_i , $i \in 1:m$, непрерывно дифференцируемы и для задачи $(\mathcal{D}, f)$ выполняются условия регулярности при $b = b^0 \in \mathcal{B}$. Кроме того, предположим, что $I(\bar{x}(b)) = I(\bar{x}(b^0))$ в некоторой окрестности точки b^0 (виды ресурсов, лимитирующие оптимальный план $\bar{x}(b^0)$, лимитируют и $\bar{x}(b)$, если b мало отличается от b^0).

Пусть $\bar{u}$ — вектор двойственных оценок для плана $\bar{x}(b^0)$. Тогда (см. 1° и критерий п. 1.7), $\bar{u} \geq 0$,

$$\nabla f(\bar{x}(b^0)) = \sum_{i=1}^m \bar{u}[i] \nabla g_i(\bar{x}(b^0)), \quad (19)$$

$$\bar{u}[i] = 0, \quad i \in I(\bar{x}(b^0)). \quad (20)$$

Докажем, что при сделанных предположениях

$$\partial f(\bar{x}(b^0))/\partial b[k] = \bar{u}[k], \quad k \in 1:m. \quad (21)$$

Равенства (21) позволяют интерпретировать $\bar{u}$ как вектор оценок (цен) ресурсов (при $b = b^0$): $\bar{u}[k]$ — прирост стоимости произведенной продукции на единицу прироста k -го ресурса. Величина $\Phi(x, \bar{u}) = f(x) - \sum_{i=1}^m \bar{u}[i] g_i(x) + \bar{u}b^0$ также получает любопытное толкование: $\Phi(x, \bar{u})$ отличается от стоимости вновь произведенной продукции (при ценах на ресурсы, связанных с оптимальным планом производства) лишь на постоянную — стоимость заданных ресурсов.

Докажем равенства (21). По правилу дифференцирования сложной функции ($k \in 1:m$)

$$\frac{\partial f(\bar{x}(b^0))}{\partial b[k]} = \sum_{j=1}^n \frac{\partial f(\bar{x}(b^0))}{\partial x[j]} \frac{\partial \bar{x}(b^0)[j]}{\partial b[k]},$$

откуда в силу (19) и (20)

$$\begin{aligned} \frac{\partial f(\bar{x}(b^0))}{\partial b[k]} &= \sum_{j=1}^n \left(\sum_{i=1}^m \bar{u}[i] \frac{\partial g_i(\bar{x}(b^0))}{\partial x[j]} \right) \frac{\partial \bar{x}(b^0)[j]}{\partial b[k]} = \\ &= \sum_{i=1}^m \bar{u}[i] \sum_{j=1}^n \frac{\partial g_i(\bar{x}(b^0))}{\partial x[j]} \frac{\partial \bar{x}(b^0)[j]}{\partial b[k]} = \sum_{i \in I(\bar{x}(b^0))} \bar{u}[i] \frac{\partial g_i(\bar{x}(b^0))}{\partial b[k]} \end{aligned} \quad (22)$$

Если $i \in I(\bar{x}(b^0))$, то в силу сделанного предположения $g_i(\bar{x}(b)) = b[i]$ для всех b из некоторой окрестности точки b^0 . Дифференцируя это равенство по $b[k]$, получаем

$$\frac{\partial g_i(\bar{x}(b^0))}{\partial b[k]} = \begin{cases} 1 & \text{при } i=k, \\ 0 & \text{при } i \neq k. \end{cases} \quad (23)$$

Из (22) и (23) следует (21). ■

Приведенную экономическую интерпретацию двойственных оценок полезно сравнить со сказанным по этому поводу в п. 1.11.

3°. *Существование условия регулярности.* Рассмотрим экстремальную задачу $(\mathcal{D}, f): f(x) \rightarrow \max, \mathcal{D} = \{x \in R: x^2 \leq 0\}$. Очевидно, что $(\mathcal{D}, f)$ является задачей выпуклого программирования и для нее выполняются все условия, при которых был выведен критерий оптимальности п. 1.7, кроме условия С. Область $\mathcal{D}$ состоит из единственной точки $\bar{x} = 0$, которая и является решением задачи. Однако вектора u из R^+ , удовлетворяющего равенству $1 = 0 \cdot u$, не существует.

Построим функцию Лагранжа задачи $(\mathcal{D}, f): \Phi(x, u) = x - ux^2$. Точка $(0, \bar{u})$ не может быть седловой для функции Φ ни при каком $\bar{u} \in R^+$, так как неравенство $\Phi(x, \bar{u}) \leq \Phi(0, \bar{u}) = 0$ не может выполняться для всех $x \in R$ ни при каком неотрицательном $\bar{u}$.

4°. *Критерии оптимальности и выпуклость.* Пусть $(\mathcal{D}, f):$

$$f(x) \rightarrow \max, \mathcal{D} = \{x \in R^m: f_i(x) \leq 0, i \in I_n, f_i(x) = 0, i \in I_y\}$$

— экстремальная задача и $\Phi(x, u) = f(x) - \sum_{i=1}^m u[i] f_i(x)$ — функция Лагранжа. Теорема 1 п. 2.2 утверждает достаточность выполнения неравенств седловой точки для оптимальности $\bar{x} \in R^m$. Напомним, что утверждение теоремы верно без каких бы то ни было предположений о свойствах задачи $(\mathcal{D}, f)$, в частности без предположений о выпуклости функций f и $f_i, i \in 1:m$. В то же время утверждения теорем 2 и 3 п. 2.2 перестают быть верными, если задача $(\mathcal{D}, f)$ не является задачей выпуклого программирования. Например, в задаче $(\mathcal{D}, f): f(x) = x^2 \rightarrow \max, \mathcal{D} = \{x \in R: -1 \leq x \leq 3\}$ точка $\bar{x} = 3$ — решение, но, как нетрудно проверить, функция Лагранжа задачи не имеет седловой точки в $R \times R^2$.

УПРАЖНЕНИЯ

В упражнениях 3.1—3.3 написать условия оптимальности допустимого плана экстремальных задач.

3.1. $(\mathcal{D}, f): f(x) = cx \rightarrow \max, \mathcal{D} = \{x \in R^n: Ax \leq b, x \geq 0\}$.

3.2. $(\mathcal{D}, f): f(x) = cx \rightarrow \min, \mathcal{D} = \{x \in R^n: Ax \geq b, x \geq 0\}$.

3.3. $(\mathcal{D}, f): f(x) = \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j] \rightarrow \min, \mathcal{D} = \{x \in R^{mn}: \sum_{j=1}^n x[i, j] \leq a[i], i \in 1:m; \sum_{i=1}^m \lambda[i, j] x[i, j] = b[j], j \in 1:n; x \geq 0\}$ (распределительная задача, или λ -задача, формулировку модели см., например, в [10]).

3.4. Написать условия оптимальности допустимого плана задачи о «диете» (см. п. 2.3.7) и задачи сглаживания (см. п. 2.3.12).

3.5. Проверить, является ли точка $(\bar{x}, \bar{u}) = (0, 0)$ седловой точкой функции $f(x, u) = x^2/a^2 + u^2/b^2$ в области $R \times R$. Является ли $(\bar{x}, \bar{u})$ точкой абсолютного минимума функции f ? Построить график функции f .

3.6. Заданы множества $X = \{1\} \subset R$, $U = [-2, -1] \subset R$ и функция $G: R^2 \rightarrow R$, $G(x, u) = x/u$. Построить сопряженные задачи относительно функции G и найти их решения.

3.7. Построить задачу, двойственную к задаче выпуклого программирования $(\mathcal{D}, f): f(x) = x_1 \rightarrow \max$, $\mathcal{D} = \{x \in R^2: x_1 \leq 1, -x_1 \leq 1, x_1^2 - x_2(1 - x_1^2) \leq 0\}$ (см. п. 3.7) и найти ее решение.

3.8. Построить задачи, двойственные к задачам пп. 2.3.2—2.3.7.

3.9. Пусть $\bar{x}$ — решение задачи линейного программирования $(\mathcal{D}, f)$, $\bar{u}$ — решение двойственной к ней задачи $(\mathcal{D}^*, f^*)$. Является ли вектор $(\bar{x}, \bar{u})$ седловой точкой функции Лагранжа задачи $(\mathcal{D}, f)$? Является ли вектор $(\bar{u}, \bar{x})$ седловой точкой функции Лагранжа задачи $(\mathcal{D}^*, f^*)$?

3.10. Пусть $(\bar{x}, \bar{u})$ — седловая точка функции Лагранжа задачи линейного программирования $(\mathcal{D}, f)$, а $(\mathcal{D}^*, f^*)$ — двойственная к ней задача. Являются ли векторы $\bar{x}$ и $\bar{u}$ решениями задач $(\mathcal{D}, f)$ и $(\mathcal{D}^*, f^*)$?

3.11. Доказать, что задачи $(\mathcal{D}, f): f(x) = cx \rightarrow \max$, $\mathcal{D} = \{x \in R^n: Ax \leq b_1, x \geq 0\}$ и $(\mathcal{D}', f'): f'(x) = cx \rightarrow \max$, $\mathcal{D}' = \{x \in R^n: Ax \leq b_2, x \geq 0\}$, где $b_1 \geq 0$, $b_2 \geq 0$, одновременно имеют или не имеют решения.

Глава 4

ОСНОВНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

§ 1. Предварительные замечания

1.1. О содержании главы. В настоящей главе рассматриваются основные вычислительные методы решения канонической задачи линейного программирования

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \quad \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\}$$

и некоторые вспомогательные вычислительные процедуры. Так как для любой задачи линейного программирования можно построить эквивалентную ей каноническую задачу линейного программирования — ее каноническую форму (см. п. 2.4.7), то эти методы являются основными вычислительными методами линейного программирования.

Напомним, что в силу теорем двойственности (см. п. 3.3.5, а также п. 3.3.7) задача линейного программирования $(\mathcal{D}, f)$ имеет решение в том и только том случае, если имеет решение двойственная к ней задача

$$(\mathcal{D}^*, f^*): f^*(u) = ub \rightarrow \min, \quad \mathcal{D}^* = \{u \in R^m: uA \geq c\}.$$

Применение методов настоящей главы фактически позволяет вместе с решением одной из этих задач строить решение и двойственной к ней задачи, если только обе они имеют решения.

1.2. О машинной реализации. Методы, рассматриваемые в данной и последующих главах, допускают машинную реализацию, так что вычисления можно проводить на ЭВМ. Более того, уже имеются составленные на основе этих методов стандартные программы, позволяющие решать задачи линейного программирования.

Составление машинных программ — трудная и кропотливая работа. Естественно, что программист должен передать ЭВМ логику алгоритма. Но, кроме того, машинная реализация требует решения целого ряда специфических проблем, связанных с организацией вычислений. Необходимо учитывать специфику

используемой ЭВМ и языка программирования, объем и состав исходной информации, точность вычислений, решать вопросы размещения массивов исходной и текущей информации и т. п. Следует иметь в виду, что при выборе метода решения реальной задачи на первый план могут выйти именно упомянутые выше проблемы. Они оказываются наиболее существенными в большинстве практических ситуаций. Однако эти проблемы по существу не обсуждаются в данном курсе. Его цель — познакомить читателя с теорией соответствующих экстремальных задач и основными принципами формирования вычислительных алгоритмов их решения.

В практической работе экономиста-математика могут встретиться задачи (в том числе и линейного программирования), решение которых в силу целого ряда причин не может быть найдено на ЭВМ с использованием имеющихся программ. В этом случае возникает необходимость в формировании вычислительной процедуры, учитывающей специфику рассматриваемой задачи, а значит, и необходимость в знании уже разработанных вычислительных методов и их теоретического обоснования. Поэтому в данном учебном пособии достаточно большое внимание уделяется различным вычислительным методам решения экстремальных задач.

Особо отметим, что вычислительные алгоритмы мы рассматриваем не только как предписания, позволяющие конструктивно строить решение экстремальной задачи, но и как математический аппарат, находящий применение в теоретических исследованиях (см., например, замечание п. 3.7).

1.3. О структуре. Параграфы, в которых излагаются вычислительные методы, имеют одинаковую структуру. В первом пункте вводятся необходимые обозначения и понятия или указываются те места книги, где они были введены впервые. Во втором приводится краткая характеристика метода. Содержание третьего пункта составляет вычислительная схема алгоритма. Трудности, которые могут возникнуть при его реализации, обсуждаются в четвертом пункте. Алгоритм всегда иллюстрируется примером, который рассматривается в пятом. Рекомендуем совместить изучение алгоритма с разбором примера. В шестом пункте формулируются утверждения, необходимые для обоснования алгоритма. В седьмом приводятся их доказательства. Восьмой пункт посвящен рассмотрению вспомогательных задач, решение которых может оказаться полезным для реализации вычислений. Такая структура параграфов допускает выборочное чтение. Так, читателю, интересующемуся только вычислительными аспектами метода, можно предложить изучение пунктов и последовательности $1 \rightarrow 3 \rightarrow 5$.

1.4. О вырожденности. В настоящей главе обоснование алгоритмов проведено для невырожденных задач $(\mathcal{D}, f)$ или $(\mathcal{D}^*, f^*)$. Задача называется невырожденной, если в множестве $\mathcal{D}$ ($\mathcal{D}^*$)

нет вырожденных планов. Определение вырожденного плана для каждой из названных задач приведено в тексте главы. Вырожденные задачи можно решать, незначительно изменив соответствующие алгоритмы. Более подробно о решении вырожденных задач см. в § 7.

1.5. Последовательность чтения алгоритмов. Алгоритмы данной и последующих глав подразделяются на этапы, каждый из которых состоит из одного или нескольких пунктов, содержащих вычислительные рекомендации. Если в пункте не указывается «адрес перехода», то после выполнения вычислений следует приступить к выполнению следующей по порядку рекомендации.

§ 2. Методы допустимых направлений для решения задач линейного программирования

В этом параграфе рассматривается общая схема вычислительных методов, которые называются методами допустимых (возможных) направлений. Определенные конкретизации этой схемы для задач линейного программирования — канонической задачи (п. 2.2.2) и задачи с ограничениями в форме неравенств, т. е. двойственной к канонической (п. 2.2.3), — приводят к основным алгоритмам решения таких задач: симплекс-методу (§ 3, 4) и двойственному симплекс-методу (§ 5). Эти методы принято называть также *методами последовательного улучшения плана*, тем самым подчеркивая, что переход от одного допустимого плана к другому осуществляется с «улучшением» значения целевой функции.

2.1. Методы допустимых направлений. Вычислительный процесс, состоящий из ряда повторяющихся операций — *итераций*, называется *итеративным*, а методы решения задач, включающие итеративные вычислительные процессы, — *итеративными методами*.

Итеративный метод решения задачи выпуклого программирования $(\mathcal{D}, f)$ называется *методом допустимых направлений*, если итерации этого метода складываются из следующих вычислительных операций.

1. Вычислительный процесс начинается с некоторой точки $x^0 \in \mathcal{D}$ — начального допустимого плана.

2. На $(k+1)$ -й итерации, $k=0, 1, 2, \dots$, после того как построены векторы $x^0, x^1, \dots, x^k$ из $\mathcal{D}$, осуществляется проверка оптимальности вектора x^k . Если $x^k = \bar{x}$ — решение задачи, то вычислительный процесс заканчивается. Остается вычислить значение $f(\bar{x})$ целевой функции на плане $\bar{x} = x^k$ и другие величины, связанные с оптимальным планом $\bar{x}$ (например, оптимальный план двойственной задачи), если они интересуют решающего.

3. Если допустимый план x^k не является оптимальным, то строится такой вектор $x^{k+1} \in \mathcal{D}$, что $f(x^{k+1}) > f(x^k)$. Процедура построения вектора x^{k+1} складывается из двух этапов:

3'. Выбора допустимого прогрессивного направления s^k в точке x^k (существование допустимого прогрессивного направления в точке x^k следует из неоптимальности плана x^k);

3''. Выбора шага, т. е. такого числа $\lambda_k > 0$, что $x^k + \lambda_k s^k \in \mathcal{D}$ и $f(x^k + \lambda_k s^k) > f(x^k)$.

Вектор x^{k+1} полагают равным $x^k + \lambda_k s^k$.

Метод допустимых направлений называется *сходящимся*, если его применение приводит к построению оптимального плана после конечного числа итераций либо такой последовательности $\{x^k\}_{k=0}^\infty$ из $\mathcal{D}$, что $f(x^k) \rightarrow \max_{x \in \mathcal{D}} f(x)$ при $k \rightarrow \infty$. На-

зовем метод допустимых направлений *конечным*, если после конечного числа итераций он приводит к построению оптимального плана или устанавливает неограниченность f на $\mathcal{D}$. В противном случае метод будем называть *бесконечным*. В настоящей главе будут рассмотрены только конечные методы решения задач линейного программирования.

2.2. Общая схема методов допустимых направлений для решения канонической задачи линейного программирования. Рассмотрим каноническую задачу линейного программирования

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\},$$

где $A = (a[i, j])$ — $m \times n$ матрица, $c = (c[1], c[2], \dots, c[n])$ и $b^* = (b[1], b[2], \dots, b[m])$. Не умаляя общности, можно считать вектор b неотрицательным ($b \geq 0$), $m \leq n$ и ранг $r(A)$ матрицы A равным m .

Как уже отмечалось, методы допустимых направлений решения задачи $(\mathcal{D}, f)$ включают начальную процедуру поиска допустимого плана x^0 и собственно итеративный процесс: проверку на каждой итерации оптимальности очередного допустимого плана задачи, выбор допустимого прогрессивного направления, если план оказался неоптимальным, определение шага и вычисление следующего допустимого плана. Рассмотрим эти вычислительные процедуры подробнее.

1°. *Нахождение начального плана.* Иногда начальный допустимый план x^0 задачи $(\mathcal{D}, f)$ удается указать сразу. Так, например, если в матрице ограничений A есть m попарно различных единичных столбцов $a^{i_1}, a^{i_2}, \dots, a^{i_m}$, то план x^0 , у которого m компонент $x[j_i] = b[i]$, $i \in 1:m$, а остальные $n - m$ компонент равны нулю, принадлежит $\mathcal{D}$. Однако в общем случае нахождение начального плана x^0 составляет самостоятельную задачу. Для ее решения обычно формулируется следующая вспомогательная задача линейного программирования:

$(\mathcal{D}', \psi): \psi(z) = c'z \rightarrow \max, \mathcal{D}' = \{z \in R^{n+m}: A'z = b, z \geq 0\}$, где

$$A' = [A] = \begin{bmatrix} a[1, 1] & a[1, 2] & \dots & a[1, n] & 1 & 0 & \dots & 0 \\ a[2, 1] & a[2, 2] & \dots & a[2, n] & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a[m, 1] & a[m, 2] & \dots & a[m, n] & 0 & 0 & \dots & 1 \end{bmatrix}$$

— $m \times (n + m)$ матрица ограничений, а

$$c' = (\underbrace{0, \dots, 0}_n, \underbrace{-1, -1, \dots, -1}_m).$$

Например, для задачи линейного программирования

$$\begin{aligned} (\mathcal{D}, f): f(x) &= 6x_1 + 4x_2 + 5x_3 + 4x_4 \rightarrow \max, \mathcal{D} = \{x \in R^4: \\ 5x_1 + 4x_2 - 5x_3 + 2x_4 &= 2, \\ 3x_1 + 4x_2 + 2x_3 + 6x_4 &= 5, \\ x_j &\geq 0, j \in 1:4\} \end{aligned}$$

строится вспомогательная задача $(\mathcal{D}', \psi): \psi(z) = -z_5 - z_6 \rightarrow \max$,

$$\begin{aligned} \mathcal{D}' = \{z \in R^6: 5z_1 + 4z_2 - 5z_3 + 2z_4 + z_5 &= 2, \\ 3z_1 + 4z_2 + 2z_3 + 6z_4 + z_6 &= 5, \\ z_j &\geq 0, j \in 1:6\}. \end{aligned}$$

Оправданием интереса к задаче $(\mathcal{D}', \psi)$ являются следующие ее свойства.

Допустимый начальный план задачи $(\mathcal{D}', \psi)$ можно указать непосредственно: действительно, план $z^0 = (0, \dots, 0, b[1], \dots, b[m]) \in \mathcal{D}'$, поскольку $b \geq 0$. ■

В рассматриваемом примере вектор $z^0 = (0, 0, 0, 0, 2, 5)$ является допустимым планом задачи $(\mathcal{D}', \psi)$.

Задача $(\mathcal{D}', \psi)$ всегда имеет решение.

В силу утверждения, доказанного в п. 3.3.7, достаточно установить непустоту множества $\mathcal{D}'$ и множества $(\mathcal{D}')^*$ допустимых планов задачи $((\mathcal{D}')^*, \psi^*)$, двойственной к задаче $(\mathcal{D}', \psi)$. Непустоту множества $\mathcal{D}'$ мы уже установили. Предоставляем читателю проверить самостоятельно, что вектор $z^* = (0, \dots, 0) \in R^m$ принадлежит $(\mathcal{D}')^*$. — Упражнение 4.1. ■

Если $\bar{z} = (\bar{z}[1], \bar{z}[2], \dots, \bar{z}[n + m])$ — решение задачи $(\mathcal{D}', \psi)$ и $\psi(\bar{z}) = 0$, то $x^0 = (\bar{z}[1], \dots, \bar{z}[n]) \in \mathcal{D}$. Если $\psi(\bar{z}) < 0$, то $\mathcal{D} = \emptyset$.

В самом деле, в силу неотрицательности вектора $\bar{z}$ из равенства $0 = \psi(\bar{z}) = c'\bar{z} = -\sum_{i=n+1}^{n+m} \bar{z}[i]$ следует $\bar{z}[n + 1] = \dots = \bar{z}[n + m] = 0$. Поэтому $x^0 \geq 0$ и $Ax^0 = A'\bar{z} = b$.

Если $\mathcal{D} \neq \emptyset$ и $x' = (x'[1], \dots, x'[n]) \in \mathcal{D}$, то очевидно $u' = (x'[1], \dots, x'[n], 0, \dots, 0) \in \mathcal{D}'$, $\psi(u') = c'u' = 0 > \psi(\bar{u})$, что противоречит оптимальности $\bar{u}$ в задаче $(\mathcal{D}', \psi)$. ■

Итак, если для решения задачи $(\mathcal{D}, \psi)$ применяется некоторый конечный итеративный метод, то вычисления можно начать с плана z^0 , и за конечное число итераций будет найдено решение $\bar{z}$, позволяющее сформировать план $x^0 \in \mathcal{D}$ либо обнаружить пустоту множества $\mathcal{D}$.

Приведенный выше метод нахождения начального допустимого плана канонической задачи линейного программирования называется *методом минимизации невязок*.

2°. Проверка оптимальности. Для проверки оптимальности или неоптимальности плана x^k , $k = 0, 1, \dots$, воспользуемся критерием п. 3.1.7. Рассмотрим множество индексов $J(x^k) = \{j \in 1:n: x^k[j] = 0\}$. Составим систему из n линейных уравнений и неравенств

$$a^j = c[j], \quad j \in J(x^k), \quad (1)$$

$$a^j \geq c[j], \quad j \in J(x^k), \quad (2)$$

где a^j — j -й столбец матрицы A . Проверка оптимальности вектора x^k сводится к проверке совместности системы (1), (2): если система совместна, то x^k — решение задачи $(\mathcal{D}, f)$; в противном случае в точке x^k существует допустимое прогрессивное направление.

Проверку совместности системы (1), (2) легко осуществить, если система уравнений (1) имеет единственное решение $\bar{u}$ (именно этот случай представляет практический интерес). Очевидно, что при этом условии система (1), (2) совместна в том и только том случае, если вектор $\bar{u}$ является решением системы неравенств (2). Поэтому для проверки оптимальности вектора x^k достаточно решить систему уравнений (1) и проверить, удовлетворяет ли ее решение — вектор $\bar{u}$ — системе неравенств (2).

Проверим оптимальность допустимого плана $x^k = (0, 1/8, 0, 3/4)$ рассмотренной выше задачи. Множество $J(x^k) = \{1, 3\}$, и для оптимальности x^k необходимо и достаточно существование такого $u = (u_1, u_2)$, что

$$4u_1 + 4u_2 = 4, \quad 2u_1 + 6u_2 = 4, \quad (1')$$

$$5u_1 + 3u_2 \geq 6, \quad -5u_1 + 2u_2 \geq 5. \quad (2')$$

Система (1') имеет единственное решение $\bar{u} = (1/2, 1/2)$, однако этот вектор не удовлетворяет системе неравенств (2'). Следовательно, план x^k не оптимален и в точке x^k существует допустимое прогрессивное направление.

3°. Поиск допустимого прогрессивного направления. Нахождение допустимого прогрессивного направления z^k в точке $x^k \in \mathcal{D}$ осуществляется в соответствии с условиями допустимости (см. п. 3.1.3) и прогрессивности (см. п. 3.1.4) направления в задаче выпуклого программирования. Для канонической

задачи линейного программирования эти условия имеют вид

$$As = 0, \quad (3)$$

$$s[j] \geq 0, \quad j \in J(x^k), \quad (4)$$

$$cs > 0. \quad (5)$$

Они получаются из условий (5)–(8) п. 3.1.3 и условия (12) п. 3.1.4 при $I_n = \emptyset$ и $\nabla f(x^k) = c$. Если x^k не является решением задачи $(\mathcal{D}, f)$, то система (3)–(5) совместна и каждое ее решение является допустимым прогрессивным направлением в точке x^k .

Мы уже проверили, что план $x^k = (0, 1/8, 0, 1/4)$ не является оптимальным и, следовательно, в точке x^k существует допустимое прогрессивное направление s^k . Множество $J(x^k) = \{1, 3\}$. Составим следующую систему уравнений и неравенств:

$$5s_1 + 4s_2 - 5s_3 + 2s_4 = 0,$$

$$3s_1 + 4s_2 + 2s_3 + 6s_4 = 0,$$

$$s_1 \geq 0, \quad s_3 \geq 0,$$

$$6s_1 + 4s_2 + 5s_3 + 4s_4 > 0.$$

Она совместна: вектор $s^k = (0, 17/8, 1, -7/4)$ — ее решение. Поэтому s^k является допустимым прогрессивным направлением в точке x^k .

Множество $S^{(k)}$ решений системы (3)–(5) состоит из пересечения многогранного выпуклого конуса допустимых направлений в точке x^k , определяемого условиями (3) и (4), с (открытым) полупространством $\{s \in R^n: cs > 0\}$. Поэтому если $s \in S^{(k)}$, то и $\lambda s \in S^{(k)}$ при $\lambda > 0$, и проблема построения допустимого прогрессивного направления сводится не просто к выбору при каждом $k = 0, 1, \dots$ некоторого $s \in S^{(k)}$, но и к указанию правила наиболее «эффективного» выбора такого s , позволяющего обеспечить минимальность числа итераций и максимальное возрастание целевой функции на каждой из них.

Если $\dim S^{(k)} = 1$, то безразлично, какое $s \in S^{(k)}$ выбирается: максимальное возрастание целевой функции можно обеспечить за счет выбора шага. Если $\dim S^{(k)} > 1$, то обычно стараются выбрать такое $s^k \in S^{(k)}$, чтобы $\nabla f(x^k)s^k = cs^k \geq cs$ для любого $s \in S^{(k)}$, удовлетворяющего, как и s^k , некоторому условию нормировки. Однако это правило может не обеспечить ни максимального возрастания целевой функции на данной итерации, поскольку вектор x^{k+1} зависит не только от s^k , но и от величины шага λ_k , ни минимальности числа итераций. Мы не будем давать здесь каких-то универсальных рекомендаций по поводу эффективного выбора $s^k \in S^{(k)}$. Более того, все вычислительные методы данной главы и большинство методов последующих глав отличаются друг от друга именно тем, каким способом осуществляется этот выбор на каждой итерации.

4°. **Выбор шага.** Если s^k — допустимое прогрессивное направление в точке x^k , то существует такое $\lambda' > 0$, что

$$x^k + \lambda s^k \in \mathcal{D}, \quad (6)$$

$$f(x^k + \lambda s^k) > f(x^k) \quad (7)$$

при любом $\lambda \in (0, \lambda']$. Более того, в силу линейности f и прогрессивности направления s^k функция $\varphi(\lambda) = f(x^k + \lambda s^k) = cx^k + \lambda cs^k$, возрастающая функция от λ и $\lim_{\lambda \rightarrow +\infty} \varphi(\lambda) = +\infty$ при $\lambda \rightarrow +\infty$. Поэтому (7) выполняется при любом $\lambda > 0$, и поскольку мы преследуем цель максимального увеличения $f(x^k + \lambda s^k)$, то λ_k следует положить равным $\max\{\lambda: x^k + \lambda s^k \in \mathcal{D}, \lambda > 0\}$. Из определения множества $\mathcal{D}$ следует, что вектор $x^k + \lambda s^k$ в том и только том случае принадлежит $\mathcal{D}$, если

$$A(x^k + \lambda s^k) = b, \quad (8)$$

$$x^k + \lambda s^k \geq 0. \quad (9)$$

Так как $Ax^k = b$ и $As^k = 0$, то соотношение (8) выполняется при любом $\lambda \in R$. Если при этом $s^k \geq 0$, то неравенство (9) верно при любом $\lambda > 0$ (вектор s^k принадлежит конусу неограниченности многогранного выпуклого множества $\mathcal{D}$ (см. п. 1.6.2)), и поскольку $cs^k > 0$, то функция f не ограничена сверху на $\mathcal{D}$ (см. п. 1.8.2, 6°).

Предположим теперь, что найдется $j \in 1:n \setminus J(x^k)$, для которого $s^k[j] < 0$. Для такого j неравенство $x^k[j] + \lambda s^k[j] \geq 0$ нарушается при $\lambda > x^k[j] / -s^k[j]$. Поэтому λ_k следует положить равным наибольшему из чисел λ , для которых выполняются все неравенства $x^k[j] + \lambda s^k[j] \geq 0$ при $j \in 1:n \setminus J(x^k)$, $s^k[j] < 0$, т. е.

$$\lambda_k = \min_{j: s^k[j] < 0} x^k[j] / -s^k[j]. \quad (10)$$

Итак, если существует такое $j \in 1:n \setminus J(x^k)$, для которого $s^k[j] < 0$, то можно определить шаг λ_k по формуле (10) и положить $x^{k+1} = x^k + \lambda_k s^k$. Это завершает $(k+1)$ -ю итерацию метода допустимых направлений: осуществлен переход от плана $x^k \in \mathcal{D}$ к плану $x^{k+1} \in \mathcal{D}$ задачи $(\mathcal{D}, f)$.

В рассматриваемом примере в точке $x^k = (0, 1/8, 0, 3/4)$ было выбрано допустимое прогрессивное направление $s^k = (0, 17/8, 1, -7/4)$. Напоминаем, что здесь $J(x^k) = \{1, 3\}$, $1:n \setminus J(x^k) = \{2, 4\}$, и поскольку $s^k[4] = -7/4 < 0$, то (см. (10))

$$\begin{aligned} \lambda_k &= 3/4 : 7/4 = 3/7, \quad x^{k+1} = x^k + \lambda_k s^k = (0, 1/8, 0, 3/4) + \\ &+ 3/7 (0, 17/8, 1, -7/4) = (0, 29/28, 3/7, 0), \quad f(x^{k+1}) = 44/7 > \\ &> 7/2 = f(x^k). \end{aligned}$$

Легко проверить, что план x^{k+1} неоптимален, так что и в точке x^{k+1} существует допустимое прогрессивное направление. Предлагаем продолжить решение задачи до нахождения ее оптимального плана и сравнить полученные результаты с решением этой задачи в п. 4.9.

5°. О сходимости методов. Как уже отмечалось, в настоящей главе рассматриваются только конечные итеративные методы решения задач линейного программирования: в результате вычислений должна быть сформирована такая конечная последовательность

$$x^0, x^1, x^2, \dots, x^q \quad (11)$$

допустимых планов, что либо x^q — решение задачи, либо на q -й итерации устанавливается неограниченность сверху f на $\mathcal{D}$. Только в таком случае будем говорить, что *метод сходится для задачи* $(\mathcal{D}, f)$, и будем говорить, что он *сходится на некотором классе задач*, если он сходится для каждой задачи этого класса.

Если метод сходится на некотором классе задач, то при решении каждой задачи этого класса получается свое число итераций. Для задач линейного программирования оно в первую очередь зависит от размерности задачи (т. е. от числа строк и числа столбцов ее матрицы ограничений). Поэтому можно говорить о скорости сходимости метода как о некоторой его средней характеристике для однотипных задач из данного класса. Обычно такая оценка метода вычисляется экспериментально.

Сходимость методов допустимых направлений может достигаться различными средствами. В методах настоящей главы сходимость обусловлена целенаправленным перебором специальных допустимых планов. В симплекс-методе решения канонической задачи линейного программирования такими планами являются экстремальные точки множества $\mathcal{D}$, в двойственном симплекс-методе — экстремальные точки множества $\mathcal{D}^*$.

Остановимся подробнее на случае канонической задачи линейного программирования. Если $\mathcal{D} \neq \emptyset$, то $\mathcal{D}$ — многогранное выпуклое множество и $\text{ext } \mathcal{D} \neq \emptyset$, так как $\mathcal{D} \subset R^{n+}$, и, следовательно, его конус неограниченности не содержит прямых (см. п. 1.6.3). Будет показано (см. п. 3.8), что метод минимизации невязок позволяет в этом случае найти начальный допустимый план $x^0 \in \text{ext } \mathcal{D}$. Если теперь на каждой итерации метода осуществляется переход от одной экстремальной точки множества $\mathcal{D}$ к другой и для каждого $k = 0, 1, \dots$ выполняется неравенство $cx^k < cx^{k+1}$, то (11) — последовательность попарно различных экстремальных точек множества $\mathcal{D}$. Но многогранное выпуклое множество может иметь только конечное число экстремальных точек, и если задача $(\mathcal{D}, f)$ имеет решение, то среди экстремальных точек есть оптимальный план (см. п. 1.8.2, 7°). Поэтому после некоторого числа итераций будет

найденно решение задачи либо установлена неограниченность f на $\mathcal{D}$.

В рассмотренном примере осуществлен переход от специального плана x^k к специальному плану x^{k+1} : векторы x^k и x^{k+1} — экстремальные точки множества $\mathcal{D}$, так как системы столбцов матрицы ограничений, отвечающих положительным компонентам векторов x^k и x^{k+1} , линейно независимы (см. пример в п. 1.5.5).

2.3. Общая схема методов допустимых направлений для решения задачи линейного программирования с системой ограничений в форме неравенств. В этом пункте рассматриваются особенности применения методов допустимых направлений к экстремальной задаче

$$(\mathcal{D}^*, f^*) : f^*(u) = ub \rightarrow \min, \quad \mathcal{D}^* = \{u \in R^m : uA \geq c\}, \quad (12)$$

т. е. к задаче, двойственной к канонической задаче линейного программирования.

Для нахождения оптимального плана задачи $(\mathcal{D}^*, f^*)$ можно построить ее каноническую форму и для решения полученной канонической задачи линейного программирования использовать схему п. 2.2. Однако при таком подходе значительно возрастает число переменных. Кроме того, и это самое существенное, метод непосредственного решения задачи $(\mathcal{D}^*, f^*)$ можно использовать для решения двойственной к ней задачи $(\mathcal{D}, f)$ (см. § 5). Поэтому анализ задачи $(\mathcal{D}^*, f^*)$ представляет особый интерес.

Конечный метод допустимых направлений для решения задачи $(\mathcal{D}^*, f^*)$ состоит в формировании последовательности

$$u^0, u^1, \dots, u^q \quad (13)$$

ее допустимых планов, основанном на выполнении вычислительных операций 1—3 п. 2.1. Так как общие принципы построения такой итеративной последовательности фактически не отличаются от рассмотренных в предыдущем пункте, то мы остановимся лишь на некоторых технических особенностях.

1°. Нахождение начального плана. Иногда начальный допустимый план задачи $(\mathcal{D}^*, f^*)$ указать легко. Так, если $c \leq 0$, то план $u^0 = (0, \dots, 0) \in \mathcal{D}^*$; если $a_i > 0$ при некотором $i \in 1:m$, то $u^0 = (\underbrace{0, \dots, 0}_{i-1}, \max_{j \in 1:n} c[j]/a[i, j], \underbrace{0, \dots, 0}_{i+1:n}) \in \mathcal{D}^*$.

В общем случае для нахождения плана $u^0 \in \mathcal{D}^*$ формируется вспомогательная задача

$$(\mathcal{G}, g) : g(z) = z[0] \rightarrow \min,$$

$$\mathcal{G} = \{z = (z[0], z[1], \dots, z[m]) \in R^{m+1} : z[0] + \sum_{i=1}^m z[i] a[i, j] \geq c[j], j \in 1:n\}. \quad (14)$$

Справедливы следующие утверждения.

Множество $\mathcal{G} \neq \emptyset$. В самом деле, если $z^0[0] = \max \{c[j], j \in 1:n\}$, то вектор $z^0 = (z^0[0], 0, \dots, 0) \in \mathcal{G}$. ■

Если в множестве $\mathcal{G}$ существует такой план $\hat{z} = (\hat{z}[0], \dots, \hat{z}[m])$, что $\hat{z}[0] \leq 0$, то $u^0 = (\hat{z}[1], \dots, \hat{z}[m]) \in \mathcal{D}^*$. Если $\bar{z}$ — решение задачи $(\mathcal{G}, g)$ и $\bar{z}[0] > 0$, то $\mathcal{D}^* = \emptyset$.

Действительно, если $\hat{z} \in \mathcal{G}$ и $\hat{z}[0] \leq 0$, то

$$u^0 a^j = \sum_{i=1}^m \hat{z}[i] a[i, j] \geq \hat{z}[0] + \sum_{i=1}^m \hat{z}[i] a[i, j] \geq c[j], \quad j \in 1:n,$$

и $u^0 \in \mathcal{D}^*$. Для доказательства второй части утверждения предположим, что $\bar{z}[0] > 0$ и $\mathcal{D}^* \neq \emptyset$. Пусть $u = (u[1], \dots, u[m]) \in \mathcal{D}^*$. Тогда очевидно $z = (0, u[1], \dots, u[m]) \in \mathcal{G}$ и $g(z) = 0 < \bar{z}[0] = g(\bar{z})$, что противоречит оптимальности $\bar{z}$. ■

Доказанные утверждения позволяют предложить следующий метод нахождения начального допустимого плана задачи $(\mathcal{D}^*, f^*)$. Задача $(\mathcal{G}, g)$ решается каким-либо конечным методом допустимых направлений. В качестве начального допустимого плана может быть выбран вектор z^0 . Если на некоторой итерации получается такой план z^k , что $z^k[0] \leq 0$, то вектор $u^0 = (z^k[1], \dots, z^k[m]) \in \mathcal{D}^*$ (заметим, что такой план заведомо будет получен, если g не ограничена снизу на $\mathcal{G}$). Если найден оптимальный план $\bar{z}$ и $\bar{z}[0] > 0$, то $\mathcal{D}^* = \emptyset$.

2°. *Проверка оптимальности.* Для проверки оптимальности или неоптимальности плана $u^k, k = 0, 1, \dots$, воспользуемся критерием п. 3.1.7. Рассмотрим множество индексов $J(u^k) = \{j \in 1:n : u^k a^j = c[j]\}$. Составим следующую систему уравнений и неравенств:

$$Ax = b, \quad (15)$$

$$x[j] = 0, \quad j \in J(u^k), \quad (16)$$

$$x \geq 0. \quad (17)$$

Если система (15)–(17) совместна, то план u^k оптимальный. В противном случае в точке u^k существует допустимое прогрессивное направление.

Проверку совместности системы (15)–(17) осуществить просто, если система уравнений (15), (16) имеет единственное решение: в этом случае достаточно проверить его неотрицательность.

3°. *Поиск допустимого прогрессивного направления.* Нахождение допустимого прогрессивного направления s^k в точке $u^k \in \mathcal{D}^*$ осуществляется в соответствии с условиями допустимости (см. п. 3.1.3) и прогрессивности (см. п. 3.1.4) направления в задаче выпуклого программирования. Для задачи $(\mathcal{D}^*, f^*)$ эти условия имеют вид

$$-a^j s \leq 0, \quad j \in J(u^k), \quad (18)$$

$$bs < 0, \quad (19)$$

где $J(u^k) = \{j \in 1:n : u^k a^j = c[j]\}$. Они получаются из условий (5)–(8) п. 3.1.3 и условия (12') п. 3.1.4 при $I_a = \emptyset$, $I_y = \emptyset$, $J = \emptyset$ и $\nabla f^*(u) = b$. Если u^k не является решением задачи $(\mathcal{D}^*, f^*)$, то система (18), (19) совместна и каждое ее решение является допустимым прогрессивным направлением в точке u^k .

Мы не будем давать универсальных рекомендаций по поводу эффективного выбора s^k , удовлетворяющего условиям (18), (19): здесь можно было бы лишь повторить сказанное в п. 2.2, 3°.

4°. **Выбор шага.** Если s^k — допустимое прогрессивное направление в точке u^k , то существует такое $\lambda' > 0$, что

$$u^k + \lambda s^k \in \mathcal{D}^*, \quad (20)$$

$$f^*(u^k + \lambda s^k) < f^*(u^k) \quad (21)$$

при любом $\lambda \in (0, \lambda']$. Так как неравенство (21) выполняется при любом $\lambda > 0$ и $\lim_{\lambda \rightarrow \infty} f^*(u^k + \lambda s^k) = -\infty$ при $\lambda \rightarrow \infty$ (см. аналогичное утверждение в п. 2.2), то λ_k следует положить равным $\max \{\lambda : u^k + \lambda s^k \in \mathcal{D}^*, \lambda > 0\}$.

Из определения множества $\mathcal{D}^*$ следует, что вектор $u^k + \lambda s^k$ в том и только том случае принадлежит $\mathcal{D}^*$, если

$$(u^k + \lambda s^k) A \geq c. \quad (22)$$

Положим $a(u^k) = u^k A - c$, $\delta(s^k) = s^k A$. Тогда неравенство (22) равносильно неравенству

$$a(u^k) + \lambda \delta(s^k) \geq 0. \quad (23)$$

Так как $u^k \in \mathcal{D}^*$ и s^k — допустимое прогрессивное направление в точке u^k , то $a(u^k) \geq 0$ и $a(u^k)[j] = 0$, $\delta(s^k)[j] \geq 0$ для $j \in J(u^k)$. Если при этом $\delta(s^k)[j] \geq 0$, $j \in 1:n \setminus J(u^k)$, то неравенство (23) выполняется при любом $\lambda > 0$ (вектор s^k принадлежит конусу неограниченности множества $\mathcal{D}^*$), и поскольку $f^*(u^k + \lambda s^k)$ — убывающая функция от λ , то целевая функция задачи $(\mathcal{D}^*, f^*)$ не ограничена снизу на $\mathcal{D}^*$ (см. п. 1.8.2, 6°).

Пусть теперь найдется такое $j \in 1:n \setminus J(u^k)$, что $\delta(s^k)[j] < 0$. Для такого j неравенство $a(u^k)[j] + \lambda \delta(s^k)[j] \geq 0$ нарушается при $\lambda > a(u^k)[j] / -\delta(s^k)[j]$. Поэтому λ_k следует положить равным наибольшему из чисел λ , для которых выполняются все неравенства $a(u^k)[j] + \lambda \delta(s^k)[j] \geq 0$ при $j \in 1:n \setminus J(u^k)$, $\delta(s^k)[j] < 0$, т. е.

$$\lambda_k = \min_{j: \delta(s^k)[j] < 0} a(u^k)[j] / -\delta(s^k)[j]. \quad (24)$$

Итак, если найдется такое $j \in 1:n \setminus J(u^k)$, что $\delta(s^k)[j] < 0$, то по формуле (24) можно вычислить шаг λ_k и положить $u^{k+1} = u^k + \lambda_k s^k$.

5°. Сходимость методов допустимых направлений обсуждалась в п. 2.2, 5°.

§ 3. Симплекс-метод

3.1. Основные обозначения и определения. Рассмотрим каноническую задачу линейного программирования

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \quad \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\}, \quad (1)$$

где $A = (a[i, j]) - m \times n$ матрица, ранг которой $r(A) = m$ ($m \leq n$), $c = (c[1], c[2], \dots, c[n])$, $x^T = (x[1], x[2], \dots, x[n])$, $b^T = (b[1], b[2], \dots, b[m])$. Как и раньше, $a^j - j$ -й столбец матрицы A , $a_i - i$ -я строка. Не умаляя общности, будем считать, что $b \geq 0$.

Введем следующие обозначения. Через $\bar{A}$ обозначим матрицу, которая получается присоединением к матрице A сверху строки $a_0 = c$; j -й столбец матрицы $\bar{A}$ обозначим через $\bar{a}^j$. Через $\bar{b}$ обозначим вектор из R^{m+1} , который получается из вектора b присоединением к нему сверху компоненты $b[0] = 0$, т. е. $\bar{b}^T = (0, b[1], b[2], \dots, b[m])$.

Все вводимые в данном параграфе понятия и конструкции будем иллюстрировать примерами. Для этого рассмотрим задачу линейного программирования *

$$(\mathcal{D}, f): f(x) = -x_1 - 2x_2 + x_3 - 3x_4 + 4x_5 \rightarrow \max, \quad \mathcal{D} = \{x \in R^5:$$

$$3x_1 + x_3 + x_4 - x_5 = 3,$$

$$x_1 + x_2 + 4x_3 + 3x_5 = 2,$$

$$x_j \geq 0, \quad j \in 1:5\}.$$

Матрица ограничений этой задачи

$$A = \begin{pmatrix} 3 & 0 & 1 & 1 & -1 \\ 1 & 1 & 4 & 0 & 3 \end{pmatrix},$$

вектор коэффициентов целевой функции $c = (-1, -2, 1, -3, 4)$, вектор ограничений $b^T = (3, 2)$, матрица

$$\bar{A} = \begin{pmatrix} -1 & -2 & 1 & -3 & 4 \\ 3 & 0 & 1 & 1 & -1 \\ 1 & 1 & 4 & 0 & 3 \end{pmatrix}.$$

вектор $\bar{b} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$, и, например, вектор $\bar{a}^4 = \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}$.

* Напоминаем, что в примерах для переменных мы используем обозначения x_j , а не $x[j]$.

Систему из m линейно независимых столбцов матрицы A назовем *базисом задачи* (1). Базис задачи часто будем называть просто *базисом*.

Произвольному базису $\mathcal{B} = \{a^{j_1}, a^{j_2}, \dots, a^{j_m}\}$ ставится в соответствие ряд объектов 1)–7), необходимых для описания алгоритмов настоящего и следующего параграфов.

1) Множество $N(\mathcal{B}) = \{j_1, j_2, \dots, j_m\}$ номеров столбцов, входящих в $\mathcal{B}$. В таблицах, которыми иллюстрируются вычислительные процедуры, $N(\mathcal{B})$ будет записываться в виде вектор-столбца, i -я компонента которого равна j_i .

Например, система столбцов $\mathcal{B}_1 = \{a^4, a^2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$.

$\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ очевидно является базисом рассматриваемой задачи.

Здесь $j_1 = 4$, $j_2 = 2$ и $N(\mathcal{B}_1) = \{4, 2\}$. Нетрудно убедиться, что базисом задачи $(\mathcal{D}, f)$ является также система столбцов $\mathcal{B}_2 = \{a^1, a^5\} = \left\{ \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\}$, здесь уже $j_1 = 1$, $j_2 = 5$, $N(\mathcal{B}_2) = \{1, 5\}$.

2) Матрица $B = \{a^{j_1} a^{j_2} \dots a^{j_m}\}$.

Базису $\mathcal{B}_1$ отвечает матрица $B_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Аналогично

$$B_2 = \begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix}.$$

3) Вектор-строка $c(\mathcal{B}) = (c[j_1], c[j_2], \dots, c[j_m])$.

Для базиса $\mathcal{B}_1$ вектор $c(\mathcal{B}_1) = (-3, -2)$, точно так же $c(\mathcal{B}_2) = (-1, 4)$.

4) Матрица $\bar{B} = (-e^1 \bar{a}^{j_1} \bar{a}^{j_2} \dots \bar{a}^{j_m})$, где $e^1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in R^{m+1}$. Оче-

видно, что $\bar{B}$ — невырожденная матрица и

$$\bar{B}^{-1} = \left(-e^1 \left| \begin{array}{c} u(\mathcal{B}) \\ B^{-1} \end{array} \right. \right) = \left(\begin{array}{c|c} -1 & u(\mathcal{B}) \\ \hline 0 & B^{-1} \end{array} \right), \quad (2)$$

где

$$u(\mathcal{B}) = c(\mathcal{B}) B^{-1}. \quad (3)$$

— Упражнение 4.2.

Строку матрицы $\bar{B}^{-1}$ с номером k обозначим через $d_k(\mathcal{B})$, $k \in 0:m$; $d(\mathcal{B})[k, j]$ — j -й элемент этой строки. Из (2) и (3) следует, что

$$d_0(\mathcal{B}) = (-1, c(\mathcal{B}) B^{-1}). \quad (4)$$

Базисам $\mathcal{B}_1$ и $\mathcal{B}_2$ отвечают матрицы

$$\bar{B}_1^{-1} = \begin{pmatrix} -1 & -3 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \bar{B}_1^{-1} = \begin{pmatrix} -1 & -7/10 & 11/10 \\ 0 & 3/10 & 1/10 \\ 0 & -1/10 & 3/10 \end{pmatrix},$$

$$B_1^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B_2^{-1} = \begin{pmatrix} 3/10 & 1/10 \\ -1/10 & 3/10 \end{pmatrix}, \quad u(\mathcal{B}_1) = (-3, -2),$$

$$d_0(\mathcal{B}_1) = (-1, -3, -2), \quad d_0(\mathcal{B}_2) = (-1, -7/10, 11/10).$$

5) Матрица $\bar{A}(\mathcal{B}) = \bar{B}^{-1}\bar{A}$, элементы которой обозначим через $a(\mathcal{B})[i, j]$ ($i \in 0:m; j \in 1:n$), а строки и столбцы соответственно — через $a_i(\mathcal{B})$ и $\bar{a}^j(\mathcal{B})$.

Матрица $A(\mathcal{B}) = B^{-1}A$ состоит из последних m строк матрицы $\bar{A}(\mathcal{B})$. Столбцы матрицы $A(\mathcal{B})$ обозначим через $a^j(\mathcal{B})$ ($j \in 1:n$). Очевидно, что вектор $a^j(\mathcal{B})$ есть последовательность коэффициентов разложения вектора a^j по базису $\mathcal{B}$, а вектор $\bar{a}^j(\mathcal{B})$ есть последовательность коэффициентов разложения вектора $\bar{a}^j$ по базису $\{-e^1\} \cup \mathcal{B}$ пространства R^{m+1} , где $\bar{\mathcal{B}} = \{\bar{a}^1, \bar{a}^2, \dots, \bar{a}^m\}$. — Упражнение 4.3. Отметим также, что $a_0(\mathcal{B}) = d_0(\mathcal{B})\bar{A}$.

В рассматриваемом примере

$$a_0(\mathcal{B}_1) = (-10, 0, -12, 0, -7),$$

$$A(\mathcal{B}_1) = \begin{pmatrix} 3 & 0 & 1 & 1 & -1 \\ 1 & 1 & 4 & 0 & 3 \end{pmatrix},$$

$$\bar{A}(\mathcal{B}_1) = \begin{pmatrix} -10 & 0 & -12 & 0 & -7 \\ 3 & 0 & 1 & 1 & -1 \\ 1 & 1 & 4 & 0 & 3 \end{pmatrix}$$

и соответственно

$$a_0(\mathcal{B}_2) = (0, 31/10, 27/10, 23/10, 0),$$

$$A(\mathcal{B}_2) = \begin{pmatrix} 1 & 1/10 & 7/10 & 3/10 & 0 \\ 0 & 3/10 & 11/10 & -1/10 & 1 \end{pmatrix}, \quad \bar{A}(\mathcal{B}_2) =$$

$$= \begin{pmatrix} 0 & 31/10 & 27/10 & 23/10 & 0 \\ 1 & 1/10 & 7/10 & 3/10 & 0 \\ 0 & 3/10 & 11/10 & -1/10 & 1 \end{pmatrix}.$$

6) Вектор-столбец $\bar{b}(\mathcal{B}) = \bar{B}^{-1}\bar{b}$, i -я компонента которого обозначается через $b(\mathcal{B})[i]$, $i \in 0:m$, и вектор $b(\mathcal{B}) = B^{-1}b$.

Очевидно, что

$$\begin{aligned}\bar{b}^*(\mathcal{B}) &= (b(\mathcal{B})[0], b(\mathcal{B})[1], \dots, b(\mathcal{B})[m]) = \\ &= (b(\mathcal{B})[0], b^*(\mathcal{B})) = (d_0(\mathcal{B})b, b^*(\mathcal{B})).\end{aligned}$$

Базису $\mathcal{B}_1$ соответствует вектор $\bar{b}^*(\mathcal{B}_1) = (-13, 3, 2)$, где $b(\mathcal{B}_1)[0] = -13$, а $b^*(\mathcal{B}_1) = (3, 2)$. Аналогично

$$\bar{b}^*(\mathcal{B}_2) = (1/10, 11/10, 3/10), \quad b(\mathcal{B}_2)[0] = 1/10, \quad b^*(\mathcal{B}_2) = (11/10, 3/10).$$

7) План $x(\mathcal{B}) = \{x(\mathcal{B})[j] = b(\mathcal{B})[i], i \in 1:m; x(\mathcal{B})[j] = 0, j \in N(\mathcal{B})\}$, который назовем *базисным*. Точнее, вектор $x \in R^n$ назовем базисным планом, если существует такой базис $\mathcal{B}$ задачи (1), что $x = x(\mathcal{B})$. Компоненты $x(\mathcal{B})[j]$ этого плана с $j \in N(\mathcal{B})$ назовем *базисными*, а остальные — *небазисными*. Очевидно, что $Ax(\mathcal{B}) = b$ и базисный план является допустимым в том и только том случае, если $b(\mathcal{B}) \geq 0$. — Упражнение 4.4.

Базис $\mathcal{B}$ задачи называется *допустимым*, если ему отвечает допустимый базисный план, и *оптимальным*, если $x(\mathcal{B})$ — решение задачи (1).

Допустимый базисный план $x(\mathcal{B})$ называется *невырожденным*, если $b(\mathcal{B}) > 0$, и *вырожденным*, если последнее неравенство не выполняется. Задача (1) называется *невырожденной*, если в множестве $\mathcal{D}$ нет вырожденных базисных планов (см. п. 1.4).

По базису $\mathcal{B}_1$ можно сформировать базисный план $x(\mathcal{B}_1) = (0, 2, 0, 3, 0)$, его базисные компоненты: $x(\mathcal{B}_1)[2] = b(\mathcal{B}_1)[2] = 2$, $x(\mathcal{B}_1)[4] = b(\mathcal{B}_1)[1] = 3$; небазисные компоненты: $x(\mathcal{B}_1)[1] = x(\mathcal{B}_1)[3] = x(\mathcal{B}_1)[5] = 0$. Поскольку $b^*(\mathcal{B}_1) = (3, 2) > 0$, то план $x(\mathcal{B}_1)$ является допустимым и невырожденным, а базис $\mathcal{B}_1$ — допустимым.

Аналогично базисный план $x(\mathcal{B}_2) = (11/10, 0, 0, 0, 3/10)$ является допустимым и невырожденным. Нетрудно проверить, что вектор $x(\mathcal{B}_2)$ — решение рассматриваемой задачи, т. е. допустимый базис $\mathcal{B}_2$ является и оптимальным.

3.2. Предварительные замечания. Симплекс-метод позволяет за конечное число q итераций найти оптимальный план невырожденной задачи (1) или обнаружить, что в задаче $(\mathcal{D}, f)$ функция f не ограничена сверху на $\mathcal{D}$. Число итераций q не известно до окончания вычислений. Оно зависит от выбора начального допустимого базиса $\mathcal{B}_1$. Выбор такого базиса — отдельная задача. Методы ее решения приведены в п. 3.8 (см. также п. 2.2, 1°).

Название метода связано с одной из его геометрических интерпретаций (см. п. 3.3.6). Каждому базису $\mathcal{B} = \{a^{i_1}, a^{i_2}, \dots, a^{i_m}\}$ в пространстве R^{m+1} отвечают $(m-1)$ -мерный симплекс $C_{\mathcal{B}} = \{ \bar{a}^{i_1}, \bar{a}^{i_2}, \dots, \bar{a}^{i_m} \}$ и многогранный выпуклый конус $K_{\mathcal{B}}$ с

вершиной в нуле, натянутый на симплекс $C_{\mathcal{B}}$. Вектор $d_0(\mathcal{B})$ (см. (4)) является нормальным вектором m -мерной гиперплоскости, содержащей конус $K_{\mathcal{B}}$. Информация о взаимном расположении в R^{m+1} нормали $d_0(\mathcal{B})$ и векторов $\bar{a}^j$, $j \in 1:n$, позволяет сделать вывод об оптимальности плана $x(\mathcal{B})$ или в случае неоптимальности $x(\mathcal{B})$ выбрать направление перехода от плана $x(\mathcal{B})$ к плану $x(\mathcal{B}')$. При этом в невырожденной задаче (1) всегда выполняется неравенство $f(x(\mathcal{B}')) > f(x(\mathcal{B}))$.

В процессе вычислений будут построены последовательность допустимых базисов $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_q$ и последовательность таблиц

$$T^1(\mathcal{B}_1), T^1(\mathcal{B}_2), \dots, T^1(\mathcal{B}_q) \quad (5)$$

вида

$$T^1(\mathcal{B}) = \left[\begin{array}{c|c|c} N(\mathcal{B}) & \bar{b}(\mathcal{B}) & \bar{A}(\mathcal{B}) \end{array} \right], \quad (6)$$

отвечающая последовательности допустимых базисов. Базисы $\mathcal{B}_2, \dots, \mathcal{B}_q$ будут строиться вместе с соответствующими таблицами.

3.3. Алгоритм. Этап 0. Вычисления начинаем с допустимого базиса $\mathcal{B}_1$. Полагаем $\mathcal{B}$ равным $\mathcal{B}_1$. По базису $\mathcal{B}$ формируется табл. $T^1(\mathcal{B})$ (см. (6)). Переходим к этапу 1.

Этап 1. Опишем одну итерацию вычислительного процесса. Она состоит из проверки оптимальности «текущего» базисного плана $x(\mathcal{B})$ и последующего перехода к новому базису $\mathcal{B}'$ и плану $x(\mathcal{B}')$, если $x(\mathcal{B})$ не является решением задачи (1).

1. В табл. $T^1(\mathcal{B})$ рассматриваются строка $a_0(\mathcal{B})$ — верхняя строка матрицы $\bar{A}(\mathcal{B})$ и множество индексов $J^-(\mathcal{B}) = \{j \in 1:n : a(\mathcal{B})[0, j] < 0\}$. Возможны два случая.

1'. Множество $J^-(\mathcal{B}) = \emptyset$, т. е. $a_0(\mathcal{B}) \geq 0$. Вычислительный процесс закончен. Выписываем оптимальный план $x(\mathcal{B})$: в столбце $N(\mathcal{B})$ табл. $T^1(\mathcal{B})$ расположены номера базисных компонент этого плана; рядом — в столбце $\bar{b}(\mathcal{B})$ — их значения. Остальные $n - m$ компонент плана $x(\mathcal{B})$ равны нулю. Максимальное значение $f(x(\mathcal{B}))$ функции f равно $b(\mathcal{B})[0]$.

1''. Множество $J^-(\mathcal{B}) \neq \emptyset$. Переходим к рекомендации 2.

2. Выбираем некоторое l из $J^-(\mathcal{B})$.

3. Рассматривается вектор $a^l(\mathcal{B})$. Возможны два случая.

3'. $a^l(\mathcal{B}) \leq 0$. Вычислительный процесс закончен. Функция f не ограничена сверху на $\mathcal{D}$.

3''. Существует такое $i \neq 0$, что $a(\mathcal{B})[i, l] > 0$. Переходим к рекомендации 4.

4. Вычисляются такие $\lambda(\mathcal{B})$ и $r \in 1:m$, что

$$\lambda(\mathcal{B}) = \min_{\substack{i: a(\mathcal{B})[i, l] > 0 \\ i \neq 0}} \frac{b(\mathcal{B})[i]}{a(\mathcal{B})[i, l]} = \frac{b(\mathcal{B})[r]}{a(\mathcal{B})[r, l]}. \quad (7)$$

Элемент $a(\mathcal{B})[r, l]$ назовем *ведущим*. Будет показано, что в невырожденной задаче (1) формула (7) определяет его однозначно. Столбец a^{lr} выводится из базиса $\mathcal{B}$. Формируется новый базис $\mathcal{B}' = \{a^{l_1}, \dots, a^{l_{r-1}}, a^l, a^{l_{r+1}}, \dots, a^{l_m}\}$ и $N(\mathcal{B}') = \{j_1, \dots, j_{r-1}, l, j_{r+1}, \dots, j_m\}$.

5. Заполняется табл. $T^1(\mathcal{B}')$. В нее записывается столбец $N(\mathcal{B}')$. Остальные элементы таблицы вычисляются по формулам:

$$a(\mathcal{B}')[r, j] = a(\mathcal{B})[r, j] / a(\mathcal{B})[r, l], \quad j \in 1:n, \quad (8)$$

$$b(\mathcal{B}')[r] = b(\mathcal{B})[r] / a(\mathcal{B})[r, l], \quad (9)$$

$$a(\mathcal{B}')[i, j] = a(\mathcal{B})[i, j] - a(\mathcal{B})[i, l] a(\mathcal{B}')[r, j], \quad i \in 0:m, \quad i \neq r, \quad j \in 1:n, \quad (10)$$

$$b(\mathcal{B}')[i] = b(\mathcal{B})[i] - a(\mathcal{B})[i, l] b(\mathcal{B}')[r], \quad i \in 0:m, \quad i \neq r. \quad (11)$$

6. Полагаем $\mathcal{B} = \mathcal{B}'$. Переходим к рекомендации 1. ■

З а м е ч а н и е. Вычисления, связанные с заполнением табл. $T^1(\mathcal{B}')$, можно проводить в следующем порядке. Табл. $T^1(\mathcal{B}')$ удобно располагать под табл. $T^1(\mathcal{B})$ (см. пример в п. 3.5). Прежде всего заполняется столбец $N(\mathcal{B}')$: в него последовательно переписываются все элементы столбца $N(\mathcal{B})$, кроме r -го, на r -е место ставится индекс l . Затем заполняется r -я строка: в соответствии с формулами (8), (9) ее элементы получаются делением элементов r -й строки табл. $T^1(\mathcal{B})$ на ведущий элемент. Полученную строку назовем *ведущей строкой*, а l -й столбец табл. $T^1(\mathcal{B})$ — *ведущим столбцом*. После этого вычисляются остальные элементы табл. $T^1(\mathcal{B}')$: элемент $a(\mathcal{B}')[i, j]$ (или $b(\mathcal{B}')[i]$) при $i \neq r$ получается (см. (10), (11)) вычитанием из $a(\mathcal{B})[i, j]$ (из $b(\mathcal{B})[i]$) произведения элемента ведущего столбца, стоящего в той же строке i , на элемент ведущей строки, стоящий в том же столбце j (или в столбце $b(\mathcal{B}')$, если вычисляется $b(\mathcal{B}')[i]$). Отметим также, что после заполнения ведущей строки сразу следует заполнить верхнюю строку табл. $T^1(\mathcal{B}')$ и проверить, есть ли в ней отрицательные элементы. Если таковых нет, то $\mathcal{B}'$ — оптимальный базис; тогда заполняется только столбец $b(\mathcal{B}')$, выписывается оптимальный план $x(\mathcal{B}')$, и $f(x(\mathcal{B}')) = b(\mathcal{B}')[0]$. Если же в строке $a_0(\mathcal{B}')$ есть отрицательные элементы, то заполняется вся табл. $T^1(\mathcal{B}')$. Полезно также иметь в виду, что

$$(\bar{a}^l(\mathcal{B}'))^r = (\underbrace{0, \dots, 0}_r, 1, 0, \dots, 0) \in R^{m+1}.$$

3.4. Обсуждение. Рассмотрим следующие две проблемы. На каждой итерации множество индексов $J^-(\mathcal{B})$ может состоять из более чем одного элемента. Тогда в силу произвольности выбора индекса l (см. рекомендацию 2 в этапе 1) появляется возмож-

ность формировать разные последовательности допустимых базисов и таблиц (5) и возникает проблема сокращения числа итераций q за счет более рационального выбора l .

Существующие на этот счет рекомендации, как правило, основаны на практике вычислений. Обычно рекомендуется выбирать такое l , чтобы

$$a(\mathcal{B})[0, l] = \min_{l \in J^-(\mathcal{B})} a(\mathcal{B})[0, l] = \gamma.$$

Предлагаемый выбор l может существенно сократить число итераций, необходимых для решения задачи. Однако и такая рекомендация может не обеспечить однозначность выбора l , если множество $S = \{s \in 1:n: a(\mathcal{B})[0, s] = \gamma\}$ состоит из более чем одного элемента. В этом случае для определенности можно положить l равным наименьшему или наибольшему из номеров $s \in S$. Может оказаться эффективным и случайный выбор l из S (или из $J^-(\mathcal{B})$) на каждой итерации симплекс-метода.

Отметим также, что если существует по крайней мере одно такое j из $J^-(\mathcal{B})$, для которого $a^l(\mathcal{B}) \leq 0$, то в соответствии с рекомендацией 3 этапа 1 вычисления нужно прекратить: целевая функция f не ограничена сверху на $\mathcal{D}$. Поэтому такую проверку также следует включать в вычислительную процедуру.

Вторая проблема связана с решением вырожденных задач. В алгоритме п. 3.3 отмечено, что выбор ведущего элемента по формуле (7) в невырожденных задачах осуществляется однозначно. В случае вырожденной задачи однозначность может не иметь места, и в частности $\lambda(\mathcal{B})$ может быть равным нулю. Тем не менее и вырожденные задачи мы предлагаем решать, используя алгоритм п. 3.3, а в случае неоднозначного выбора r по формуле (7) полагать r равным, например, наименьшему из номеров p , для которых $b(\mathcal{B})[p]/a(\mathcal{B})[p, l] = \lambda(\mathcal{B})$. Возникающие при этом теоретические осложнения обсуждаются в § 7. Однако при решении подавляющего большинства вырожденных задач такая простая рекомендация позволяет решить задачу. Эффективным является и случайный выбор r (подробнее о вырожденности см. в § 7).

3.5. Пример. В табл. (12), состоящей из табл. $T^1(\mathcal{B}_1)$, $T^1(\mathcal{B}_2)$, $T^1(\mathcal{B}_3)$ и $T^1(\mathcal{B}_4)$, приведено решение задачи

$$\begin{aligned} (\mathcal{D}, f): f(x) &= -x_1 - 2x_2 + x_3 - 3x_4 + 4x_5 \rightarrow \max, \\ \mathcal{D} &= \{x \in R^5: 3x_1 + x_3 + x_4 - x_5 = 3, \\ &\quad x_1 + x_2 + 4x_3 + 3x_5 = 2, \\ &\quad x_j \geq 0, j \in 1:5\}, \end{aligned}$$

которую мы уже рассматривали в п. 3.1. В таблицах отмечены ведущие элементы и выделены ведущие столбцы и строки. Рассмотрим подробнее этап 0 и первую итерацию.

$$\begin{aligned}
 T^1(\mathcal{B}_1) &= \begin{array}{c|ccc|ccc} & -13 & -10 & 0 & -12 & 0 & -7 \\ \hline 4 & 3 & 3 & 0 & 1 & 1 & -1 \\ 2 & 2 & 1 & 1 & 4^* & 0 & 3 \end{array} \\
 T^1(\mathcal{B}_2) &= \begin{array}{c|cc|cc|cc} & -7 & -7 & 3 & 0 & 0 & 2 \\ \hline 4 & 5/2 & 11/4^* & -1/4 & 0 & 1 & -7/4 \\ 3 & 1/2 & 1/4 & 1/4 & 1 & 0 & 3/4 \end{array} \\
 T^1(\mathcal{B}_3) &= \begin{array}{c|cc|cc|cc} & -7/11 & 0 & 26/11 & 0 & 28/11 & -27/11 \\ \hline 1 & 10/11 & 1 & -1/11 & 0 & 4/11 & -7/11 \\ 3 & 3/11 & 0 & 3/11 & 1 & -1/11 & 10/11^* \end{array} \\
 T^1(\mathcal{B}_4) &= \begin{array}{c|cc|cc|cc} & 1/10 & 0 & 31/10 & 27/10 & 22/10 & 0 \\ \hline 1 & 11/10 & 1 & 1/10 & 7/10 & 3/10 & 0 \\ 5 & 3/10 & 0 & 3/10 & 11/10 & -1/10 & 1 \end{array}
 \end{aligned} \tag{12}$$

Решение начато с базиса $\mathcal{B}_1$, для которого $N(\mathcal{B}_1) = \{4, 2\}$. Как было отмечено в п. 3.1,

$$B_1 = B_1^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad c(\mathcal{B}_1) = (-3, -2),$$

$$\bar{B}_1^{-1} = \left(\begin{array}{c|cc} -1 & -3 & -2 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) = \begin{pmatrix} -1 & -3 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\bar{A} = \begin{pmatrix} c \\ A \end{pmatrix} = \begin{pmatrix} -1 & -2 & 1 & -3 & 4 \\ 3 & 0 & 1 & 1 & -1 \\ 1 & 1 & 4 & 0 & 3 \end{pmatrix},$$

$$\bar{A}(\mathcal{B}_1) = \bar{B}_1^{-1} \bar{A} = \begin{pmatrix} -10 & 0 & -12 & 0 & -7 \\ 3 & 0 & 1 & 1 & -1 \\ 1 & 1 & 4 & 0 & 3 \end{pmatrix},$$

$$\bar{b} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}, \quad \bar{b}(\mathcal{B}_1) = \bar{B}_1^{-1} \bar{b} = \begin{pmatrix} -13 \\ 3 \\ 2 \end{pmatrix}, \text{ и можно заполнить табл. (6):}$$

$$T^1(\mathcal{B}_1) = \begin{array}{c|ccccc|c} & -13 & -10 & 0 & -12 & 0 & -7 \\ \hline 4 & 3 & 3 & 0 & 1 & 1 & -1 \\ 2 & 2 & 1 & 1 & 4^* & 0 & 3 \end{array}.$$

Рассмотрим верхнюю строку $a_0(\mathcal{B}_1) = (-10, 0, -12, 0, 7)$ матрицы $\bar{A}(\mathcal{B}_1)$. Здесь $J^-(\mathcal{B}_1) = \{1, 3, 5\}$. Поскольку $J^-(\mathcal{B}_1) \neq \emptyset$, то базисный план $x(\mathcal{B}_1) = (0, 2, 0, 3, 0)$ неоптимален. В качестве l можно взять любой элемент из множества $J^-(\mathcal{B}_1)$. В табл. (12) на первой итерации $l = 3$, поскольку $\min_{j \in J^-(\mathcal{B}_1)} a(\mathcal{B}_1)[0, j] = -12 = a(\mathcal{B}_1)[0, 3]$. Этого правила выбора l мы придерживаемся на всех итерациях. Предлагаем читателю самостоятельно произвести дальнейшие вычисления при $l = 1$ или $l = 5$.

Вектор $a^3(\mathcal{B}_1) = \begin{pmatrix} 1 \\ 4 \end{pmatrix} > 0$, поэтому вычисляется $\lambda(\mathcal{B}_1) = \min \{3/1, 2/4\} = 1/2 = b(\mathcal{B}_1)[2]/a(\mathcal{B}_1)[2, 3]$, так что $r = 2$. Элемент $a(\mathcal{B}_1)[2, 3] = 4$ является ведущим (он отмечен), столбец $\bar{a}^3(\mathcal{B}_1) = \begin{pmatrix} -12 \\ 1 \\ 4 \end{pmatrix}$ — ведущий, он выделен.

Табл. $T^1(\mathcal{B}_2)$ заполняется по правилам, указанным в рекомендации 5 этапа 1 (см. замечание к п. 3.3). Поскольку $l = 3$, а $r = 2$, то из базиса $\mathcal{B}_1$ выводится вектор a^2 и вместо него вводится вектор a^3 , т. е. $N(\mathcal{B}_2) = \{4, 3\}$. По формулам (8), (9) заполняется ведущая строка (она выделена в табл. $T^1(\mathcal{B}_2)$), остальные элементы табл. $T^1(\mathcal{B}_2)$ вычислены по формулам (10), (11). На этом заканчивается первая итерация. Осуществлен переход к новому базису $\mathcal{B}_2$; базисный план $x(\mathcal{B}_2) = (0, 0, 1/2, 5/2, 0)$, $f(x(\mathcal{B}_2)) = -7 > -13 = f(x(\mathcal{B}_1))$.

Все последующие вычисления производятся аналогично, их результаты приведены в табл. $T^1(\mathcal{B}_3)$ и $T^1(\mathcal{B}_4)$. Из табл. $T^1(\mathcal{B}_4)$ можно выписать оптимальный план $x(\mathcal{B}_4) = (11/10, 0, 0, 0, 3/10)$ и максимальное значение целевой функции $f(x(\mathcal{B}_4)) = 1/10$, $q = 4$.

3.6. Обоснование алгоритма. Приведенное в п. 3.3 описание алгоритма симплекс-метода не позволяет сразу увидеть идеи, на которых основана вычислительная процедура. Покажем, и это будет служить обоснованием алгоритма, что он является специальной реализацией метода допустимых направлений. Последовательность (11) п. 2.2 в рассматриваемом случае есть последовательность $x(\mathcal{B}_1), \dots, x(\mathcal{B}_q)$ экстремальных точек (вершин)

множества $\mathcal{D}$. Переход от $x(\mathcal{B}_k)$ к $x(\mathcal{B}_{k+1})$ осуществляется в допустимом прогрессивном направлении.

Перейдем к точным формулировкам. Утверждения 1°—12°, формулируемые в данном пункте, доказываются в п. 3.7.

Пусть $\mathcal{B}$ — допустимый базис задачи $(\mathcal{D}, f)$. Тогда допустимый базисный план $x(\mathcal{B})$ удовлетворяет системе линейных уравнений $Ax = b$; $x[j] = 0$, $j \in 1:n \setminus N(\mathcal{B})$ (см. п. 3.1, 7) и, следовательно, является экстремальной точкой (вершиной) многогранного выпуклого множества $\mathcal{D}$ (см. пример в п. 1.5.5). Для каждого $l \in 1:n \setminus N(\mathcal{B})$ определим вектор $s^l(\mathcal{B}) \in R^n$. Его компоненты:

$$\begin{aligned} s^l(\mathcal{B})[j] &= -a(\mathcal{B})[l, j], \quad j \in 1:m, \\ s^l(\mathcal{B})[l] &= 1; \quad s^l(\mathcal{B})[j] = 0, \quad j \in 1:n \setminus (N(\mathcal{B}) \cup \{l\}). \end{aligned} \quad (13)$$

Справедливы следующие утверждения.

1°. Если задача $(\mathcal{D}, f)$ невырождена, то для любого $l \in 1:n \setminus N(\mathcal{B})$ вектор $s^l(\mathcal{B})$ — допустимое направление в точке $x(\mathcal{B})$.

2°. Если $a^l(\mathcal{B}) \leq 0$, то $s^l(\mathcal{B})$ — допустимое направление в точке $x(\mathcal{B})$ и $x(\mathcal{B}) + \lambda s^l(\mathcal{B}) \in \mathcal{D}$ при любом $\lambda \geq 0$.

Для $l \in 1:n \setminus N(\mathcal{B})$ положим

$$H_l = \{x \in R^n : Ax = b, x[j] = 0, j \in 1:n \setminus (N(\mathcal{B}) \cup \{l\})\}.$$

Множество H_l — аффинное множество размерности 1, при этом $H_l = \{x \in R^n : x = x(\mathcal{B}) + \lambda s^l(\mathcal{B}), \lambda \in R\}$. — Упражнение 4.5.

Положим $\mathcal{D}_l = \mathcal{D} \cap H_l$. Так как $x(\mathcal{B}) \in \mathcal{D}_l$, то $\mathcal{D}_l$ — грань (см. п. 1.7.5) многогранного выпуклого множества $\mathcal{D}$ и $\dim \mathcal{D}_l \leq 1$. Если $\dim \mathcal{D}_l = 1$, то $\mathcal{D}_l$ — ребро. В силу 1° для невырожденной задачи $\dim \mathcal{D}_l = 1$ при любом $l \in 1:n \setminus N(\mathcal{B})$.

3°. Для любого $l \in 1:n \setminus N(\mathcal{B})$ справедливо равенство $cs^l(\mathcal{B}) = -a(\mathcal{B})[0, l]$, и вектор $s^l(\mathcal{B})$ в том и только том случае является прогрессивным направлением в точке $x(\mathcal{B})$, если $a(\mathcal{B})[0, l] < 0$.

4°. Если $a_0(\mathcal{B}) \geq 0$, то $x(\mathcal{B})$ — решение задачи $(\mathcal{D}, f)$.

Из утверждения 4° следует, что для оптимальности допустимого базисного плана $x(\mathcal{B})$ достаточно (без предположения о невырожденности задачи $(\mathcal{D}, f)$), чтобы среди векторов $s^l(\mathcal{B})$, $l \in 1:n \setminus N(\mathcal{B})$, не было прогрессивных в точке $x(\mathcal{B})$ направлений. Это обстоятельство вместе с утверждением 5° служит обоснованием рекомендации 1' этапа 1.

5°. $f(x(\mathcal{B})) = b(\mathcal{B})[0]$.

Для невырожденной задачи $(\mathcal{D}, f)$ в силу 4° условие $a_0(\mathcal{B}) \geq 0$ очевидно является и необходимым (см. п. 3.1.2). Предлагаем читателю самостоятельно построить пример вырожденной задачи, для которой 4° не было бы необходимым условием оптимальности. — Упражнение 4.6.

Пусть $l \in 1:n \setminus N(\mathcal{B})$. Из 2° следует, что при $a^l(\mathcal{B}) \leq 0$ вектор $s^l(\mathcal{B})$ принадлежит конусу неограниченности многогранного выпуклого множества $\mathcal{D}$ (см. п. 1.6.2). Если к тому же

$a(\mathcal{B})[0, l] < 0$, то $s^l(\mathcal{B})$ — прогрессивное направление в точке $x(\mathcal{B})$ (см. 3°) и линейная функция f не ограничена сверху на $\mathcal{D}$ (см. п. 2.2, 4°). Тем самым доказано следующее утверждение — обоснование рекомендации 3' этапа 1.

6°. Если $a(\mathcal{B})[0, l] < 0$ и $a^l(\mathcal{B}) \leq 0$, то f — не ограничена сверху на $\mathcal{D}$.

Следующее утверждение служит обоснованием процедуры выбора шага в алгоритме симплекс-метода.

7°. Пусть существует такое $i \in 1:m$, что $a(\mathcal{B})[i, l] > 0$, и число $\lambda(\mathcal{B})$ определяется формулой (7). Тогда $\lambda(\mathcal{B}) = \max \{ \lambda : x(\mathcal{B}) + \lambda s^l(\mathcal{B}) \in \mathcal{D}, \lambda \geq 0 \}$ и $\lambda(\mathcal{B}) > 0$ для невырожденной задачи. Если к тому же $a(\mathcal{B})[0, l] < 0$, то $f(x(\mathcal{B}) + \lambda(\mathcal{B})s^l(\mathcal{B})) > f(x(\mathcal{B}))$.

Предположим теперь, что задача невырождена, $a(\mathcal{B})[0, l] < 0$ и существует такое $i \in 1:m$, что $a(\mathcal{B})[i, l] > 0$. Тогда вектор $s^l(\mathcal{B})$ — допустимое прогрессивное направление в точке $x(\mathcal{B})$ (см. 1° и 3°). В соответствии с рекомендациями п. 2.2, 4° и в силу 7° следует перейти из точки $x(\mathcal{B})$ в точку $x(\mathcal{B}) + \lambda(\mathcal{B})s^l(\mathcal{B})$. Оказывается, именно этот переход осуществляется в алгоритме: $x(\mathcal{B}) + \lambda(\mathcal{B})s^l(\mathcal{B})$ — допустимый базисный план (экстремальная точка множества $\mathcal{D}$), отвечающий допустимому базису $\mathcal{B}' = (\mathcal{B} \setminus \{a^l\}) \cup \{a^i\}$, где r определяется формулой (7). Это следует из утверждений 8°, 9° и является обоснованием рекомендации 4 алгоритма.

8°. Пусть $r \in 1:m$, $l \in 1:n \setminus N(\mathcal{B})$ и $a(\mathcal{B})[r, l] \neq 0$. Тогда система векторов $\mathcal{B}' = (\mathcal{B} \setminus \{a^l\}) \cup \{a^i\}$ — базис задачи $(\mathcal{D}, f)$.

9°. Пусть $l \in 1:n \setminus N(\mathcal{B})$, существует такое $i \in 1:m$, что $a(\mathcal{B})[i, l] > 0$, а $\lambda(\mathcal{B})$ и r определяются формулой (7). Тогда $\mathcal{B}' = (\mathcal{B} \setminus \{a^l\}) \cup \{a^i\}$ — допустимый базис и $x(\mathcal{B}') = x(\mathcal{B}) + \lambda(\mathcal{B})s^l(\mathcal{B})$.

Формулы (8) — (11), используемые при вычислении элементов табл. $T^1(\mathcal{B}')$, следуют из утверждения 10°.

10°. Пусть $a(\mathcal{B})[r, l] \neq 0$ при некотором $r \in 1:m$ и $\mathcal{B}'$ — базис задачи $(\mathcal{D}, f)$, полученный из базиса $\mathcal{B}$ заменой столбца a^l на столбец a^i (см. 8°). Обозначим через $C = (c[i, j], i \in 0:m, j \in 0:m)$ матрицу, столбцы которой, кроме столбца с номером r , совпадают с векторами канонического базиса пространства R^{m+1} , т. е. $c[j, j] = 1, j \in 0:m, j \neq r; c[i, j] = 0, i \in 0:m, j \in 0:m, i \neq j, j \neq r$, а элементы столбца c^r вычисляются по формулам

$$c[r, r] = 1/a(\mathcal{B})[r, l], \quad c[i, r] = -a(\mathcal{B})[i, l]/a(\mathcal{B})[r, l], \\ i \in 0:m, \quad i \neq r.$$

Тогда $B'^{-1} = CB^{-1}$ и элементы табл. $T^1(\mathcal{B}')$ связаны с элементами табл. $T^1(\mathcal{B})$ (см. (6)) формулами (8) — (11).

Обоснование алгоритма завершает доказательство следующих утверждений.

11°. Если задача $(\mathcal{D}, f)$ невырождена, то формула (7) определяет r однозначно.

12°. Для невырожденной задачи число итераций q симплекс-метода конечно.

3.7. Доказательство утверждений 1°–12°.

1°. Направление $s^i(\mathcal{B})$ допустимо в точке $x(\mathcal{B}) \in \mathcal{D}$, если $As^i(\mathcal{B}) = 0$ и $s^i(\mathcal{B})[j] \geq 0$ для $j \in J(x(\mathcal{B}))$ (см. п. 2.2, 3°). Первое из этих соотношений следует из равенства

$$As^i(\mathcal{B}) = -Ba^i(\mathcal{B}) + a^i = -BB^{-1}a^i + a^i = 0, \quad (14)$$

а неравенства $s^i(\mathcal{B})[j] \geq 0$, $j \in J(x(\mathcal{B}))$, — из определения вектора $s^i(\mathcal{B})$ (см. (13)), поскольку для невырожденной задачи $J(x(\mathcal{B})) = 1:n \setminus N(\mathcal{B})$. ■

2°. Покажем, что в рассматриваемом случае $x(\mathcal{B}) + \lambda s^i(\mathcal{B}) \in \mathcal{D}$ при любом $\lambda \geq 0$. Действительно, равенство $A(x(\mathcal{B}) + \lambda s^i(\mathcal{B})) = b$ выполняется при любом $\lambda \in R$, поскольку $x(\mathcal{B}) \in \mathcal{D}$ и $As^i(\mathcal{B}) = 0$ (см. (14)). Кроме того, из условия $a^i(\mathcal{B}) \leq 0$ следует $s^i(\mathcal{B}) \geq 0$ (см. (13)), и так как $x(\mathcal{B}) \geq 0$, то $x(\mathcal{B}) + \lambda s^i(\mathcal{B}) \geq 0$ при любом $\lambda \geq 0$. ■

3°. Условием прогрессивности вектора $s^i(\mathcal{B})$ является неравенство $cs^i(\mathcal{B}) > 0$ (см. п. 3.1.4 или п. 2.2, 3°), которое в силу равенства $cs^i(\mathcal{B}) = -c(\mathcal{B})a^i(\mathcal{B}) + c[l] = -a(\mathcal{B})[0, l]$ выполняется в том и только том случае, если $a(\mathcal{B})[0, l] < 0$. ■

4°. По определению матрицы $\bar{A}(\mathcal{B})$ и в силу (4) и (2)

$$a(\mathcal{B})[0, j] = d_0(\mathcal{B})\bar{a}^j = -1 \cdot c[j] + u(\mathcal{B})a^j, \quad j \in 1:n.$$

Поэтому $u(\mathcal{B})a^j = c[j]$ для $j \in N(\mathcal{B})$, и если $a_0(\mathcal{B}) \geq 0$, то $u(\mathcal{B})a^j \geq c[j]$ для $j \in 1:n$. Так как $J(x(\mathcal{B})) \subset 1:n \setminus N(\mathcal{B})$ (независимо от того, вырожден или невырожден план $x(\mathcal{B})$), то для допустимого плана $x(\mathcal{B})$ выполнены условия оптимальности п. 2.2, 2° (см. также п. 3.1.9). ■

5°. По определению векторов $b(\mathcal{B})$, $d_0(\mathcal{B})$ (см. (4)) и $x(\mathcal{B})$ $b(\mathcal{B})[0] = d_0(\mathcal{B})\bar{b} = -1 \cdot 0 + c(\mathcal{B})B^{-1}b = c(\mathcal{B})b(\mathcal{B}) = f(x(\mathcal{B}))$. ■

6°. Утверждение доказано в п. 3.6. ■

7°. В силу (14) равенство $A(x(\mathcal{B}) + \lambda s^i(\mathcal{B})) = b$ выполняется при любом $\lambda \in R$. Поэтому $\bar{\lambda} = \max \{\lambda : x(\mathcal{B}) + \lambda s^i(\mathcal{B}) \in \mathcal{D}, \lambda \geq 0\} = \max \{\lambda : x(\mathcal{B}) + \lambda s^i(\mathcal{B}) \geq 0\}$, и равенство $\bar{\lambda} = \lambda(\mathcal{B})$ очевидно следует из определения векторов $x(\mathcal{B})$ (см. п. 3.1, 7°), $s^i(\mathcal{B})$ (см. (13)) и числа $\lambda(\mathcal{B})$ (см. (7)). Предлагаем также сравнить формулу (7) с формулой (10) п. 2.2, 4°.

Так как для невырожденной задачи $b(\mathcal{B}) > 0$, то очевидно $\lambda(\mathcal{B}) > 0$. Если к тому же $a(\mathcal{B})[0, l] < 0$, то (см. 3°)

$$f(x(\mathcal{B}) + \lambda(\mathcal{B})s^i(\mathcal{B})) = f(x(\mathcal{B})) + \lambda(\mathcal{B})cs^i(\mathcal{B}) > f(x(\mathcal{B})). \quad \blacksquare$$

8°. Вектор $a^i(\mathcal{B})$ есть последовательность коэффициентов разложения вектора a^i по базису $\mathcal{B}$ (см. п. 3.1, 5°), т. е.

$a' = \sum_{k=1}^m a(\mathcal{B})[k, l] a'^k$, и поскольку $a(\mathcal{B})[r, l] \neq 0$, то $\mathcal{B}'$ — базис (см. упражнение 1.1). ■

9°. Так как $\mathcal{B}'$ — базис задачи $(\mathcal{D}, f)$ (см. 8°), то достаточно доказать равенство $x(\mathcal{B}') = x(\mathcal{B}) + \lambda(\mathcal{B}) s^l(\mathcal{B})$, т. е. проверить, что вектор $x(\mathcal{B}) + \lambda(\mathcal{B}) s^l(\mathcal{B})$ является решением системы уравнений $Ax = b$, $x[j] = 0$, $j \in 1:n \setminus N(\mathcal{B}') = 1:n \setminus ((N(\mathcal{B}) \setminus \{j_r\}) \cup \{l\})$ (см. п. 3.1, 7°). Очевидно, что $A(x(\mathcal{B}) + \lambda(\mathcal{B}) s^l(\mathcal{B})) = b$ (см. доказательство 2° или 7°). Очевидно также, что j -я компонента $(x(\mathcal{B}) + \lambda(\mathcal{B}) s^l(\mathcal{B}))[j]$ вектора $x(\mathcal{B}) + \lambda(\mathcal{B}) s^l(\mathcal{B})$ равна нулю при $j \in 1:n \setminus (N(\mathcal{B}) \cup \{l\})$, так как $x(\mathcal{B})[j] = s^l(\mathcal{B})[j] = 0$ для таких j . Наконец, из (7) следует, что $(x(\mathcal{B}) + \lambda(\mathcal{B}) s^l(\mathcal{B}))[j_r] =$
 $= x(\mathcal{B})[j_r] + \lambda(\mathcal{B}) s^l(\mathcal{B})[j_r] = b(\mathcal{B})[j_r] - \frac{b(\mathcal{B})[j_r]}{a(\mathcal{B})[r, l]} a(\mathcal{B})[r, l] = 0$. ■

10°. Пусть $\mathcal{B} = \{a^1, a^2, \dots, a^m\}$ — базис задачи $(\mathcal{D}, f)$; B и $\bar{B}$ — матрицы, составленные в соответствии с 2) и 4) п. 3.1, $A(\mathcal{B})$ и $\bar{A}(\mathcal{B})$ — матрицы, вычисленные в соответствии с 5) п. 3.1. Предполагаем, что столбцы и строки матриц $\bar{B}$, $\bar{B}^{-1}$, B' , B'^{-1} занумерованы индексами $0, 1, \dots, m$. Матрица B' отличается от матрицы B только столбцом с номером r . Следовательно, столбцы матрицы $\bar{C} = \bar{B}^{-1} \bar{B}'$, кроме столбца с номером r , — векторы канонического базиса пространства R^{m+1} , а столбец с номером r — вектор $\bar{B}^{-1} \bar{a}' = \bar{a}'(\mathcal{B})$ (строки и столбцы матрицы $\bar{C}$ также нумеруются индексами $0, 1, \dots, m$). Значит, $B'^{-1} = \bar{C}^{-1} \bar{B}^{-1}$. Непосредственной проверкой устанавливается, что $\bar{C}^{-1} = C$, где C — матрица, описанная в условии.

Формулы (8) — (11) следуют из равенств $\bar{A}(\mathcal{B}') = \bar{B}'^{-1} \bar{A} =$
 $= C \bar{B}^{-1} \bar{A} = C \bar{A}(\mathcal{B})$ и $\bar{b}(\mathcal{B}') = \bar{B}'^{-1} \bar{b} = C \bar{B}^{-1} \bar{b} = C \bar{b}(\mathcal{B})$, так как r -я строка матрицы C имеет вид $c_r = (0, \dots, 0, \underbrace{1/a(\mathcal{B})[r, l]}_r, 0,$

$\dots, 0)$, а каждая из остальных строк c_i матрицы C , $i \in 0:m$, $i \neq r$, содержит ровно два ненулевых элемента: $c[i, i] = 1$ и $c[i, r] = -a(\mathcal{B})[i, l]/a(\mathcal{B})[r, l]$. ■

11°. Если $\lambda(\mathcal{B}) = b(\mathcal{B})[r_1]/a(\mathcal{B})[r_1, l] = b(\mathcal{B})[r_2]/a(\mathcal{B})[r_2, l]$ при $r_1 \neq r_2$, то, как следует из 10° и формулы (11), по крайней мере одна компонента вектора $b(\mathcal{B}')$ равна нулю, что противоречит невырожденности задачи $(\mathcal{D}, f)$. ■

12°. В алгоритме симплекс-метода вычисления заканчиваются в двух случаях: если на некоторой итерации $a_0(\mathcal{B}) \geq 0$ (рекомендация 1' этапа 1) или если $a^l(\mathcal{B}) \leq 0$ и $a(\mathcal{B})[0, l] < 0$ (рекомендация 3' этапа 1). В первом случае получен оптимальный план задачи (см. 4°), во втором, как уже отмечалось в п. 3.6, функция f не ограничена сверху на $\mathcal{D}$. В противном случае в невырожденной задаче осуществляется переход от экстремальной точки $x(\mathcal{B})$ к экстремальной точке $x(\mathcal{B}')$ вдоль допустимого прогрессивного направления $s^l(\mathcal{B})$ (см. 1°, 3° и 9°). Кроме того, $f(x(\mathcal{B}')) > f(x(\mathcal{B}))$ (см. 7°) и ни одна из встретившихся в последовательности $x(\mathcal{B}_1), x(\mathcal{B}_2), \dots$ экстремальных точек не мо-

жет повторно встретиться в этой последовательности. Так как многогранное выпуклое множество имеет лишь конечное множество экстремальных точек (см. п. 1.5.5), или иначе: каждая из точек $x(\mathcal{B}_k)$, $k = 1, 2, \dots$, однозначно определяется допустимым базисом $\mathcal{B}_k$, а из столбцов матрицы A можно составить лишь конечное число различных базисов, то $x(\mathcal{B}_1), x(\mathcal{B}_2), \dots$ — конечная последовательность. Значит, в невырожденной задаче вычисления заканчиваются после конечного числа итераций. ■

З а м е ч а н и е. Из алгоритма симплекс-метода следует: если невырожденная задача (1) имеет решение, то среди ее оптимальных планов есть базисный план. Это утверждение верно и для вырожденной задачи (1). — Упражнение 4.7.

3.8. Построение допустимого базиса. Вычисления в алгоритме симплекс-метода начинаются после построения допустимого базиса $\mathcal{B}_1$. Иногда построение такого базиса удается осуществить без затруднений. Например, если матрица A содержит единичную подматрицу порядка m и $b > 0$, то столбцы этой подматрицы очевидно составляют допустимый базис задачи (1). Именно такая задача рассматривалась в п. 3.5. Однако в общем случае нахождение начального допустимого базиса составляет самостоятельную задачу.

Для ее решения можно использовать метод минимизации невязок (см. п. 2.2, 1°). Для этого, как и в п. 2.2, 1°, рассмотрим вспомогательную задачу

$$(\mathcal{D}', \psi): \psi(z) = c'z \rightarrow \max, \quad \mathcal{D}' = \{z \in R^{n+m}: A'z = b, z \geq 0\}, \quad (15)$$

$m \times (n+m)$ матрица ограничений которой $A' = [A \ I]$ содержит единичную подматрицу I порядка m и

$$c' = (\underbrace{0, \dots, 0}_n, \underbrace{-1, \dots, -1}_m).$$

Предположим, что $b > 0$ и задача (15) невырождена. Тогда множество $\mathcal{B}_0$ столбцов матрицы A' с $N(\mathcal{B}_0) = \{n+1, n+2, \dots, n+m\}$ является допустимым базисом задачи (15). Поскольку задача (15) всегда имеет решение (см. п. 2.2, 1°), то применение алгоритма п. 3.3 приведет к построению оптимального базисного плана $\bar{z}(\mathcal{B}^*)$ этой задачи. Если $N(\mathcal{B}^*) \subset 1:n$, то $\mathcal{B}^*$ очевидно является допустимым базисом задачи (1). В противном случае $\mathcal{D} = \emptyset$ (см. п. 2.2, 1°).

Несколько рекомендаций. Решение задачи (15) начинается с базиса $\mathcal{B}_0$; ему соответствует матрица

$$\bar{B}_0^{-1} = \left(\begin{array}{c|ccc} -1 & -1 & -1 & \dots & -1 \\ \hline 0 & & & & 1 \end{array} \right).$$

Если на k -й итерации из базиса выводится вектор $\bar{a}^j$ и $j \in N(\mathcal{B}_0)$, то в каждой из табл. $T^l(\mathcal{B}_0)$

при $p > k$ столбец $\bar{a}'(\mathcal{B}_p)$ можно не заполнять. Если $\mathcal{B}^*$ — оптимальный базис задачи (15) и $N(\mathcal{B}^*) \subset 1:n$, то полагаем $\mathcal{B}_1 = \mathcal{B}^*$ и приступаем к решению задачи (1). При этом полезно иметь в виду, что все элементы табл. $T^1(\mathcal{B}_1)$, кроме верхней строки, можно переписать из табл. $T^1(\mathcal{B}^*)$, составленной для задачи (15), а верхняя строка табл. $T^1(\mathcal{B}_1)$ заполняется по формулам $b(\mathcal{B}_1)[0] = c(\mathcal{B}^*)b(\mathcal{B}^*)$, $a(\mathcal{B}_1)[0, j] = -c[j] + c(\mathcal{B}^*)a'(\mathcal{B}^*)$ (см. доказательства 3° и 7° п. 3.7).

Как уже говорилось, не умаляя общности, можно считать, что $b \geq 0$. Если по крайней мере одна из компонент вектора b равна нулю, то задача $(\mathcal{D}, \psi)$ является вырожденной. Однако это совсем не исключает возможности ее решения с помощью алгоритма симплекс-метода (см. п. 3.4 и § 7).

§ 4. Симплекс-метод — вычислительная схема, связанная с преобразованием обратных матриц

4.1. Основные обозначения и определения введены в п. 3.1.

4.2. Краткая характеристика метода. В настоящем параграфе предлагается вычислительная схема алгоритма симплекс-метода для решения невырожденной канонической задачи линейного программирования $(\mathcal{D}, f)$ (см. (1) и 7) п. 3.1), позволяющая существенно сократить объем вычислений по сравнению с вычислительной схемой п. 3.3, если число строк в матрице ограничений A значительно меньше числа ее столбцов. Читателю после знакомства с методом предлагается самостоятельно оценить, при каких соотношениях между m и n объем вычислений на одной итерации алгоритма п. 4.3 меньше соответствующего объема вычислений, связанных с применением алгоритма п. 3.3. — Упражнение 4.8.

Как и в п. 3.3, решение невырожденной канонической задачи линейного программирования $(\mathcal{D}, f)$ начинают после построения допустимого базиса $\mathcal{B}_1$. Задача построения такого базиса рассмотрена в п. 3.8. Если план $x(\mathcal{B}_1)$ неоптимален, то так же, как и в предыдущем параграфе, можно перейти от базиса $\mathcal{B}_1$ к допустимому базису $\mathcal{B}_2$, отличающемуся от $\mathcal{B}_1$ только одним вектором, и т. д. Алгоритм позволяет за конечное число итераций q получить оптимальный план или обнаружить неограниченность сверху f на $\mathcal{D}$.

Пусть $\mathcal{B}'$ — базис, следующий за базисом $\mathcal{B}$ в последовательности $\mathcal{B}_1, \mathcal{B}_2, \dots$, и $\bar{B}, \bar{B}'$ — матрицы, отвечающие базисам $\mathcal{B}, \mathcal{B}'$ (см. п. 3.1, 4)). Тогда $\bar{B}'^{-1} = C\bar{B}^{-1}$ (см. п. 3.6, 10°). Возможность такого представления матрицы $\bar{B}'^{-1}$ по существу используется в алгоритме, к изложению которого приступаем.

4.3. Алгоритм. В процессе вычислений будут построены последовательность допустимых базисов $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_q$, заполнены таблицы вида

$d_0(\mathcal{B}_1)$	$\bar{A}$
$\vdots$	
$d_0(\mathcal{B}_q)$	$a_0(\mathcal{B}_1)$
	$\vdots$
	$a_0(\mathcal{B}_q)$

(1)

и последовательность таблиц $T^2(\mathcal{B}_1), T^2(\mathcal{B}_2), \dots, T^2(\mathcal{B}_q)$ вида

$$T^2(\mathcal{B}) = \left[\begin{array}{c|c|c|c|c} N(\mathcal{B}) & \bar{b}(\mathcal{B}) & \bar{B}^{-1} & \bar{a}^l & \bar{a}^l(\mathcal{B}) \end{array} \right],$$

отвечающая последовательности допустимых базисов. Базисы $\mathcal{B}_2, \dots, \mathcal{B}_q$ строятся вместе с соответствующими таблицами.

Этап 0. В табл. (1) заполняется матрица $\bar{A}$. Строится допустимый базис $\mathcal{B}_1$ (см. пп. 3.8 и 4.8). Полагаям $\mathcal{B}$ равным $\mathcal{B}_1$. По базису $\mathcal{B}$ формируется табл. $T^2(\mathcal{B})$, два ее последних столбца не заполняются. Переходим к этапу 1.

Этап 1. Как и в алгоритме п. 3.3, здесь будет описана одна итерация вычислительного процесса. Она состоит из проверки оптимальности «текущего» базисного плана $x(\mathcal{B})$ и последующего перехода к новому базису $\mathcal{B}'$ и плану $x(\mathcal{B}')$, если $x(\mathcal{B})$ не является решением задачи $(\mathcal{D}, f)$.

1. Из табл. $T^2(\mathcal{B})$ в табл. (1) переписывается $d_0(\mathcal{B})$ — верхняя строка матрицы $\bar{B}^{-1}$.

2. В табл. (1) заполняется строка $a_0(\mathcal{B})$. Элементы этой строки получаются скалярным умножением вектора $d_0(\mathcal{B})$ на столбцы матрицы $\bar{A}$, т. е. $a_0(\mathcal{B}) = d_0(\mathcal{B})\bar{A}$.

3. В табл. (1) рассматривается строка $a_0(\mathcal{B})$. Формируется множество индексов $J^-(\mathcal{B}) = \{j \in 1:n : a_0(\mathcal{B})[0, j] < 0\}$. Возможны два случая.

3'. Множество $J^-(\mathcal{B}) = \emptyset$, т. е. $a_0(\mathcal{B}) \geq 0$. Вычислительный процесс закончен. Выписываем оптимальный план $x(\mathcal{B})$: в столбце $N(\mathcal{B})$ расположены номера базисных компонент этого плана, рядом — в столбце $\bar{b}(\mathcal{B})$ — их значения; остальные $n - m$ компонент плана $x(\mathcal{B})$ равны нулю. Максимальное значение $f(x(\mathcal{B}))$ функции f равно $\bar{b}(\mathcal{B})[0]$.

3''. Множество $J^-(\mathcal{B}) \neq \emptyset$. Переходим к рекомендации 4.

4. Выбираем некоторое l из $J^-(\mathcal{B})$.

5. Из табл. (1) в табл. $T^2(\mathcal{B})$ переписывается вектор $\bar{a}^l$.

6. В табл. $T^2(\mathcal{B})$ заполняется столбец $\bar{a}^l(\mathcal{B}) = \bar{B}^{-1}\bar{a}^l$.

7. Рассматривается вектор $\bar{a}^l(\mathcal{B})$. Возможны два случая.

7'. $a'(\mathcal{B}) \leq 0$. Вычислительный процесс закончен: целевая функция f не ограничена сверху на $\mathcal{D}$.

7''. Существует такое $i \neq 0$, что $a(\mathcal{B})[i, l] > 0$. Переходим к рекомендации 8.

8. Вычисляется

$$\lambda(\mathcal{B}) = \min_{\substack{i: a(\mathcal{B})[i, l] > 0 \\ i \neq 0}} \frac{b(\mathcal{B})[l]}{a(\mathcal{B})[i, l]} = \frac{b(\mathcal{B})[r]}{a(\mathcal{B})[r, l]}. \quad (2)$$

Элемент $a(\mathcal{B})[r, l]$ является ведущим. В невырожденной задаче формула (2) определяет его однозначно. Столбец a^{lr} выводится из базиса. Формируются новый базис $\mathcal{B}' = \{a^{l1}, a^{l2}, \dots, a^{lr-1}, a^{lr}, a^{lr+1}, \dots, a^{lm}\}$ и множество $N(\mathcal{B}') = \{j_1, j_2, \dots, j_{r-1}, l, j_{r+1}, \dots, j_m\}$.

9. Заполняется табл. $T^2(\mathcal{B}')$: записывается столбец $N(\mathcal{B}')$, остальные элементы табл. $T^2(\mathcal{B}')$, кроме элементов двух последних столбцов, вычисляются по формулам

$$b(\mathcal{B}')[r] = b(\mathcal{B})[r] / a(\mathcal{B})[r, l], \quad (3)$$

$$d(\mathcal{B}')[r, j] = d(\mathcal{B})[r, j] / a(\mathcal{B})[r, l], \quad j \in 0:m, \quad (4)$$

$$b(\mathcal{B}')[i] = b(\mathcal{B})[i] - a(\mathcal{B})[i, l] b(\mathcal{B}')[r], \quad i \in 0:m, \quad i \neq r, \quad (5)$$

$$d(\mathcal{B}')[i, j] = d(\mathcal{B})[i, j] - a(\mathcal{B})[i, l] d(\mathcal{B}')[r, j], \\ i \in 0:m, \quad i \neq r, \quad j \in 0:m. \quad (6)$$

10. Полагаем $\mathcal{B} = \mathcal{B}'$. Переходим к 1. ■

Замечание. Вычисления, связанные с заполнением табл. $T^2(\mathcal{B}')$ по формулам (3) — (6), можно проводить в соответствии с рекомендациями к заполнению табл. $T^1(\mathcal{B}')$ (см. замечание п. 3.3). Аналогично сказанному в этом замечании рекомендуем после заполнения ведущей строки заполнить строку $d_0(\mathcal{B})$, переписать ее в табл. (1) и вычислить $a_0(\mathcal{B})$. Если $a_0(\mathcal{B}) \geq 0$, то в табл. $T^2(\mathcal{B}')$ заполняется только столбец $\bar{b}(\mathcal{B}')$, выписывается оптимальный план $x(\mathcal{B}')$ и $f(x(\mathcal{B}')) = b(\mathcal{B}')[0]$. Если в строке $a_0(\mathcal{B}')$ есть отрицательные элементы, то заполняется вся табл. $T^2(\mathcal{B}')$ (кроме двух последних столбцов).

4.4. Обсуждение. Обращаем внимание на проблему рационального выбора l из множества $J^-(\mathcal{B})$ и проблему решения вырожденных задач, которые рассматривались в п. 3.4. Эти проблемы существенны и для алгоритма п. 4.3, а рекомендации, которые были даны по их поводу, остаются в силе. Отметим лишь, что в отличие от алгоритма п. 3.3 здесь нет возможности обнаружить неограниченность сверху f на $\mathcal{D}$ до выбора l и вычисления вектора $a^l(\mathcal{B})$.

4.5. Пример. В табл. (7) и (8) приведено решение задачи $(\mathcal{D}, f)$, которая уже решена симплекс-методом в п. 3.5. Как и

$$\begin{aligned}
 T^2(\mathcal{B}_1) &= \begin{array}{c|ccc|cc} & -13 & -1 & -3 & -2 & 1 & -12 \\ \hline 4 & 3 & 0 & 1 & 0 & 1 & 1 \\ 2 & 2 & 0 & 0 & 1 & 4 & 4^* \end{array} \\
 T^2(\mathcal{B}_2) &= \begin{array}{c|cc|ccc|cc} & -7 & -1 & -3 & 1 & -1 & -7 \\ \hline 4 & 5/2 & 0 & 1 & -1/4 & 3 & 11/4^* \\ 3 & 1/2 & 0 & 0 & 1/4 & 1 & 1/4 \end{array} \\
 T^2(\mathcal{B}_3) &= \begin{array}{c|cc|ccc|cc} & -7/11 & -1 & -5/11 & 4/11 & 4 & -27/11 \\ \hline 1 & 10/11 & 0 & 4/11 & -1/11 & -1 & -7/11 \\ 3 & 3/11 & 0 & -1/11 & 3/11 & 3 & 10/11^* \end{array} \\
 T^2(\mathcal{B}_4) &= \begin{array}{c|cc|ccc|cc} & 1/10 & -1 & -7/10 & 11/10 & & \\ \hline 1 & 11/10 & 0 & 3/10 & 1/10 & & \\ 5 & 3/10 & 0 & -1/10 & 3/10 & & \end{array}
 \end{aligned} \tag{7}$$

$$\begin{array}{c|ccc|ccccc}
 -1 & -3 & -2 & -1 & -2 & 1 & -3 & 4 \\
 \hline
 -1 & -3 & 1 & 3 & 0 & 1 & 1 & -1 \\
 \hline
 -1 & 5/11 & 4/11 & 1 & 1 & 4 & 0 & 3 \\
 \hline
 -1 & 7/10 & 11/10 & -10 & 0 & -12 & 0 & -7 \\
 \hline
 & & & -7 & 3 & 0 & 0 & 2 \\
 \hline
 & & & 0 & 26/11 & 0 & 28/11 & -27/11 \\
 \hline
 & & & 0 & 31/10 & 27/10 & 23/10 & 0
 \end{array} \tag{8}$$

там, решение начато с базиса $\mathcal{B}_1 = \{a^4, a^2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ и базисного плана $x(\mathcal{B}_1) = (0, 2, 0, 3, 0)$. Таблица

$$T^2(\mathcal{B}_1) = \begin{array}{c|ccc|cc|cc} & -13 & -1 & -3 & -2 & & & \\ \hline 4 & 3 & 0 & 1 & 0 & & & \\ 2 & 2 & 0 & 0 & 1 & & & \end{array}$$

формируется по правилам и формулам п. 3.1. В табл. (8) переписываются исходные данные задачи: вектор c и матрица A .

Первая итерация состоит из следующих операций. Из табл. $T^2(\mathcal{B}_1)$ переписываем в (8) строку $d_0(\mathcal{B}_1) = (-1, -3, -2)$ матрицы B_1^{-1} и, скалярно умножая $d_0(\mathcal{B}_1)$ на столбцы матрицы $\bar{A}$, вычисляем элементы строки $a_0(\mathcal{B}_1) = (-10, 0, -12, 0, -7)$. Рассматриваются элементы строки $a_0(\mathcal{B}_1)$ и формируется множество $J^-(\mathcal{B}_1)$. Здесь $J^-(\mathcal{B}_1) = \{1, 3, 5\}$. Поскольку $J^-(\mathcal{B}_1) \neq \emptyset$, то базисный план $x(\mathcal{B}_1)$ неоптимален. Можно выбрать любой индекс l из $J^-(\mathcal{B}_1)$ и соответствующий столбец матрицы $\bar{A}$ переписать в табл. $T^2(\mathcal{B}_1)$. Как и в п. 3.5, выбираем l отвечающим наименьшему из $a(\mathcal{B}_1)[0, j]$, $j \in J^-(\mathcal{B}_1)$ (см. пп. 3.4 и 4.4). Поэтому $l = 3$, в табл. $T^2(\mathcal{B}_1)$ переписан столбец $\bar{a}^3$ и заполнен ведущий столбец $\bar{a}^3(\mathcal{B}_1) = \bar{B}_1^{-1}a^3$ (см. табл. $T^2(\mathcal{B}_1)$ в (7)). Поскольку $a^3(\mathcal{B}_1) = \begin{pmatrix} 1 \\ 4 \end{pmatrix} > 0$, вычисляется $\lambda(\mathcal{B}_1) = \min \{^3/1, ^2/4\} = 1/2$ (вы-

бирается минимальное из отношений элементов столбца $b(\mathcal{B}_1)$ к соответствующим положительным элементам столбца $a^3(\mathcal{B}_1)$). Элемент $a(\mathcal{B}_1)[2, 3] = 4$ является ведущим (он отмечен), столбец a^2 выводится из базиса, вводится в базис столбец a^3 , так что $\mathcal{B}_2 = \{a^4, a^3\}$, $N(\mathcal{B}_2) = \{4, 3\}$. По формулам (3) и (4) заполняется ведущая строка: она выделена в табл. $T^2(\mathcal{B}_2)$ (см. (7)); остальные элементы (кроме элементов двух последних столбцов) вычисляются по формулам (5), (6). Осуществлен переход к новому базису $\mathcal{B}_2$; $x(\mathcal{B}_2) = (0, 0, 1/2, 5/2, 0)$ $f(x(\mathcal{B}_2)) = -7 > -13 = f(x(\mathcal{B}_1))$.

Все последующие вычисления производятся аналогично, их результаты сведены в табл. (7) и (8). Из табл. $T^2(\mathcal{B}_4)$ можно выписать оптимальный план $x(\mathcal{B}_4) = (11/10, 0, 0, 0, 3/10)$ и максимальное значение целевой функции $f(x(\mathcal{B}_4)) = 1/10$; $q = 4$. Рекомендуем сравнить объем вычислений по данному алгоритму и алгоритму п. 3.3. Вектор $u(\mathcal{B}_4) = (-7/10, 11/10)$ является решением задачи $(\mathcal{D}^*, f^*)$, двойственной к задаче $(\mathcal{D}, f)$ (см. замечание к п. 4.6).

4.6. Обоснование алгоритма. Мы уже отмечали, что алгоритм этого параграфа отличается от алгоритма § 3 лишь схемой вычислений. Так же, как и в § 3, здесь строятся последовательность $\mathcal{B}_1, \mathcal{B}_2, \dots$ допустимых базисов и отвечающая этим базисам последовательность $x(\mathcal{B}_1), x(\mathcal{B}_2), \dots$ допустимых базисных планов — экстремальных точек множества $\mathcal{D}$. Проверка оптимальности базиса $\mathcal{B}$ из последовательности $\mathcal{B}_1, \mathcal{B}_2, \dots$ осуществляется на основе критерия 4° п. 3.6. Если условие этого критерия не выполняется, то для невырожденной задачи векторы $s^l(\mathcal{B})$, для которых $a(\mathcal{B})[0, l] < 0$, являются допустимыми прогрессивными направлениями (см. 1° и 3° п. 3.6). Тогда либо устанавливается неограниченность f на $\mathcal{D}$ (см. 6° п. 3.6), либо осуществляется переход к новому допустимому базису $\mathcal{B}'$, так

что $x(\mathcal{B}') = x(\mathcal{B}) + \lambda(\mathcal{B})s^i(\mathcal{B})$, $f(x(\mathcal{B}')) > f(x(\mathcal{B}))$ (см. 7°—9° п. 3.6). Формулы (3)—(6) следуют из утверждения 10° п. 3.6. Конечность алгоритма установлена в 12° того же пункта.

З а м е ч а н и е. Из приведенного в п. 3.7 доказательства утверждения 4° п. 3.6 следует, что если $a_0(\mathcal{B}) \geq 0$, то вектор $u(\mathcal{B}) = c(\mathcal{B})B^{-1}$ (см. п. 3.1, 4) является решением задачи $(\mathcal{D}^*, f^*)$, двойственной к задаче $(\mathcal{D}, f)$. — Упражнение 4.9. Поскольку вектор $u(\mathcal{B})$ составлен из последних m компонент вектора $d_0(\mathcal{B})$, то применение алгоритма п. 4.3 для решения задачи $(\mathcal{D}, f)$ приводит к одновременному вычислению оптимальных планов задач $(\mathcal{D}, f)$ и $(\mathcal{D}^*, f^*)$ (если f не ограничена сверху на $\mathcal{D}$, то $\mathcal{D}^* = \emptyset$). — См. п. 3.3.7).

4.7. Доказательство утверждений 1°—12° п. 3.6 приведено в п. 3.7.

4.8. Построение допустимого базиса. Задача построения допустимого базиса методом минимизации невязок была рассмотрена в п. 3.8. Если для решения задачи $(\mathcal{D}, f)$ применяется алгоритм симплекс-метода п. 4.3, то для решения вспомогательной задачи $(\mathcal{D}', \psi)$ также целесообразно использовать этот алгоритм.

После того как найден оптимальный базис $\mathcal{B}^*$ вспомогательной задачи $(\mathcal{D}', \psi)$, переход к решению канонической задачи $(\mathcal{D}, f)$ осуществляется следующим образом. Из табл. $T^2(\mathcal{B}^*)$ в табл. $T^2(\mathcal{B}_1)$ ($\mathcal{B}_1 = \mathcal{B}^*$) полностью переписываются m последних строк, а строку $(b(\mathcal{B}_1)[0], d_0(\mathcal{B}_1))$ можно заполнить по формулам

$$b(\mathcal{B}_1)[0] = c(\mathcal{B}^*)b(\mathcal{B}^*), \quad d_0(\mathcal{B}_1) = (-1, c(\mathcal{B}^*)B^{*-1}).$$

4.9. Геометрическая интерпретация. В п. 3.3.6 была приведена одна из геометрических интерпретаций канонической задачи линейного программирования

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \quad \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\}.$$

Было установлено, что решение задачи $(\mathcal{D}, f)$ сводится к нахождению параметра λ «последней» точки пересечения направленной прямой $L = \{b + \lambda e_{m+1}, \lambda \in R\}$ с многогранным выпуклым конусом $K_0 = \{z \in R^{m+1}: z = \lambda u, \lambda \geq 0, u \in v\{\bar{a}^1, \bar{a}^2, \dots, \bar{a}^n\}\}$ и неотрицательных коэффициентов разложения этой точки по образующим векторам $\bar{a}^i$ конуса K_0 .

Решение задачи $(\mathcal{D}, f)$ симплекс-методом начинается после построения некоторого допустимого базиса $\mathcal{B} = \{a^{1'}, a^{2'}, \dots, a^{m'}\}$. В пространстве R^{m+1} по базису $\mathcal{B}$ можно построить $(m-1)$ -мерный симплекс $C_{\mathcal{B}} = v\{\bar{a}^{1'}, \bar{a}^{2'}, \dots, \bar{a}^{m'}\}$ и многогранный выпуклый конус $K_{\mathcal{B}}$ с вершиной в нуле, натянутый на симплекс $C_{\mathcal{B}}$, т. е.

$$K_{\mathcal{B}} = \{z \in R^{m+1}: z = \lambda u, \lambda \geq 0, u \in C_{\mathcal{B}}\}.$$

$K_{\mathcal{B}} \subset K_0$. Поскольку $d_0(\mathcal{B}) \neq 0$ и $d_0(\mathcal{B}) \bar{a}^i = 0$ для $i \in 1:m$, то $d_0(\mathcal{B})$ является нормальным вектором m -мерной гиперплоскости, проходящей через конус $K_{\mathcal{B}}$.

В силу допустимости $x(\mathcal{B})$ и утверждения 3° п. 3.3.6 параметр точки пересечения прямой L с конусом $K_{\mathcal{B}}$ равен значению целевой функции в точке $x(\mathcal{B})$, т. е. $\lambda = f(x(\mathcal{B}))$.

Гиперплоскость с нормальным вектором $d_0(\mathcal{B})$, проходящая через конус $K_{\mathcal{B}}$ (и симплекс $C_{\mathcal{B}}$), разбивает множество небазисных векторов $\bar{a}^j$ ($j \in 1:n \setminus N(\mathcal{B})$) на два подмножества. В первое входят векторы, образующие с нормалью $d_0(\mathcal{B})$ тупой угол ($d_0(\mathcal{B}) \bar{a}^j < 0$), во второе — векторы, образующие нетупой угол ($d_0(\mathcal{B}) \bar{a}^j \geq 0$). Ясно, что если первое подмножество непусто, то существует точка на прямой L , принадлежащая конусу K_0 , с параметром $\lambda > f(x(\mathcal{B}))$. Тогда в силу утверждения 4° п. 3.3.6 план $x(\mathcal{B})$ не является оптимальным. Любой из векторов $\bar{a}^j$ из первого подмножества можно вводить в базис. Переход по алгоритму симплекс-метода к новому допустимому базису $\mathcal{B}'$ позволяет построить симплекс $C_{\mathcal{B}'}$, конус $K_{\mathcal{B}'}$ и m -мерную гиперплоскость с нормалью $d_0(\mathcal{B}')$, проходящую через конус $K_{\mathcal{B}'}$.

Точка пересечения прямой L с конусом $K_{\mathcal{B}'}$ соответствует значению параметра $\lambda' = f(x(\mathcal{B}'))$. В невырожденной задаче $\lambda' > \lambda = f(x(\mathcal{B}))$, и если $d_0(\mathcal{B}') \bar{a}^j \geq 0$ для $j \in 1:n \setminus N(\mathcal{B}')$, то получен оптимальный план задачи $(\mathcal{D}, f)$. В противном случае переходим к новому базису $\mathcal{B}''$, симплексу $C_{\mathcal{B}''}$ и т. д. до нахождения гиперплоскости, опорной к конусу K_0 в последней точке пересечения K_0 с прямой L . Для такой гиперплоскости и соответствующего базиса $\mathcal{B}$ выполняется условие $d_0(\mathcal{B}) \bar{a}^j \geq 0$, $j \in 1:n$, или, что то же, $a_0(\mathcal{B}) \geq 0$.

§ 5. Двойственный симплекс-метод

5.1. Обозначения и определения. Рассмотрим пару двойственных задач (см. п. 3.3.4):

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \quad \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\}, \quad (1)$$

$$(\mathcal{D}^*, f^*): f^*(u) = ub \rightarrow \min, \quad \mathcal{D}^* = \{u \in R^m: uA \geq c\}, \quad (2)$$

где $A = (a[i, j]) — m \times n$ матрица, ранг которой $r(A) = m$, $c \in R^n$, $b \in R^m$. В п. 3.1 было введено понятие базиса задачи (1) и показано, как по произвольному базису $\mathcal{B} = \{a^1, \dots, a^m\}$ формируются множество $N(\mathcal{B}) = \{j_1, \dots, j_m\}$, матрицы B , $\bar{B}$, $\bar{B}^{-1}$, векторы $c(\mathcal{B})$ и $u(\mathcal{B}) = c(\mathcal{B})B^{-1}$. Строку матрицы $\bar{B}^{-1}$ с номером $i \in 0:m$, как и прежде, обозначаем через $\bar{d}_i(\mathcal{B})$, а i -ю строку матрицы B^{-1} — через $\underline{d}_i(\mathcal{B})$, $i \in 1:m$, «подчеркивая», что она состоит из последних m компонент вектора $d_i(\mathcal{B})$.

Введем несколько определений. План $u(\mathcal{B}) \in R^m$ назовем *базисным планом задачи (2)*. Он является *допустимым планом за-*

дачи (2), если $u(\mathcal{B})A \geq c$. Базис $\mathcal{B}$, для которого $u(\mathcal{B})A \geq c$, назовем *двойственно допустимым*.

Допустимый базисный план $u(\mathcal{B})$ называется *невыврожденным*, если $u(\mathcal{B})a^j > c[j]$ для $j \in N(\mathcal{B})$ (из определения вектора $u(\mathcal{B})$ следует, что $u(\mathcal{B})a^j = c[j]$ для $j \in N(\mathcal{B})$), и *вырожденным* в противном случае. Задачу (2) назовем *невыврожденной*, если она не имеет вырожденных планов (см. п. 1.5). Задачу (1) будем называть *двойственно невыврожденной*, если не вырождена задача (2).

5.2. Предварительные замечания. Двойственный симплекс-метод позволяет за конечное число итераций q найти оптимальный план двойственно невыврожденной задачи $(\mathcal{D}, f)$ или обнаружить, что $\mathcal{D} = \emptyset$. Как и в предыдущих методах, число итераций q неизвестно до окончания вычислений. Оно зависит от выбора двойственно допустимого базиса $\mathcal{B}_1$, с которого начинается решение. Нахождение такого базиса — отдельная задача, методы ее решения рассматриваются в п. 5.8 и § 6.

При решении задачи $(\mathcal{D}, f)$ двойственным симплекс-методом одновременно строится и оптимальный план задачи $(\mathcal{D}^*, f^*)$ или устанавливается неограниченность снизу f^* на $\mathcal{D}^*$, если $\mathcal{D} = \emptyset$ (см. п. 3.3.7). Более того, двойственный симплекс-метод является реализацией метода допустимых направлений для решения задачи $(\mathcal{D}^*, f^*)$ (см. п. 2.3). В процессе решения строится последовательность $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_q$ двойственно допустимых базисов. В случае двойственной невыврожденности задачи $(\mathcal{D}, f)$ для допустимых базисных планов $u(\mathcal{B}_1), u(\mathcal{B}_2), \dots, u(\mathcal{B}_q)$ задачи $(\mathcal{D}^*, f^*)$ выполняются неравенства $f^*(u(\mathcal{B}_{k+1})) < f^*(u(\mathcal{B}_k))$, $k = 1, 2, \dots, q-1$, и либо $u(\mathcal{B}_q)$ — оптимальный базисный план задачи $(\mathcal{D}^*, f^*)$, либо на q -й итерации устанавливается неограниченность снизу f^* на $\mathcal{D}^*$. В первом случае базис $\mathcal{B}_q$ оказывается допустимым базисом задачи $(\mathcal{D}, f)$, а $x(\mathcal{B}_q)$ — оптимальным планом задачи $(\mathcal{D}, f)$.

Вычислительная схема алгоритма аналогична схеме § 4. Вычислительную схему алгоритма двойственного симплекс-метода, аналогичную схеме § 3, мы не рассматриваем.

5.3. Алгоритм. В процессе вычислений будут построена последовательность двойственно допустимых базисов $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_q$, заполнены табл. (3) (см. табл. на с. 170) и последовательность таблиц $T^2(\mathcal{B}_1), T^2(\mathcal{B}_2), \dots, T^2(\mathcal{B}_q)$ вида

$$T^2(\mathcal{B}) = \left[\begin{array}{c|c|c|c|c} N(\mathcal{B}) & \bar{b}(\mathcal{B}) & \bar{B}^{-1} & \bar{a}^t & \bar{a}^t(\mathcal{B}) \end{array} \right],$$

отвечающая последовательности двойственно допустимых базисов. Базисы $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_q$ будут строиться вместе с соответствующими таблицами.

Этап 0. В табл. (3) заполняется матрица $\bar{A}$. Дальнейшие вычисления производятся после построения двойственно допу-

стимого базиса $\mathcal{B}_1$ (см. п. 5.8 и § 6). Полагаем $\mathcal{B}$ равным $\mathcal{B}_1$. По базису $\mathcal{B}$ формируется табл. $T^2(\mathcal{B})$. Два последних ее столбца не заполняются. Переходим к этапу 1.

$d_r(\mathcal{B}_1)$	$\bar{A}$
$d_0(\mathcal{B}_1)$	
$\vdots$	
$d_r(\mathcal{B}_{q-1})$	
$d_0(\mathcal{B}_{q-1})$	$a_0(\mathcal{B}_1)$
	$\delta(\mathcal{B}_1)$
	$\vdots$
	$a_0(\mathcal{B}_{q-1})$
	$\delta(\mathcal{B}_{q-1})$

(3)

Этап 1. Опишем одну итерацию вычислительного процесса. Она состоит из проверки оптимальности (допустимости) «текущего» базисного плана $x(\mathcal{B})$ и последующего перехода в случае неоптимальности $x(\mathcal{B})$ к новому двойственно допустимому базису $\mathcal{B}'$ и плану $x(\mathcal{B}')$.

1. В табл. $T^2(\mathcal{B})$ рассматриваются вектор $b(\mathcal{B})$ и множество $I^-(\mathcal{B}) = \{i \in 1:m : b(\mathcal{B})[i] < 0\}$. Возможны два случая.

1'. $b(\mathcal{B}) \geq 0$ ($I^-(\mathcal{B}) = \emptyset$). Вычислительный процесс закончен. Из табл. $T^2(\mathcal{B})$ можно выписать оптимальный план $x(\mathcal{B})$ задачи (1): в столбце $N(\mathcal{B})$ расположены номера базисных переменных этого плана, рядом — в столбце $\bar{b}(\mathcal{B})$ — их значения; остальные $n - m$ компонент плана $x(\mathcal{B})$ равны нулю. Максимальное значение $f(x(\mathcal{B}))$ функции f равно $b(\mathcal{B})[0]$. Из этой же таблицы можно выписать решение $u(\mathcal{B})$ задачи (2). Вектор $u(\mathcal{B})$ расположен в строке $d_0(\mathcal{B}) = (-1, u(\mathcal{B}))$. Минимальное значение $f^*(u(\mathcal{B}))$ функции f^* равно $b(\mathcal{B})[0]$.

1''. Существует $i \neq 0$, для которого $b(\mathcal{B})[i] < 0$. Переходим к рекомендации 2.

2. Выбираем некоторое r из множества индексов $I^-(\mathcal{B})$.

3. Из табл. $T^2(\mathcal{B})$ в табл. (3) переписывается строка $d_r(\mathcal{B})$, по формуле

$$\delta(\mathcal{B})[j] = d_r(\mathcal{B}) \bar{a}^j \quad (4)$$

вычисляются элементы строки $\delta(\mathcal{B})$ табл. (3) с номерами $j \in 1:n \setminus N(\mathcal{B})$. Остальные элементы строки $\delta(\mathcal{B})$ не вычисляются.

4. Рассматривается строка $\delta(\mathcal{B})$. Возможны два случая.

4'. $\delta(\mathcal{B})[j] \geq 0$ для $j \in 1:n \setminus N(\mathcal{B})$. Вычислительный процесс закончен: функция f^* не ограничена снизу на $\mathcal{D}^*$, а $\mathcal{D} = \emptyset$.

4''. Существует такое $j \in 1:n \setminus N(\mathcal{B})$, что $\delta(\mathcal{B})[j] < 0$. Переходим к рекомендации 5.

5. Из табл. $T^2(\mathcal{B})$ в табл. (3) переписывается вектор $d_0(\mathcal{B})$ и заполняется строка $a_0(\mathcal{B}) = d_0(\mathcal{B})\bar{A}$.

6. Вычисляется

$$\lambda(\mathcal{B}) = \min_{j: \delta(\mathcal{B})[j] < 0} \frac{a(\mathcal{B})[0, j]}{-\delta(\mathcal{B})[j]} = \frac{a(\mathcal{B})[0, l]}{-\delta(\mathcal{B})[l]}. \quad (5)$$

Будет показано, что в двойственно невырожденной задаче (1) (в невырожденной задаче (2)) формула (5) определяет l однозначно.

7. Из табл. (3) в табл. $T^2(\mathcal{B})$ переписывается вектор $\bar{a}^l$ и заполняется столбец $\bar{a}^l(\mathcal{B}) = B^{-1}\bar{a}^l$. Элемент $a(\mathcal{B})[r, l] = \delta(\mathcal{B})[l] < 0$ является ведущим. Из базиса $\mathcal{B}$ выводится вектор $a^l r$. Новый базис $\mathcal{B}' = \{a^{l_1}, a^{l_2}, \dots, a^{l_{r-1}}, a^l, a^{l_{r+1}}, \dots, a^{l_m}\}$ и $N(\mathcal{B}') = \{j_1, j_2, \dots, j_{r-1}, l, j_{r+1}, \dots, j_m\}$.

8. Заполняется табл. $T^2(\mathcal{B}')$. В табл. $T^2(\mathcal{B}')$ записывается столбец $N(\mathcal{B}')$. Остальная часть табл. $T^2(\mathcal{B}')$, кроме элементов двух последних столбцов, заполняется по формулам

$$b(\mathcal{B}')[r] = b(\mathcal{B})[r]/a(\mathcal{B})[r, l], \quad (6)$$

$$d(\mathcal{B}')[r, j] = d(\mathcal{B})[r, j]/a(\mathcal{B})[r, l], \quad j \in 0:m, \quad (7)$$

$$b(\mathcal{B}')[i] = b(\mathcal{B})[i] - a(\mathcal{B})[i, l]b(\mathcal{B}')[r], \quad i \in 0:m, \quad i \neq r, \quad (8)$$

$$d(\mathcal{B}')[i, j] = d(\mathcal{B})[i, j] - a(\mathcal{B})[i, l]d(\mathcal{B}')[r, j], \quad (9)$$

$$i \in 0:m, \quad i \neq r, \quad j \in 0:m.$$

9. Полагаем $\mathcal{B} = \mathcal{B}'$. Переходим к 1. ■

З а м е ч а н и е. Вычисления, связанные с заполнением табл. $T^2(\mathcal{B}')$ по формулам (6)–(9), предлагается проводить в соответствии с рекомендациями к заполнению табл. $T^1(\mathcal{B}')$ (см. замечания пп. 3.3 и 4.3).

5.4. Обсуждение. Как и в п. 3.4, рассмотрим две проблемы. Первая из них связана с неоднозначностью выбора r из множества $I^-(\mathcal{B})$, если это множество состоит из более чем одного элемента. Предлагаем выбирать r так, чтобы $b(\mathcal{B})[r] = \min_{i \in I^-(\mathcal{B})} b(\mathcal{B})[i] = \gamma$. Такой выбор может сократить число итераций. Однако не исключено, что найдется несколько номеров i , для которых $b(\mathcal{B})[i] = \gamma$. В этом случае r можно положить равным наименьшему из таких i или осуществить случайный выбор r из множества $I^-(\mathcal{B})$.

Вторая проблема связана с возможной двойственной вырожденностью задачи. В этом случае нельзя гарантировать однозначность выбора l в соответствии с формулой (5), и $\lambda(\mathcal{B})$ мо-

жет оказаться равным нулю. Тем не менее, предлагается решать задачу и в этом случае, используя алгоритм п. 5.3. Если в процессе вычислений выяснится, что существует несколько p , для которых $a(\mathcal{B})[0, p] / -\delta(\mathcal{B})[p] = \lambda(\mathcal{B})$, то l выбирается равным одному из таких p (случайный выбор) или наименьшему из них. Подробнее об этом см. в § 7.

5.5. Пример. В табл. (11), (12) приведено решение задачи

$$((\mathcal{D}, f) : f(x) = -3x_1 - \frac{1}{2}x_4 \rightarrow \max, \mathcal{D} = \{x \in R^5 : \begin{array}{l} x_1 - 2x_4 + x_5 = -3, \\ -3x_1 + x_2 + x_4 = -6, \\ -x_1 + x_3 - x_4 = -8, \\ x_j \geq 0, \quad j \in 1:5 \end{array} \quad (10)$$

двойственным симплекс-методом:

$T^2(\mathcal{B}_1) =$		0	-1	0	0	0	-1/2	1/2
	5	-3	0	1	0	0	-2	-2
	2	-6	0	0	1	0	1	1
	3	-8	0	0	0	1	-1	-1*
$T^2(\mathcal{B}_2) =$		-4	-1	0	0	1/2	-3	5/2
	5	13	0	1	0	-2	1	3
	2	-14	0	0	1	1	-3	-4*
	4	8	0	0	0	-1	-1	1
$T^2(\mathcal{B}_3) =$		-51/4	-1	0	5/8	9/8		
	5	5/2	0	1	3/4	-5/4		
	1	7/2	0	0	-1/4	-1/4		
	4	9/2	0	0	1/4	-3/4		

Решение начато с двойственно допустимого базиса $\mathcal{B}_1$, $N(\mathcal{B}_1) = \{5, 2, 3\}$ и допустимого базисного плана $u(\mathcal{B}_1) = (0, 0, 0)$ задачи, двойственной к задаче (10) (все элементы строки $a_0(\mathcal{B}_1) = (3, 0, 0, 1/2, 0)$ (см. (3) и табл. (12)) неотрицательные).

Табл. (11) состоит из табл. $T^2(\mathcal{B}_1)$, $T^2(\mathcal{B}_2)$, $T^2(\mathcal{B}_3)$. Заполнение первой из них начинается на нулевом этапе алгоритма и осуществляется по базису $\mathcal{B}_1$. В табл. (12) на нулевом этапе заполняется только матрица $\bar{A}$. В дальнейшем табл. (12) на каждой итерации расширяется.

-1	0	0	0	-3	0	0	-1/2	0
0	0	0	1	1	0	0	-2	1
-1	0	0	1/2	-3	1	0	1	0
0	0	1	1	-1	0	1	-1	0
				3	0	0	1/2	0
				-1				-1
				5/2	0	1/2	0	0
				-4				1

(12)

Первая итерация состоит в следующем. Поскольку $I-(\mathcal{B}_1) = \{1, 2, 3\}$, допустимый базисный план $u(\mathcal{B}_1)$ задачи, двойственной к задаче (10), неоптимален. Базисный план (иногда говорят *псевдоплан*) $x(\mathcal{B}_1) = (0, -6, -8, 0, -3) \in \mathcal{D}$ и r можно выбрать равным 1, 2 или 3. В приведенном примере $r = 3$, поскольку $\min \{-3, -6, -8\} = -8 = b(\mathcal{B}_1)$ [3].

Предлагаем самостоятельно рассмотреть варианты, связанные с выбором $r = 1$ или $r = 2$.

Из табл. $T^2(\mathcal{B}_1)$ в табл. (12) переписана строка $d_3(\mathcal{B}_1) = (0, 0, 0, 1)$ и в табл. (12) вычислены элементы строки $\delta(\mathcal{B}_1)$ с номерами $j \in 1:5 \setminus N(\mathcal{B}_1) = \{1, 4\}$, $\delta(\mathcal{B}_1)[1] = -1$, $\delta(\mathcal{B}_1)[4] = -1$. Поскольку они отрицательные, можно приступить к вычислению строки $a_0(\mathcal{B}_1) = d_0(\mathcal{B}_1)\bar{A}$. Для этого из табл. $T^2(\mathcal{B}_1)$ в табл. (12) переписывается строка $d_0(\mathcal{B}_1) = (-1, 0, 0, 0)$; вычисляются элементы строки $a_0(\mathcal{B}_1) = (3, 0, 0, 1/2, 0)$ и $\lambda(\mathcal{B}_1) = \min \{a(\mathcal{B}_1)[0, 1]/-\delta(\mathcal{B}_1)[1], a(\mathcal{B}_1)[0, 4]/-\delta(\mathcal{B}_1)[4]\} = 1/2$. Следовательно, $l = 4$. Из табл. (12) в табл. $T^2(\mathcal{B}_1)$ переписывается

столбец $\bar{a}^4 = \begin{bmatrix} -1/2 \\ -2 \\ 1 \\ -1 \end{bmatrix}$, и вычисляется вектор $\bar{a}^4(\mathcal{B}_1) =$

$= \bar{B}_1^{-1}\bar{a}^4 = \begin{bmatrix} 1/2 \\ -2 \\ 1 \\ -1 \end{bmatrix}$. Поскольку $r = 3$, элемент $a(\mathcal{B}_1)[3, 4] = -1$ яв-

ляется ведущим, в табл. $T^2(\mathcal{B}_1)$ он отмечен; столбец $\bar{a}^4(\mathcal{B}_1)$ — ведущий, он выделен. По формулам (6) — (9) заполняется табл. $T^2(\mathcal{B}_2)$ (см. также замечание в п. 5.3). Заполнение табл. $T^2(\mathcal{B}_2)$ завершает первую итерацию.

Осуществлен переход от базиса $\mathcal{B}_1$ и базисного плана $u(\mathcal{B}_1)$ (плана $x(\mathcal{B}_1)$) к базису $\mathcal{B}_2 = \{a^5, a^2, a^4\}$ и базисному плану $u(\mathcal{B}_2) = (0, 0, 1/2)$ (плану $x(\mathcal{B}_2) = (0, -14, 0, 8, 13)$, $f^*(u(\mathcal{B}_2)) = -4 < 0 = f^*(u(\mathcal{B}_1))$). План $u(\mathcal{B}_2)$ неоптимален: $b(\mathcal{B}_2)[2] = -14 < 0$.

Все последующие итерации осуществляются аналогично. Напоминаем, что в табл. (11) выделены ведущие строки и столбцы.

Из табл. $T^2(\mathcal{B}_3)$ можно выписать оптимальный план $x(\mathcal{B}_3) = (7/2, 0, 0, 9/2, 5/2)$ задачи (10), максимальное значение целевой функции $f(x(\mathcal{B}_3)) = -51/4$, оптимальный план $u(\mathcal{B}_3) = (0, 5/8, 9/8)$ двойственной задачи и минимальное значение ее целевой функции $f^*(u(\mathcal{B}_3)) = -51/4$. В табл. $T^2(\mathcal{B}_3)$ выделены компоненты плана $u(\mathcal{B}_3)$.

5.6. Обоснование алгоритма. Покажем, что двойственный симплекс-метод, как и все предыдущие методы настоящей главы, является методом допустимых направлений (см. п. 2.1). Поэтому фактически будет рассмотрен тот же круг вопросов, который уже обсуждался в п. 2.3, они лишь конкретизируются с учетом специфики рассматриваемой вычислительной схемы (т. е. с учетом того, как осуществляется проверка оптимальности очередного плана, как выбираются допустимое прогрессивное направление, шаг и т. п.). Утверждения 1° — 12°, которые формулируются ниже, доказываются в п. 5.7.

Пусть $\mathcal{B}$ — двойственно допустимый базис (см. п. 5.1). Тогда допустимый базисный план $u(\mathcal{B})$ задачи (2) удовлетворяет системе уравнений $ua^j = c[j]$, $j \in N(\mathcal{B})$ и, следовательно (см. теорему п. 1.5.5), является экстремальной точкой (вершиной) многогранного выпуклого множества $\mathcal{D}^*$. Справедливы следующие утверждения.

1°. Если задача (2) невырождена, то при каждом $r \in 1 : m$ вектор $d_r(\mathcal{B})$ — допустимое направление в точке $u(\mathcal{B})$.

2°. Если $\delta(\mathcal{B})[j] = d_r(\mathcal{B})a^j \geq 0$ для каждого $j \in 1 : n \setminus N(\mathcal{B})$ (см. (4)), то $d_r(\mathcal{B})$ — допустимое направление в точке $u(\mathcal{B})$ и $u(\mathcal{B}) + \lambda d_r(\mathcal{B}) \in \mathcal{D}^*$ при любом $\lambda \geq 0$.

Для $r \in 1 : m$ положим $H_r^* = \{u \in \mathbb{R}^m : ua^j = c[j], j \in N(\mathcal{B}) \setminus \{j_r\}\}$. Тогда $\dim H_r^* = 1$. Если задача (2) невырождена, то грань $\mathcal{D}_r^* = \mathcal{D}^* \cap H_r^*$ — ребро многогранного выпуклого множества $\mathcal{D}^*$ (см. п. 1.7.5), исходящее из вершины $u(\mathcal{B})$, и множество $\{\mathcal{D}_r^*, r \in 1 : m\}$ содержит все ребра, исходящие из этой вершины. — Упражнение 4.10. В вырожденной задаче не все множества $\mathcal{D}_r^*$ обязаны быть ребрами, и из вершины $u(\mathcal{B})$ может исходить больше чем m ребер.

3°. Для любого $r \in 1 : m$ справедливо равенство $b(\mathcal{B})[r] = d_r(\mathcal{B})b$ и вектор $d_r(\mathcal{B})$ в том и только том случае является прогрессивным направлением в точке $u(\mathcal{B})$, если $b(\mathcal{B})[r] < 0$.

4°. Если $b(\mathcal{B}) \geq 0$, то $u(\mathcal{B})$ — решение задачи $(\mathcal{D}^*, f^*)$, а $x(\mathcal{B})$ — решение задачи $(\mathcal{D}, f)$.

Из утверждения 4° следует, что для оптимальности допусти-

мого базисного плана $u(\mathcal{B})$ задачи (2) достаточно (без предположения о невырожденности задачи (2)), чтобы среди векторов $\underline{d}_r(\mathcal{B})$ не было прогрессивных в точке $u(\mathcal{B})$ направлений. Это условие вместе с утверждением 5° является обоснованием рекомендации 1' этапа 1.

5°. $f(x(\mathcal{B})) = b(\mathcal{B})[0] = f^*(u(\mathcal{B}))$.

Если задача (2) невырождена, то условие 4° очевидно является и необходимым для оптимальности базисного плана $u(\mathcal{B})$ (см. п. 3.1.2). Предлагаем читателю самостоятельно построить пример (вырожденной) задачи (2), для которой 4° не будет необходимым условием оптимальности. — Упражнение 4.11.

Из 2° следует, что если $\delta(\mathcal{B})[j] = \underline{d}_r(\mathcal{B})a^j \geq 0$ при каждом $j \in 1:n \setminus N(\mathcal{B})$, то вектор $\underline{d}_r(\mathcal{B})$ принадлежит конусу неограниченности многогранного выпуклого множества $\mathcal{D}^*$ (см. п. 1.6.2). Если к тому же $b(\mathcal{B})[r] < 0$, то $\underline{d}_r(\mathcal{B})$ — прогрессивное направление в точке $u(\mathcal{B})$ и линейная функция f^* не ограничена снизу на $\mathcal{D}^*$ (см. п. 2.3, 4°), а $\mathcal{D} = \emptyset$ (см. п. 3.3.7). Тем самым доказано утверждение 6° — обоснование рекомендации 4' этапа 1.

6°. Если $b(\mathcal{B})[r] < 0$ и $\delta(\mathcal{B})[j] \geq 0$ для $j \in 1:n \setminus N(\mathcal{B})$, то f^* не ограничена снизу на $\mathcal{D}^*$, а $\mathcal{D} = \emptyset$.

Утверждение 7° служит обоснованием процедуры выбора шага.

7°. Пусть $r \in 1:t$, существует такое $j \in 1:n \setminus N(\mathcal{B})$, что $\delta(\mathcal{B})[j] < 0$, и $\lambda(\mathcal{B})$ определяется формулой (5). Тогда $\lambda(\mathcal{B}) = \max \{ \lambda : u(\mathcal{B}) + \lambda \underline{d}_r(\mathcal{B}) \in \mathcal{D}^*, \lambda \geq 0 \}$. Если к тому же задача $(\mathcal{D}^*, f^*)$ невырождена и $b(\mathcal{B})[r] < 0$, то $\lambda(\mathcal{B}) > 0$ и $f^*(u(\mathcal{B}) + \lambda(\mathcal{B}) \underline{d}_r(\mathcal{B})) < f^*(\mathcal{B})$.

Предположим теперь, что задача (2) невырождена, $b(\mathcal{B})[r] < 0$ и существует такое $j \in 1:n \setminus N(\mathcal{B})$, что $\delta(\mathcal{B})[j] < 0$. Тогда вектор $\underline{d}_r(\mathcal{B})$ — допустимое прогрессивное направление в точке $u(\mathcal{B})$ (см. 1°, 3°). В этом случае в соответствии с рекомендациями п. 2.3 следует перейти из точки $u(\mathcal{B})$ в точку $u(\mathcal{B}) + \lambda(\mathcal{B}) \underline{d}_r(\mathcal{B})$. Именно такой переход и определяется в алгоритме: обоснованием этого служат утверждения 8°, 9°.

8°. Если $\delta(\mathcal{B})[l] = \underline{d}_r(\mathcal{B})a^l \neq 0$ при некотором $r \in 1:t$ и некотором $l \in 1:n \setminus N(\mathcal{B})$, то система векторов $\mathcal{B}' = (\mathcal{B} \setminus \{a^l\}) \cup \{a^l\}$ — базис.

9°. Пусть $r \in 1:t$, существует такое $j \in 1:n \setminus N(\mathcal{B})$, что $\delta(\mathcal{B})[j] < 0$ (см. (4)), и $\lambda(\mathcal{B})$ и l определяются формулой (5). Тогда $\mathcal{B}' = (\mathcal{B} \setminus \{a^l\}) \cup \{a^l\}$ — двойственно допустимый базис (см. п. 5.1) и $u(\mathcal{B}') = u(\mathcal{B}) + \lambda(\mathcal{B}) \underline{d}_r(\mathcal{B})$.

Обоснованием рекомендаций 8 этапа 1, связанных с вычислением элементов табл. $T^2(\mathcal{B}')$, служит утверждение 10°.

10°. Элементы табл. $T^2(\mathcal{B}')$ связаны с элементами табл. $T^2(\mathcal{B})$ формулами (6) — (9).

Доказательство утверждений 11°, 12° завершает обоснование алгоритма.

11°. Если задача (2) невырождена, то формула (5) определяет l однозначно.

12°. Для невырожденной задачи (2) число итераций q двойственного симплекс-метода конечно.

5.7. Доказательство утверждений 1°–12°.

1°. В силу условий (18) п. 2.3, 3° достаточно проверить, что $-d_r(\mathcal{B})a^i \leq 0$ для любого $j \in J(u(\mathcal{B}))$ (напомним, что $J(\bar{u}(\mathcal{B})) = \{j \in 1:n : u(\mathcal{B})a^j = c[j]\}$). Но для невырожденной задачи (2) (см. п. 5.1) $u(\mathcal{B})a^j > c[j]$, $j \in N(\mathcal{B})$. Следовательно, $J(u(\mathcal{B})) = N(\mathcal{B})$, и так как $d_r(\mathcal{B})$ — строка матрицы B^{-1} , то $d_r(\mathcal{B})a^j = 0$ для $j \in N(\mathcal{B})$, $j \neq j_r$, и $d_r(\mathcal{B})a^{j_r} = 1$. ■

2°. Достаточно показать, что $u(\mathcal{B}) + \lambda d_r(\mathcal{B}) \in \mathcal{D}^*$ при $\lambda \geq 0$, если $d_r(\mathcal{B})a^j \geq 0$ для $j \in 1:n \setminus N(\mathcal{B})$. Это действительно так, поскольку

$$(u(\mathcal{B}) + \lambda d_r(\mathcal{B}))a^j - c[j] = u(\mathcal{B})a^j - c[j] + \lambda d_r(\mathcal{B})a^j > 0, \quad (13)$$

$$j \in 1:n \setminus N(\mathcal{B}),$$

$$(u(\mathcal{B}) + \lambda d_r(\mathcal{B}))a^j - c[j] = u(\mathcal{B})a^j - c[j] = 0 \quad (14)$$

$$\text{для } j \in N(\mathcal{B}), \quad j \neq j_r,$$

$$(u(\mathcal{B}) + \lambda d_r(\mathcal{B}))a^{j_r} - c[j_r] = u(\mathcal{B})a^{j_r} - c[j_r] + \lambda = \lambda \geq 0. \quad \blacksquare \quad (15)$$

3°. В силу условия (19) п. 2.3, 3° (см. также п. 3.1.4) направление $\underline{d}_r(\mathcal{B})$ в том и только том случае прогрессивно в точке $u(\mathcal{B})$, если $-\underline{d}_r(\mathcal{B})b = -b(\mathcal{B})[r] > 0$. ■

4°, 5°. Проверим сначала равенство 5°. Из определения матрицы $\bar{B}$, векторов $\bar{b}$, $\bar{b}(\mathcal{B})$ и $x(\mathcal{B})$ (см. п. 3.1) следует, что

$$b(\mathcal{B})[0] = d_0(\mathcal{B})\bar{b} = -1 \cdot 0 + u(\mathcal{B})b = f^*(u(\mathcal{B})),$$

$$d_0(\mathcal{B})\bar{b} = -1 \cdot 0 + c(\mathcal{B})B^{-1}b = c(\mathcal{B})b(\mathcal{B}) = f(x(\mathcal{B})).$$

Если $b(\mathcal{B}) \geq 0$, то $x(\mathcal{B}) \in \mathcal{D}$, $u(\mathcal{B}) \in \mathcal{D}^*$ и $f(x(\mathcal{B})) = f^*(u(\mathcal{B}))$. Следовательно, выполняются условия теоремы 2 п. 3.3.5, и $x(\mathcal{B})$ — решение задачи $(\mathcal{D}, f)$, а $u(\mathcal{B})$ — решение задачи $(\mathcal{D}^*, f^*)$. ■

6°. Утверждение доказано в п. 5.6. ■

7°. В силу (13)–(15) $u(\mathcal{B}) + \lambda d_r(\mathcal{B}) \in \mathcal{D}^*$ при $\lambda \geq 0$ тогда и только тогда, когда $u(\mathcal{B})a^j - c[j] + \lambda d_r(\mathcal{B})a^j \geq 0$, $j \in 1:n \setminus N(\mathcal{B})$. Но

$$u(\mathcal{B})a^j - c[j] = -1 \cdot c[j] + c(\mathcal{B})B^{-1}a^j = d_0(\mathcal{B})\bar{a}^j = a(\mathcal{B})[0, j],$$

а $d_r(\mathcal{B})a^j = \delta(\mathcal{B})[j]$. Поэтому

$$\max \{ \lambda : u(\mathcal{B}) + \lambda d_r(\mathcal{B}) \in \mathcal{D}^*, \lambda \geq 0 \} =$$

$$= \min_{j: \delta(\mathcal{B})[j] < 0} \frac{\bar{a}(\mathcal{B})[0, j]}{-\delta(\mathcal{B})[j]} = \lambda(\mathcal{B})$$

(см. (5)). Эта формула имеет смысл, поскольку множество $\{j \in 1:n \setminus N(\mathcal{B}) : \delta(\mathcal{B})[j] < 0\} \neq \emptyset$. Так как для невырожденной задачи (2) $a(\mathcal{B})[0, j] > 0$, $j \in 1:n \setminus N(\mathcal{B})$, то $\lambda(\mathcal{B}) > 0$. Если к тому же $b(\mathcal{B})[r] < 0$, то

$$f^*(u(\mathcal{B}) + \lambda(\mathcal{B}) \underline{d}_r(\mathcal{B})) = f^*(u(\mathcal{B})) + \lambda(\mathcal{B}) b(\mathcal{B})[r] < f^*(u(\mathcal{B})). \blacksquare$$

8°. Так как $\delta(\mathcal{B})[l] = \underline{d}_r(\mathcal{B}) a^l = a(\mathcal{B})[r, l]$, то это утверждение равносильно утверждению 8° п.3.6. $\blacksquare$

9°. Из 8° следует, что $\mathcal{B}'$ — базис. Из 7° следует, что $u(\mathcal{B}) + \lambda(\mathcal{B}) \underline{d}_r(\mathcal{B}) \in \mathcal{D}^*$. Покажем, что $u(\mathcal{B}') = u(\mathcal{B}) + \lambda(\mathcal{B}) \underline{d}_r(\mathcal{B})$. Действительно, очевидно

$$\begin{aligned} (u(\mathcal{B}) + \lambda(\mathcal{B}) \underline{d}_r(\mathcal{B})) a^i &= c[j_i] + \lambda(\mathcal{B}) \underline{d}_r a^i = c[j_i] \text{ при } i \neq r, \\ (u(\mathcal{B}) + \lambda(\mathcal{B}) \underline{d}_r(\mathcal{B})) a^l &= c(\mathcal{B}) B^{-1} a^l + \lambda(\mathcal{B}) \delta(\mathcal{B})[l] = \\ &= a(\mathcal{B})[0, l] + c[l] - \frac{a(\mathcal{B})[0, l]}{\delta(\mathcal{B})[l]} \delta(\mathcal{B})[l] = c[l]. \blacksquare \end{aligned} \quad (16)$$

10°. См. утверждение 10° п. 3.6. $\blacksquare$

11°. Предположим, что формула (5) определяет l неоднозначно, т. е. существуют такие l_1 и l_2 из $1:n \setminus N(\mathcal{B})$, $l_1 \neq l_2$, что

$$\lambda(\mathcal{B}) = \frac{a(\mathcal{B})[0, l_1]}{-\delta(\mathcal{B})[l_1]} = \frac{a(\mathcal{B})[0, l_2]}{-\delta(\mathcal{B})[l_2]}. \quad (17)$$

Положим в (16) l равным l_1 и l равным l_2 . Тогда

$$u(\mathcal{B}') a^{l_1} = (u(\mathcal{B}) + \lambda(\mathcal{B}) \underline{d}_r(\mathcal{B})) a^{l_1} = c[l_1], \quad u(\mathcal{B}') a^{l_2} = c[l_2],$$

что противоречит невырожденности задачи (2). $\blacksquare$

12°. В алгоритме двойственного симплекс-метода вычисления заканчиваются в двух случаях: если на некоторой итерации $b(\mathcal{B}) \geq 0$ (рекомендация 1' этапа 1) или если $\delta(\mathcal{B})[j] \geq 0$ при $j \in 1:n \setminus N(\mathcal{B})$ (рекомендация 4' этапа 1). В первом случае получены оптимальные планы задач $(\mathcal{D}, f)$ и $(\mathcal{D}^*, f^*)$ (см. 4°), во втором — функция f^* не ограничена снизу на $\mathcal{D}^*$, а $\mathcal{D} = \emptyset$ (см. 6°). В противном случае в невырожденной задаче осуществляется переход от вершины $u(\mathcal{B})$ множества $\mathcal{D}^*$ к его вершине $u(\mathcal{B}')$ вдоль допустимого прогрессивного направления $\underline{d}_r(\mathcal{B})$ (см. 1°, 3° и 9°), так что $f^*(u(\mathcal{B}')) < f^*(u(\mathcal{B}))$. Следовательно, ни одна из встретившихся в последовательности $u(\mathcal{B}_1), u(\mathcal{B}_2), \dots$ вершин множества $\mathcal{D}^*$ не может повториться в этой последовательности. Так как многогранное выпуклое множество имеет лишь конечное число экстремальных точек (см. п. 1.5.5, или иначе: каждая из точек $u(\mathcal{B}_k)$, $k = 1, 2, \dots$, однозначно определяется двойственно допустимым базисом $\mathcal{B}_k$, а из столбцов матрицы A можно составить лишь конечное число различных базисов), то $u(\mathcal{B}_1), u(\mathcal{B}_2), \dots$ — конечная последовательность. Значит, в невырожденной задаче вычисления заканчиваются после конечного числа итераций q . $\blacksquare$

5.8. Нахождение двойственно допустимого базиса. Вычисления в алгоритме п. 5.3 начинаются после того, как сформирован двойственно допустимый базис задачи (1). Иногда построение такого базиса удается осуществить без затруднений. Действительно, пусть матрица A содержит единичную подматрицу порядка m . Из ее столбцов можно сформировать базис $\mathcal{B}_1$, и если $c[j] = 0$ для $j \in N(\mathcal{B}_1)$, $c[j] < 0$ для $j \in 1:n \setminus N(\mathcal{B}_1)$, то $\mathcal{B}_1$ является двойственно допустимым базисом, поскольку вектор $u(\mathcal{B}_1) = (0, 0, \dots, 0) \in \mathcal{D}^*$. План $u(\mathcal{B}_1)$ — невырожденный. Именно с такого плана начато решение задачи в п. 5.5.

В общем случае нахождение двойственно допустимого базиса — отдельная задача. Один из методов ее решения — градиентный метод — приведен в следующем параграфе.

5.9. Геометрические интерпретации. В этом пункте будут рассмотрены две геометрические интерпретации двойственного симплекс-метода. Первая из них предполагает знакомство с п. 3.3.6 и п. 4.9. Вторая связана с общими положениями п. 3.6, которые здесь будут проиллюстрированы примером.

Двойственный симплекс-метод — движение по $(m-1)$ -мерным симплексам. В процедуре двойственного симплекс-метода вычисления начинаются с двойственно допустимого базиса, и осуществляется последовательный перебор двойственно допустимых базисов, пока среди них не будет найден допустимый базис.

Ранее было показано (см. п. 3.3.6 и п. 4.9), как построить в R^{m+1} соответствующие задаче $(\mathcal{D}, f)$ многогранный выпуклый конус K_0 и прямую L , а также отвечающие базису $\mathcal{B}$ $(m-1)$ -мерный симплекс $S_{\mathcal{B}}$ и m -мерный многогранный выпуклый конус $K_{\mathcal{B}} \subset K_0$, натянутый на этот симплекс. Поскольку в двойственном симплекс-методе $\mathcal{B}$ — двойственно допустимый базис, то $d_0(\mathcal{B})\bar{a}^j = 0$, $j \in N(\mathcal{B})$, $d_0(\mathcal{B})\bar{a}^j \geq 0$, $j \in 1:n \setminus N(\mathcal{B})$, и

$$f^*(u(\mathcal{B})) \geq \max_{x \in \mathcal{D}} f(x) = \min_{u \in \mathcal{D}^*} f^*(u).$$

Поэтому гиперплоскость $H_{\mathcal{B}}$ в R^{m+1} , проходящая через конус $K_{\mathcal{B}}$, имеет уравнение $d_0(\mathcal{B})u = 0$, является опорной к конусу K_0 и пересекает прямую L в точке $\bar{b} + f^*(u(\mathcal{B}))e_{m+1}$, расположенной над конусом K_0 . Если $\bar{b} + f^*(u(\mathcal{B}))e_{m+1} \in K_0$, то вектор $u(\mathcal{B})$ является решением задачи $(\mathcal{D}^*, f^*)$, а $x(\mathcal{B})$ — решением задачи $(\mathcal{D}, f)$. В противном случае осуществляется перебор $(m-1)$ -мерных симплексов, которым отвечают проходящие через эти симплексы гиперплоскости, опорные к конусу K_0 и пересекающие L в точке, расположенной над K_0 . Процесс продолжается до построения конуса $K_{\mathcal{B}}$, пересекающегося с прямой L , если только $K_0 \cap L \neq \emptyset$ (см. п. 3.3.6, 2°). Если $K_0 \cap L = \emptyset$, то $\mathcal{D} = \emptyset$ и f^* не ограничена снизу на $\mathcal{D}^*$. Отметим, что в отличие от симплекс-метода, в котором движение производится по внутренним симплексам и все гиперплоскости, за исключением

последней (иногда и первой), пересекают конус K_0 , в двойственном симплекс-методе движение осуществляется по внешним симплексам и внешним гиперплоскостям, а конус K_0 всегда расположен в одном из порождаемых ими полупространств.

Двойственный симплекс-метод — движение по смежным вершинам. В п. 5.6 установлено, что двойственный симплекс-метод является специальной реализацией методов допустимых направлений для решения задачи $(\mathcal{D}^*, f^*)$, двойственной к канонической задаче линейного программирования. Некоторые из высказанных там соображений проиллюстрируем примером. Рассмотрим задачу

$$(\mathcal{D}, f): f(x) =$$

$$= -x_1 + 4x_2 + x_3 - 4x_4 \rightarrow \max,$$

$$\mathcal{D} = \{x \in R^4:$$

$$2x_1 + x_2 - x_4 = 4,$$

$$-x_1 + x_2 + x_3 + x_4 = 1,$$

$$x_j \geq 0, \quad j \in 1:4\}.$$

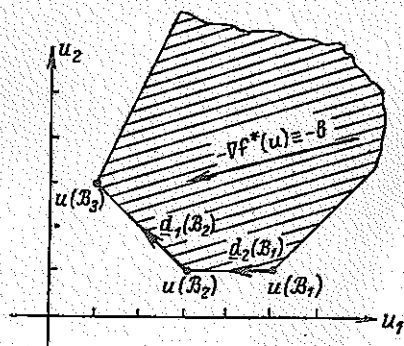


Рис. 25.

Ее решение двойственным симплекс-методом приведено в таблицах (18). На рис. 25 заштриховано множество $\mathcal{D}^* =$

	21	-1	5	1	4	2
3	5	0	1	1	1	2
4	-4	0	-1	0	1	-1*

	13	-1	3	1	-1	6
3	-3	0	-1	1	2	-3*
2	4	0	1	0	-1	2

	7	-1	1	3		
1	1	0	1/3	-1/3		
2	2	0	1/3	2/3		

-1	5	1	-1	4	1	-4
0	-1	0	2	1	0	-1
-1	3	1	-1	1	1	1
0	-1	1	10	2	0	0
			-2	-1		
			6	0	0	2
			-3			2

(18)

$= \{(u_1, u_2) \in R^2: 2u_1 - u_2 \geq -1, u_1 + u_2 \geq 4, u_2 \geq 1, -u_1 + u_2 \geq -4\}$. Решение задачи начато с двойственно допустимого базиса $\mathcal{B}_1 = \{a^3, a^4\} = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$. Ему отвечает вер-

шина $u(\mathcal{B}_1) = (5, 1)$. Вектор $d_2(\mathcal{B}_1) = (-1, 0)$ — допустимое прогрессивное направление в точке $u(\mathcal{B}_1)$. Величина шага $\lambda(\mathcal{B}_1) = 2$. Осуществляется переход из вершины $u(\mathcal{B}_1)$ в смежную вершину $u(\mathcal{B}_2)$ вдоль ребра, соединяющего эти вершины. Вектор $d_1(\mathcal{B}_2) = (-1, 1)$ — допустимое прогрессивное направление в точке $u(\mathcal{B}_2)$. Величина шага $\lambda(\mathcal{B}_2) = 2$. Осуществляется переход из вершины $u(\mathcal{B}_2)$ в смежную вершину $u(\mathcal{B}_3) = (1, 3)$. Точка $u(\mathcal{B}_3)$ — решение задачи $(\mathcal{D}^*, f^*)$. Базису $\mathcal{B}_3$ отвечает допустимый базисный план $x(\mathcal{B}_3) = (1, 2, 0, 0)$ — решение задачи $(\mathcal{D}, f)$.

§ 6. Градиентный метод

6.1. Постановка проблемы. Основные предположения. В этом параграфе приводятся один из методов построения двойственно допустимого базиса $\mathcal{B}$ (см. п. 5.1) канонической задачи линейного программирования

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\} \quad (1)$$

и вычисления допустимого базисного плана $u(\mathcal{B})$ задачи

$$(\mathcal{D}^*, f^*): f^*(u) = ub \rightarrow \min, \mathcal{D}^* = \{u \in R^m: uA \geq c\}, \quad (2)$$

двойственной к задаче $(\mathcal{D}, f)$.

Как и в п. 5.1, $A = (a[i, j])$ — $m \times n$ матрица, ее ранг $r(A) = m$, $m \leq n$, $c \in R^n$, $b \in R^m$. Предполагается, что читатель знаком с терминологией п. 5.1.

6.2. Предварительные замечания. Построение двойственно допустимого базиса задачи (2) градиентным методом начинается после вычисления допустимого плана u_1 задачи (2) (см. п. 2.3, 1°). Плану $u_1 \in \mathcal{D}^*$ поставим в соответствие следующую экстремальную задачу:

$$(\mathcal{D}^*(u_1), f^*): f^*(z) = zb \rightarrow \min, \mathcal{D}^*(u_1) = \{z \in R^m: \\ za^j = c[j], j \in J(u_1), za^j \geq c[j], j \in J(u_1)\},$$

где через $J(u_1)$, как обычно, обозначено множество тех $j \in 1:n$, для которых $u_1 a^j = c[j]$. Множество $\mathcal{D}^*(u_1)$ является гранью многогранного выпуклого множества $\mathcal{D}^*$ (см. п. 1.7.5), содержащей точку u_1 .

Предположим, что ранг системы линейных уравнений $\{za^j = c[j], j \in J(u_1)\}$ равен h_1 , и пусть $\mathcal{B}_1$ — базисная подматрица этой системы; множество номеров базисных уравнений обозначим через $N_y(\mathcal{B}_1) = \{j_1, j_2, \dots, j_{h_1}\}$, а множество базисных переменных — через $N_x(\mathcal{B}_1) = \{i_1, i_2, \dots, i_{h_1}\}$. Выразив базисные переменные через остальные и подставив их значения в каждое из неравенств $za^j \geq c[j]$, $j \in J(u_1)$, и целевую функцию f^* , перейдем к задаче $(\mathcal{D}_1, f_1, \mathcal{B}_1)$, эквивалентной задаче $(\mathcal{D}^*(u_1), f^*)$. При этом $\mathcal{D}_1 \equiv \mathcal{D}^*(u_1) \subset \mathcal{D}^*$ и $f_1(z) = f^*(z)$ при любом $z \in \mathcal{D}_1$.

Пример. Рассмотрим задачу

$$(\mathcal{D}^*, f^*) : f^*(u) = 2u[1] + 6u[2] + 7u[3] \rightarrow \min, \mathcal{D}^* = \{u \in R^3 : \\ u[1] + 2u[2] + u[3] \geq -2, u[1] + u[2] + u[3] \geq -1, \\ 2u[1] - 3u[2] + u[3] \geq 1, -u[1] + u[2] + u[3] \geq 1\}.$$

Непосредственно проверяется, что вектор $u_1 = (0, 0, 1) \in \mathcal{D}^*$ и множество $J(u_1) = \{3, 4\}$. Плану u_1 отвечает задача

$$(\mathcal{D}^*(u_1), f^*) : f^*(z) = 2z[1] + 6z[2] + 7z[3] \rightarrow \min, \mathcal{D}^*(u_1) = \{z \in R^3 : \\ z[1] + 2z[2] + z[3] \geq -2, z[1] + z[2] + z[3] \geq -1, \quad (3) \\ 2z[1] - 3z[2] + z[3] = 1, -z[1] + z[2] + z[3] = 1\}. \quad (4)$$

Легко проверить, что ранг h_1 системы уравнений (4) равен двум, $\mathcal{B}_1 = \begin{pmatrix} 2 & -3 \\ -1 & 1 \end{pmatrix}$ — базисная подматрица этой системы, $N_y(\mathcal{B}_1) = \{3, 4\}$, $N_n(\mathcal{B}_1) = \{1, 2\}$. Выразим переменную $z[1]$ из первого уравнения системы (4) через $z[2]$ и $z[3]$:

$$z[1] = \frac{3}{2}z[2] - \frac{1}{2}z[3] + \frac{1}{2} \quad (5)$$

и подставим вместо $z[1]$ выражение (5) в остальные уравнения и неравенства системы (3), (4). Получим следующую систему неравенств и уравнений, равносильную системе (3), (4):

$$\frac{7}{2}z[2] + \frac{1}{2}z[3] \geq -\frac{5}{2}, \frac{5}{2}z[2] + \frac{1}{2}z[3] \geq -\frac{3}{2}, \quad (3')$$

$$z[1] - \frac{3}{2}z[2] + \frac{1}{2}z[3] = \frac{1}{2}, -\frac{1}{2}z[2] + \frac{3}{2}z[3] = \frac{3}{2}. \quad (4')$$

Подставив выражение (5) вместо $z[1]$ в $f^*(z)$, получим

$$f^*(z) = 9z[2] + 6z[3] + 1. \quad (6)$$

Выразив теперь $z[2]$ через $z[3]$ из второго уравнения системы (4'), получим $z[2] = 3z[3] - 3$. Подставляя это выражение вместо $z[2]$ в остальные уравнения и неравенства системы (3'), (4'), а также в (6), имеем следующую систему неравенств и уравнений, равносильную системе (3'), (4'):

$$z[3] \geq \frac{8}{11}, z[3] \geq \frac{3}{4},$$

$$z[1] = 4z[3] - 4, z[2] = 3z[3] - 3,$$

а также $f_1(z) = 33z[3] - 26$. Следовательно, задача

$$(\mathcal{D}_1, f_1, \mathcal{B}_1) : f_1(z) = 33z[3] - 26 \rightarrow \min, \mathcal{D}_1 = \{z \in R^3 : 11z[3] \geq 8, \\ 8z[3] \geq 6, z[1] - 4z[3] = -4, z[2] - 3z[3] = -3\}$$

эквивалентна задаче $(\mathcal{D}^*(u_1), f^*)$, так как по построению $\mathcal{D}_1 \equiv \mathcal{D}^*(u_1)$ и $f_1(z) = f^*(z)$ при любом $z \in \mathcal{D}_1$ (см. п. 2.4.1). Отметим, что при построении задачи $(\mathcal{D}_1, f_1, \mathcal{B}_1)$ мы фактически применили метод исключения Гаусса.

Рассмотрим теперь задачу $(\mathcal{D}_1, f_1, \mathcal{B}_1)$. Очевидно, что если $u_1 \in \text{ext } \mathcal{D}^*$ (см. п. 1.5.1), то $u_1 \in \text{ext } \mathcal{D}_1$ и, более того, $u_1 \in \mathcal{D}_1^{00}$. Выберем какое-нибудь допустимое (в множестве $\mathcal{D}_1$) направление s_1 в точке u_1 . Если $f_1 \neq \text{const}$, то направление s_1 можно выбрать так, чтобы оно было прогрессивным. Тогда либо $u_1 + \lambda s_1 \in \mathcal{D}_1$ при любом $\lambda \geq 0$ и функция f_1 не ограничена снизу на $\mathcal{D}_1$, а значит, и f^* не ограничена снизу на $\mathcal{D}^*$, либо можно построить точку $u_2 = u_1 + \lambda_1 s_1$, $\lambda_1 > 0$, принадлежащую относительной границе $\partial \mathcal{D}_1$. Из леммы 2 п. 1.7.5 следует, что $\dim \mathcal{D}^*(u_2) < \dim \mathcal{D}^*(u_1)$. Кроме того, $f^*(u_2) < f^*(u_1)$. Если $\dim \mathcal{D}^*(u_2) = 0$, то $u_2 \in \text{ext } \mathcal{D}^*$ (см. п. 1.7.5). В противном случае построим задачу $(\mathcal{D}_2, f_2, \mathcal{B}_2)$, эквивалентную задаче $(\mathcal{D}^*(u_2), f^*)$, и т. д.

Одной из возможных реализаций рассмотренной выше вычислительной процедуры является алгоритм градиентного метода. Его применение позволяет за $q \leq m+1$ итераций (на каждой из которых заполняется отдельная таблица) найти допустимый базисный план задачи (2) и отвечающий этому плану двойственно допустимый базис задачи $(\mathcal{D}, f)$ или обнаружить неограниченность снизу f^* на $\mathcal{D}^*$. Поскольку алгоритм п. 6.3 дает возможность находить достаточно хорошие базисные планы, то его можно применять и как приближенный метод решения задачи (2). Отметим, что невырожденность задачи (2) (так же, как и невырожденность задачи (1)) не предполагается.

Название метода обусловлено тем, что переход от одного допустимого плана к другому осуществляется в направлении вектора $-\nabla f^*(u) = -b$ или его «проекции» на грань $\mathcal{D}^*(u)$ многогранного выпуклого множества $\mathcal{D}^*$.

6.3. Алгоритм. В процессе вычислений будут построены последовательность допустимых планов $u_1, u_2, \dots, u_q$ (некоторые из которых могут совпадать) и табл. T_k^3 , $k=1, 2, \dots, q$, вида

$$T_k^3 = \begin{array}{|c|c|c|c|} \hline & & b^k & \\ \hline \Delta(u_k) & \mu^k & A_k & c^k \\ \hline \end{array}.$$

Некоторые элементы столбца $\Delta(u_k)$ могут быть отмеченными.

Этап 0. 1. По допустимому плану u_1 задачи (2) формируется табл. T_1^3 , где $b^1 = b^T$, $c^1 = c^T$, $A_1 = A^T$, $\Delta(u_1) = (u_1 A - c)^T$. Столбец μ^1 на этом этапе не заполняется.

2. Полагаем k равным 1. Переходим к этапу 1.

Этап 1. Переход к эквивалентной задаче.

1. Если матрица A_k состоит из m различных единичных столбцов, то вычислительный процесс закончен. Из табл. T_k^3 выписываем план $u(\mathcal{B})$ и множество $N(\mathcal{B})$; i -я компонента плана $u(\mathcal{B})$ расположена в столбце c^k в той же строке, что и

единица i -го столбца матрицы A_k . Номер этой строки совпадает с индексом i -го столбца базиса $\mathcal{B}$.

В противном случае переходим к 2.

2. Среди не отмеченных элементов $\Delta(u_k)[j]$, $j \in J(u_k) = \{j \in 1:n: \Delta(u_k)[j] = 0\}$ находим $\Delta(u_k)[j]$ с минимальным индексом $j = l$. Отмечаем $\Delta(u_k)[l]$ и переходим к рекомендации 3.

Если множество не отмеченных элементов $\Delta(u_k)[j]$, $j \in J(u_k)$, пусто, то переходим к этапу 2.

3. В l -й строке матрицы A_k ищется первый по порядку (слева направо) ненулевой элемент $a^k[l, r]$.

Если все $a^k[l, i] = 0$, то переходим к рекомендации 2.

4. Заполняется табл. T_{k+1}^3 : если столбец $\Delta(u_{k+1})$ не был заполнен в соответствии с рекомендацией 4 этапа 2, то из табл. T_k^3 на место столбца $\Delta(u_{k+1})$ переписывается (со всеми отметками) столбец $\Delta(u_k)$; столбец μ^{k+1} на этом этапе не формируется. Остальная часть таблицы заполняется по формулам

$$a^{k+1}[l, i] = a^k[l, i]/a^k[l, r], \quad i \in 1:m, \quad (7)$$

$$c^{k+1}[l] = c^k[l]/a^k[l, r], \quad (8)$$

$$a^{k+1}[j, i] = a^k[j, i] - a^k[j, r]a^{k+1}[l, i], \quad j \in 1:n, j \neq l, i \in 1:m, \quad (9)$$

$$c^{k+1}[j] = c^k[j] - a^k[j, r]c^{k+1}[l], \quad j \in 1:n, j \neq l, \quad (10)$$

$$b^{k+1}[i] = b^k[i] - b^k[r]a^{k+1}[l, i], \quad i \in 1:m. \quad (11)$$

5. Полагаем k равным $k+1$. Переходим к рекомендации 1.

Этап 2. Выбор допустимого прогрессивного направления.

1. Если $b^k \neq 0$, переходим к рекомендации 2. В противном случае переходим к этапу 3.

2. Полагаем $\mu^k[j]$ равным $b^k a_j^k$ для $j \in J(u_k)$ и $\mu^k[j] = 0$ для $j \notin J(u_k)$. В столбец μ^k записываются $\mu^k[j]$ при $j \in J(u_k)$. Остальные элементы этого столбца не заполняются. Возможны два случая.

2'. $\mu^k[j] \leq 0$ для $j \in J(u_k)$. Вычислительный процесс закончен. Целевая функция f^* не ограничена снизу на $\mathcal{D}^*$.

2''. Существует j , для которого $\mu^k[j] > 0$. Переходим к рекомендации 3.

3. Вычисляется

$$\lambda(u_k) = \min_{j: \mu^k[j] > 0} \Delta(u_k)[j]/\mu^k[j]. \quad (12)$$

Полагаем l равным наименьшему из номеров p , для которых

$$\Delta(u_k)[p]/\mu^k[p] = \lambda(u_k).$$

4. Заполняется табл. T_{k+1}^3 : из табл. T_k^3 в столбец $\Delta(u_{k+1})$ переписываются со своими отметками $\Delta(u_k)[j]$, $j \in J(u_k)$. Остальные

элементы столбца $\Delta(u_{k+1})$ заполняются по формуле

$$\Delta(u_{k+1})[j] = \Delta(u_k)[j] - \lambda(u_k)\mu^k[j], \quad j \in J(u_k). \quad (13)$$

Отмечается $\Delta(u_{k+1})[l]$. Из (12) и (13) следует, что $\Delta(u_{k+1})[l] = 0$, и поскольку $\mu^k[l] > 0$, то в l -й строке матрицы A_k существует отличный от нуля элемент. Переходим к рекомендации 3 этапа 1.

Этап 3. Особый случай.

1. Среди ненулевых строк матрицы A_k , номера которых не входят в множество $J(u_k)$, выбираем строку a'_k с минимальным номером r . Будет показано, что такой выбор всегда можно осуществить. В строке a'_k находим первый по порядку элемент $a^k[r, t] \neq 0$.

2. Полагаем $\mu^k[j]$ равным $a^k[j, t] \text{sign } a^k[r, t]$ для $j \in J(u_k)$ и $\mu^k[j] = 0$ для $j \notin J(u_k)$. В столбец μ^k записываются $\mu^k[j]$ при $j \in J(u_k)$. Остальные элементы этого столбца не заполняются. Нетрудно проверить, что $\mu^k[r] > 0$. Переходим к рекомендации 3 этапа 2.

З а м е ч а н и е. Обращаем внимание на то, что для вычисления элементов табл. T_{k+1}^3 по формулам (7)–(11) удобно располагать эту таблицу под табл. T_k^3 и заполнять ее в соответствии с рекомендациями, приведенными в замечании п. 3.3, т. е. используя ведущую строку и ведущий столбец.

6.4. Пример. В табл. (14), состоящей из табл. $T_1^3, T_2^3, T_3^3, T_4^3$, градиентным методом построен допустимый базисный план задачи $(\mathcal{D}^*, f^*)$, сформулированной в примере п. 6.1. В таблицах выделены ведущие столбцы, строки и отмечены ведущие элементы. Решение начато с допустимого плана $u_1 = (0, 0, 1)$. В п. 6.1 была построена отвечающая плану u_1 задача $(\mathcal{D}_1, f_1, \mathcal{B}_1)$, эквивалентная задаче $(\mathcal{D}^*(u_1), f^*)$. Рекомендации этапа 1 приводят к выражению первой и второй переменных через третью и исключению их из выражения для целевой функции и из системы неравенств (3), определяющих вместе с системой уравнений (4) множество $\mathcal{D}^*(u_1)$. Эти вычисления проведены на первой и второй итерациях и отражены в табл. T_2^3 и T_3^3 (полезно сравнить эти таблицы с аналогичными вычислениями в примере п. 6.1). В табл. T_3^3 осуществлен переход к эквивалентной задаче $(\mathcal{D}_1, f_1, \mathcal{B}_1)$. Все $\Delta(u_1)[j]$, $j \in J(u_1) = \{3, 4\}$, отмечены. Поэтому в силу рекомендации 2 этапа 1 следующие вычисления проводятся в соответствии с рекомендациями этапа 2. Так как $b^3 = (0, 0, 33)$, то вычисляются $\mu^3[1] = b^3 a_1^3 = 363$ и $\mu^3[2] = b^3 a_2^3 = 264$. Поскольку $\mu^3[1] > 0$ и $\mu^3[2] > 0$, то вычисляется $\lambda(u_3) = \lambda(u_1) = \min \{3/363, 2/264\} = 2/264$. Таким образом, $l = 2$. По формуле (13) заполняется столбец $\Delta^T(u_4) = (1/4, 0^*, 0^*, 0^*)$, а по формулам (7)–(11) — остальная часть табл. T_4^3 . Матрица A_4 состоит из трех разных единичных столбцов. Вычислитель-

$$\begin{array}{l}
 T_1^3 = \begin{array}{|c|c|c|c|c|c|} \hline & & \boxed{2} & 6 & 7 & \\ \hline 3 & & 1 & 2 & 1 & -2 \\ 2 & & 1 & 1 & 1 & -1 \\ 0^* & & \boxed{2^*} & -3 & 1 & 1 \\ 0 & & -1 & 1 & 1 & 1 \\ \hline \end{array} \\
 \\
 T_2^3 = \begin{array}{|c|c|c|c|c|c|} \hline & & 0 & \boxed{9} & 6 & \\ \hline 3 & & 0 & 7/2 & 1/2 & -5/2 \\ 2 & & 0 & 5/2 & 1/2 & -3/2 \\ 0^* & & \boxed{1} & -3/2 & 1/2 & 1/2 \\ 0 & & 0 & \boxed{-1/2^*} & 3/2 & 3/2 \\ \hline \end{array} \\
 \\
 T_3^3 = \begin{array}{|c|c|c|c|c|c|} \hline & & 0 & 0 & \boxed{33} & \\ \hline 3 & 363 & 0 & 0 & 11 & 8 \\ 2 & 264 & 0 & 0 & 8^* & 6 \\ 0^* & & 1 & 0 & -4 & -4 \\ 0^* & & \boxed{0} & 1 & -3 & -3 \\ \hline \end{array} \\
 \\
 T_4^3 = \begin{array}{|c|c|c|c|c|c|} \hline & & 0 & 0 & 0 & \\ \hline 1/4 & & 0 & 0 & 0 & -1/4 \\ 0^* & & \boxed{0} & 0 & 1 & 3/4 \\ 0^* & & 1 & 0 & 0 & -1 \\ 0^* & & 0 & 1 & 0 & -3/4 \\ \hline \end{array}
 \end{array} \quad (14)$$

ный процесс закончен. Из табл. T_4^3 можно выписать базисный план $u(\mathcal{B}) = (-1, -3/4, 3/4)$, который (читатель может это проверить самостоятельно) является оптимальным, и базис $\mathcal{B} = \{a^3, a^4, a^2\}$ — двойственно допустимый и оптимальный для задачи $(\mathcal{D}, f)$, двойственной к задаче $(\mathcal{D}^*, f^*)$.

6.5. Обоснование алгоритма. Как и при обосновании вычислительных методов § 3—5, будет показано, что градиентный метод является методом допустимых направлений (см. пп. 2.1 и 2.3).

Решение задачи начинается после построения точки $u_1 \in \mathcal{D}^*$. Последовательное применение рекомендаций 2—4 этапа 1 поз-

воляет привести задачу $(\mathcal{D}^*(u_1), f^*)$ к виду $(\mathcal{D}_1, f_1, \mathcal{B}_1)$. Как уже отмечалось в п. 6.1, основой этих рекомендаций является метод исключения Гаусса. Выбор l и r на k -й итерации в рекомендациях 2 и 3 этапа 1 равносителен выбору нового базисного уравнения и новой базисной переменной, а преобразование табл. T_k^3 по формулам (7)–(11)—исключению этой переменной из выражения $f_k(z)$ и остальных уравнений и неравенств, определяющих множество

$$\mathcal{D}^*(u_k) = \{z \in R^m : za^j = c[j], j \in J(u_k); za^j \geq c[j], j \in J(u_k)\}.$$

Формальную проверку этого мы предоставляем читателю. — Упражнение 4.12. Естественно, что применение процедуры этапа 1 не приводит к изменению плана u_k .

Если после одной или нескольких итераций, связанных с такими преобразованиями, окажется, что матрица A_k состоит из m различных единичных столбцов, то u_k — очевидно базисный план, поскольку в этом случае множество $J(u_k)$ содержит m индексов системы линейно независимых столбцов матрицы A . Читателю предоставляется возможность самостоятельно проверить справедливость правила записи вектора $u(\mathcal{B})$ и базиса $\mathcal{B}$, предложенного в рекомендации 1 этапа 1. — Упражнение 4.13.

Предположим теперь, что табл. T_k^3 составлена в соответствии с рекомендациями этапа 1 для задачи $(\mathcal{D}_k, f_k, \mathcal{B}_k)$. Не умаляя общности, можно считать, что на k -й итерации базисная подматрица B_k расположена в левом верхнем углу матрицы A^r (этого всегда можно достигнуть перестановкой столбцов и строк матрицы A^r и такой же перестановкой коэффициентов целевой функции и компонент вектора c^r). Тогда

$$(\mathcal{D}_k, f_k, \mathcal{B}_k) : f_k(z) = zb^k + c^k[0] \rightarrow \min, \mathcal{D}_k = \{z \in R^m :$$

$$za_j^k = c^k[j], j \in J(u_k); za_j^k \geq c^k[j], j \in J(u_k)\},$$

$c^k[0] \in R$, векторы $b^k, a_j^k, j \in 1:n$, имеют вид

$$b^k = (\underbrace{0, \dots, 0}_{h_k}, b^k[h_k + 1], \dots, b^k[m]), \quad (15)$$

$$a_j^k = (\underbrace{0, \dots, 0, 1, 0, \dots, 0}_{h_k}, a^k[j, h_k + 1], \dots, a^k[j, m]), j \in 1:h_k, \quad (16)$$

$$a_j^k = (\underbrace{0, \dots, 0}_{h_k}, a^k[j, h_k + 1], \dots, a^k[j, m]), j \in (h_k + 1):n, \quad (17)$$

и поскольку B_k — базисная подматрица, то система векторов $\mathcal{B}_k = \{a^1, a^2, \dots, a^{h_k}\}$ линейно независима.

Обоснование алгоритма градиентного метода сводится к доказательству следующих утверждений.

1°. Если $b^k \neq 0$ и все $\Delta(u_k)[j], j \in J(u_k)$, отмечены, то

$$\bar{s}^k = (b^k a_1^k, \dots, b^k a_{h_k}^k, -b^k [h_k + 1], \dots, -b^k [m]) \quad (18)$$

— допустимое и прогрессивное направление в точке u_k (если $J(u_1) = \emptyset$, то $\bar{s}^k = -b$).

Если $b^k = 0, h_k \neq m$ и все $\Delta(u_k)[j], j \in J(u_k)$, отмечены, то найдутся такие $r \in J(u_k)$ и $t \geq h_k + 1$, что вектор

$$\bar{s}^k = \{\bar{s}^k[j] = a^k[j, t] \operatorname{sign} a^k[r, t], j \in 1 : h_k; \quad (19)$$

$$\bar{s}^k[j] = 0, j \in (h_k + 1) : m, j \neq t; \bar{s}^k[t] = -\operatorname{sign} a^k[r, t]\}$$

является допустимым направлением в точке u_k .

Очевидно, что $\mu^k = -\bar{s}^k A_k$, если $b^k \neq 0$, и $\mu^k = -\bar{s}^k A_k$, если $b^k = 0$ (определение μ^k см. в рекомендации 2 этапа 2 и в рекомендации 2 этапа 3).

2°. Если $\mu^k[j] \leq 0$ для всех $j \in J(u_k)$ и $b^k \neq 0$, то функция f^* не ограничена снизу на $\mathcal{D}^*$.

3°. Если $b^k \neq 0$, все $\Delta(u_k)[j], j \in J(u_k)$, отмечены, $h_k < m$ и существует такое $j \in J(u_k)$, что $\mu^k[j] > 0$, то (см. (12)) $\lambda(u_k) = \max\{\lambda : u_k + \lambda \bar{s}^k \in \mathcal{D}_k, \lambda \geq 0\}$, $\lambda(u_k) > 0$ и $f^*(u_{k+1}) < f^*(u_k)$.

Если $b^k = 0$, все $\Delta(u_k)[j], j \in J(u_k)$, отмечены и $h_k < m$, то $\lambda(u_k) = \max\{\lambda : u_k + \lambda \bar{s}^k \in \mathcal{D}_k, \lambda \geq 0\}$, $\lambda(u_k) > 0$ и $f^*(u_{k+1}) = f^*(u_k)$.

4°. Если все $\Delta(u_k)[j], j \in J(u_k)$, отмечены, $h_k < m$ и определено такое $l \in J(u_k)$, что $\mu^k[l] > 0$, то существует такое $r > h_k$, для которого $a^k[l, r] \neq 0$; система векторов $\mathcal{B}_{k+1} = \{a^1, \dots, a^{h_k}, a^l\}$ линейно независима; вычисление элементов табл. T_{k+1}^3 по формулам (7) — (11) позволяет перейти от задачи $(\mathcal{D}_k, f_k, \mathcal{B}_k)$ к задаче $(\mathcal{D}_{k+1}, f_{k+1}, \mathcal{B}_{k+1})$, эквивалентной задаче $(\mathcal{D}^*(u_{k+1}), f^*)$, где $u_{k+1} = u_k + \lambda(u_k) \bar{s}^k$ или $u_{k+1} = u_k + \lambda(u_k) \bar{s}^k$; $\mathcal{D}_{k+1} = \mathcal{D}^*(u_{k+1}) \subset \tilde{\mathcal{D}}_k$; $\Delta(u_{k+1}) = \Delta(u_k) - \lambda(u_k) \mu^k$; $J(u_k) \subset J(u_{k+1})$, $l \in J(u_{k+1}) \setminus J(u_k)$.

5°. Число итераций q конечно.

6.6. Доказательство утверждений 1°—5°.

1°. Пусть $b^k \neq 0$. В силу условий допустимости и прогрессивности направления $\bar{s}^k$ в точке $u_k \in \mathcal{D}_k = \mathcal{D}^*(u_k)$ (см. п. 3.1.3 и п. 3.1.4) достаточно проверить, что $\bar{s}^k a_j^k = 0, j \in J(u_k)$, и $-\bar{s}^k b^k > 0$. Так как все $\Delta(u_k)[j], j \in J(u_k)$, отмечены, то $a_j^k = 0$ для $j \in J(u_k) \setminus 1 : h_k$ и очевидно $\bar{s}^k a_j^k = 0$ для $j \in J(u_k) \setminus 1 : h_k$. Равенство $\bar{s}^k a_j^k = 0$ при $j \in 1 : h_k$ следует из (16) и (18), так как $b^k a_j^k = \sum_{i=h_k+1}^m b^k[i] a^k[j, i]$. Неравенство $-\bar{s}^k b^k > 0$ получаем из условия $b^k \neq 0$, поскольку

$$-\bar{s}^k b^k = \sum_{i=h_k+1}^m (b^k[i])^2 > 0. \quad (20)$$

Пусть $b^k = 0$. Так как ранг матрицы A равен m и $h_k < m$, то найдутся такие $r \in J(u_k)$ и $t \geq h_k + 1$, что $a^k[r, t] \neq 0$. По этому $\bar{s}^k \neq 0$ (см. (19)). Мы уже отмечали, что $a_j^k = 0$ для $j \in J(u_k) \setminus 1:h_k$, так что $\bar{s}^k a_j^k = 0$ для $j \in J(u_k) \setminus 1:h_k$. Из (16) и (19) следует, что $\bar{s}^k a_j^k = 0$ для $j \in 1:h_k$. Значит, $\bar{s}^k a_j^k = 0$ для $j \in J(u_k)$ и $\bar{s}^k$ — допустимое направление в точке $u_k \in \mathcal{D}_k$ (см. п. 3.1.3). ■

2°. Так как все $\Delta(u_k)[j]$, $j \in J(u_k)$, отмечены (вычисление $\mu^k[j]$ производится только в этом случае), то $a_j^k = 0$ для $j \in J(u_k) \setminus 1:h_k$. Поскольку $u_k A_k \geq c^k$, $\mu^k[j] = 0$, если $j \in J(u_k)$, и $\mu^k[j] \leq 0$ для $j \in J(u_k)$, то

$$[u_k + \lambda \bar{s}^k] A_k = u_k A_k - \lambda \mu^k \geq c^k \quad (21)$$

при любом $\lambda > 0$. Следовательно, вектор $\bar{s}^k$ принадлежит конусу неограниченности многогранного выпуклого множества $\mathcal{D}_k = \mathcal{D}^*(u_k)$ (см. п. 1.6.2), а так как $b^k \neq 0$, то $\bar{s}^k$ — прогрессивное направление в точке u_k (см. (20)). Значит, f_k не ограничена снизу на $\mathcal{D}^*(u_k)$ (см. п. 1.8.2, 6°). Но тогда и f не ограничена снизу на $\mathcal{D}^*$, поскольку $\mathcal{D}^*(u_k) \subset \mathcal{D}^*$, а $f^*(u) = f_k(u)$ для $u \in \mathcal{D}^*(u_k)$. ■

3°. Пусть $b^k \neq 0$. Тогда (см. 1°) $\bar{s}^k$ — допустимое прогрессивное направление в точке $u_k \in \mathcal{D}_k$, и $u_k + \lambda \bar{s}^k \in \mathcal{D}_k$ при $\lambda \geq 0$ в том и только том случае, если $(u_k + \lambda \bar{s}^k) a_j^k - c^k[j] \geq 0$, $j \in J(u_k)$. Но для $j \in J(u_k)$ из (15), (17) и (18) следует, что $\bar{s}^k a_j^k = b^k a_j^k = -\mu^k[j]$ (см. рекомендацию 2 этапа 2) и

$$(u_k + \lambda \bar{s}^k) a_j^k - c^k[j] = \Delta(u_k)[j] - \lambda \mu^k[j]. \quad (22)$$

Так как существует такое $j \in J(u_k)$, что $\mu^k[j] > 0$, то формула (12) имеет смысл. Следовательно, $\max \{ \lambda : u_k + \lambda \bar{s}^k \in \mathcal{D}_k, \lambda \geq 0 \} = \lambda(u_k)$. Неравенства $\lambda(u_k) > 0$ и $f^*(u_{k+1}) < f^*(u_k)$ следуют из допустимости и прогрессивности направления $\bar{s}^k$ в точке $u_k \in \mathcal{D}_k$.

Пусть $b^k = 0$. Тогда (см. 1°) $\bar{s}^k$ — допустимое направление в точке $u_k \in \mathcal{D}_k$ и $u_k + \lambda \bar{s}^k \in \mathcal{D}_k$ тогда и только тогда, когда $(u_k + \lambda \bar{s}^k) a_j^k - c^k[j] \geq 0$, $j \in J(u_k)$. Для $j \in J(u_k)$ из (15), (17) и (19) следует, что $\bar{s}^k a_j^k = -a^k[j, t] \text{ sign } a^k[r, t] = -\mu^k[j]$ (см. рекомендацию 2 этапа 3) и

$$(u_k + \lambda \bar{s}^k) a_j^k - c^k[j] = \Delta(u_k)[j] - \lambda \mu^k[j]. \quad (23)$$

Так как $\mu^k[r] > 0$, то формула (12) имеет смысл. Значит, $\max \{ \lambda : u_k + \lambda \bar{s}^k \in \mathcal{D}_k, \lambda \geq 0 \} = \lambda(\mathcal{D})$. Неравенство $\lambda(u_k) > 0$ сле-

дует из допустимости направления $\bar{s}^k$, и поскольку $b^k = 0$, то

$$f^*(u_{k+1}) = f_k(u_{k+1}) = f_k(u_k) + \lambda(u_k) b^k \bar{s}^k = f_k(u_k) = f^*(u_k). \blacksquare$$

4°. Из определения $\mu^k[l]$ следует, что $a_l^k \neq 0$ (см. рекомендации 2 этапов 2 и 3). Поэтому вектор a^l не может быть представлен в виде линейной комбинации векторов $a^1, a^2, \dots, a^{h_k}$, так что система векторов $\mathcal{B}_{k+1}$ линейно независима. Кроме того, $u_{k+1} \in \bar{\mathcal{D}}_k$ в силу определения шага $\lambda(u_k)$, а значит, $\dim \mathcal{D}^*(u_{k+1}) < \dim \mathcal{D}^*(u_k)$ (см. лемму 2 п. 1.7.5) и $\mathcal{D}^*(u_{k+1}) \subset \bar{\mathcal{D}}_k$. Формула $\Delta(u_{k+1})[j] = \Delta(u_k)[j] - \lambda(u_k) \mu^k[j]$ следует из (22) или (23) при $j \in J(u_k)$ и очевидна для $j \in J(u_{k+1})$.

В п. 6.5 отмечалось (см. упражнение 4.12), что преобразование табл. T_k^3 по формулам (7) — (11) равносильно преобразованию задачи $(\mathcal{D}^*(u_{k+1}), f^*)$ в эквивалентную ей задачу $(\mathcal{D}_{k+1}, f_{k+1}, \mathcal{B}_{k+1})$. $\blacksquare$

5°. Так как ранг матрицы A равен m и на каждой итерации порядок базисной подматрицы возрастает, то $q \leq m + 1$. $\blacksquare$

6.7. Замечание. О вычислении начального плана u_1 задачи $\mathcal{D}^*, f^*$ см. п. 2.3, 1°.

§ 7. Вырожденность. Решение вырожденных задач

7.1. Постановка проблемы. Каноническая задача линейного программирования (3.1) и двойственная к ней задача (см., например, п. 5.1) называются вырожденными, если имеют вырожденные планы. Напоминаем, что определения вырожденных планов для каждой из указанных задач приведены в пп. 3.1 и 5.1.

При решении вырожденных задач появляется возможность неоднозначного выбора индексов r или l в соответствии с формулами (7) п. 3.3, (2) п. 4.3 или (5) п. 5.3. В пп. 3.4 и 5.4 приведены некоторые правила выбора r или l в таких случаях. Так, например, в п. 3.4 рекомендовалось выбирать r равным наименьшему из номеров p , для которых $b(\mathcal{B})[p]/a(\mathcal{B})[p, l] = \lambda(\mathcal{B})$. Там же отмечалась эффективность случайного выбора r среди таких p .

Однако применение указанных правил может не дать желаемого результата. Суть дела состоит в том, что для вырожденных задач перестают быть верными некоторые из утверждений 1° — 12° пп. 3.6 или 5.6. Так, например, для вырожденной задачи не каждое из направлений $s^l(\mathcal{B})$ (направлений $\underline{d}_r(\mathcal{B})$) является допустимым, вследствие чего может оказаться, что $\lambda(\mathcal{B}) = 0$. Тогда нет гарантии, что при переходе от базиса $\mathcal{B}$ к базису $\mathcal{B}'$ осуществляется переход от вершины $x(\mathcal{B})(u(\mathcal{B}))$ многогранного выпуклого множества $\mathcal{D}(\mathcal{D}^*)$ в вершину $x(\mathcal{B}')(u(\mathcal{B}'))$, отличную от вершины $x(\mathcal{B})(u(\mathcal{B}))$. Поэтому в

принципе возможно построение такой последовательности базисов $\mathcal{B}_k, \mathcal{B}_{k+1}, \dots, \mathcal{B}_{k+t}$, что $\mathcal{B}_k = \mathcal{B}_{k+t}$, $x(\mathcal{B}_k) = x(\mathcal{B}_{k+1}) = \dots = x(\mathcal{B}_{k+t})$, а план $x(\mathcal{B}_k)$ не является оптимальным. Продолжать в такой ситуации вычисления бессмысленно: мы никогда не выйдем из указанного круга базисов. Обсуждаемое явление называется *зацикливанием*.

Возможность зацикливания при реализации на ЭВМ методов настоящей главы настораживает. Поэтому прежде всего необходимо выяснить, как часто на практике встречаются вырожденные задачи, поскольку только при их решении может произойти зацикливание, и как часто встречается зацикливание в вырожденных задачах. Здесь полезно привести следующее высказывание Дж. Данцига (см. [17, с. 225, 226]): «опыт, основанный на решении симплекс-методом тысяч практических задач линейного программирования, показывает, что почти каждая задача на некотором этапе решения оказывается вырожденной. Можно было бы думать, что так как вырождение происходит все время, то должно было бы наблюдаться много случаев зацикливания. Однако согласно данным не было ни одного случая зацикливания, за исключением специально сконструированных примеров Гофмана и Била».

Итак, по-видимому, с вычислительной точки зрения как само явление зацикливания, так и способы его устранения имеют чисто теоретический интерес, а рекомендации, приведенные в пп. 3.4 и 5.4, вполне оправдываются на практике.

Однако методы данной главы позволяют не только решать каноническую и двойственную к ней задачи линейного программирования. Они могут найти применение также при доказательстве некоторых теорем и утверждений, и здесь уже необходимо обоснование их конечности для общего случая. Поэтому приведем (без доказательства) один из методов устранения зацикливания.

7.2. Метод возмущения. Этот метод наиболее удобен для устранения зацикливания в алгоритме п. 3.3.

Если формула (7) п. 3.3. определяет r неоднозначно, то предлагается следующее правило выбора r .

1. Определяется множество

$$I_1 = \left\{ i \in 1:m : \frac{b(\mathcal{B})[i]}{a(\mathcal{B})[i, l]} = \lambda(\mathcal{B}), a(\mathcal{B})[i, l] > 0 \right\}.$$

2. p полагается равным 1.

3. Вычисляется

$$\lambda_p = \min_{i \in I_p} \frac{a(\mathcal{B})[i, p]}{a(\mathcal{B})[i, l]} = \frac{a(\mathcal{B})[r, p]}{a(\mathcal{B})[r, l]}. \quad (1)$$

Если формула (1) определяет r однозначно, то элемент $a(\mathcal{B})[r, l]$ является ведущим. Переходим к рекомендациям 5

этапа I алгоритма (п. 3.3). В противном случае переходим к рекомендации 4.

4. Определяется множество

$$I_{p+1} = \left\{ i \in I_p : \frac{a(\mathcal{B})[i, p]}{a(\mathcal{B})[i, i]} = \lambda_p \right\}.$$

5. p полагается равным $p+1$. Переходим к рекомендации 3. ■

Утверждается, что при использовании в алгоритме п. 3.3 такого правила выбора r можно гарантировать конечность числа итераций для произвольной канонической задачи линейного программирования. Обоснованием этого утверждения является следующая теорема Чарнса.

Пусть $\varepsilon \geq 0$. Рассмотрим задачу

$$(\mathcal{D}(\varepsilon), f) : f(x) = cx \rightarrow \max, \mathcal{D} = \{x \in R^n : Ax = b(\varepsilon), x \geq 0\},$$

где $b(\varepsilon) = b + \sum_{j=1}^n \varepsilon^j a^j$, а ε^j — j -я степень числа ε . Пусть $\mathcal{B}$ — базис. Обозначим через $x(\mathcal{B}, \varepsilon)$ базисный план задачи $(\mathcal{D}(\varepsilon), f)$. Очевидно, что $\mathcal{D}(0) = \mathcal{D}$ и $x(\mathcal{B}, 0) = x(\mathcal{B})$.

1. Существует такое $\varepsilon' > 0$, что для любого $\varepsilon \in (0, \varepsilon')$ задача $(\mathcal{D}(\varepsilon), f)$ является невырожденной.

2. Существует такое $\varepsilon'' > 0$, что для любого $\varepsilon \in [0, \varepsilon'')$ из допустимости базиса $\mathcal{B}$ в задаче $(\mathcal{D}(\varepsilon), f)$ следует его допустимость в задаче $(\mathcal{D}, f)$, а из оптимальности базиса $\mathcal{B}$ в задаче $(\mathcal{D}(\varepsilon), f)$ следует его оптимальность в задаче $(\mathcal{D}, f)$.

С доказательством этой теоремы можно ознакомиться, например, в [49].

Пусть теперь ε' и ε'' — числа, существование которых утверждается в теореме, и $\varepsilon \in (0, \min\{\varepsilon', \varepsilon''\})$. Задача $(\mathcal{D}(\varepsilon), f)$ невырождена, и ее можно решить методом § 3. Если $x(\mathcal{B}, \varepsilon)$ — решение задачи $(\mathcal{D}(\varepsilon), f)$, то $x(\mathcal{B})$ — решение задачи $(\mathcal{D}, f)$.

Заметим теперь, что для любого базиса $\mathcal{B}$

$$b(\mathcal{B}, \varepsilon) = B^{-1}b(\varepsilon) = b(\mathcal{B}) + \sum_{j=1}^n \varepsilon^j a^j(\mathcal{B}). \quad (2)$$

Если l — индекс вектора, который предполагается вводить в базис, то в силу (2) и формулы (7) п. 3.3

$$\begin{aligned} \lambda(\mathcal{B}, \varepsilon) &= \min_{i : a(\mathcal{B})[i, l] > 0} \left\{ \frac{b(\mathcal{B})[i]}{a(\mathcal{B})[i, l]} + \sum_{j=1}^n \varepsilon^j \frac{a(\mathcal{B})[i, j]}{a(\mathcal{B})[i, l]} \right\} = \\ &= \frac{b(\mathcal{B}, \varepsilon)[r]}{a(\mathcal{B})[r, l]}. \end{aligned} \quad (3)$$

Очевидно, что при малых значениях ε выбор r по формуле (3) согласуется с указанным выше правилом выбора r .

7.3. Геометрическая интерпретация. Возвращаясь к геометрической интерпретации задач двойственной пары (см. п. 3.3.6) и методов их решения (см. пп. 4.9 и 5.9), можно привести гео-

метрическую интерпретацию вырожденности и двойственной вырожденности, проясняющие идею метода возмущений.

Действительно, задача $(\mathcal{D}, f)$ вырождена в том и только том случае, если найдется такой допустимый базис $\mathcal{B}$, для которого не все компоненты вектора $b(\mathcal{B})$ положительны. Это означает, что прямая L пересекает грань многогранного выпуклого конуса $K_{\mathcal{B}}$ (см. п. 4.9) размерности меньше m . Теорема Чарнса утверждает возможность такого незначительного изменения вектора b в вырожденной задаче, что прямая L не будет пересекать грани конуса $K_{\mathcal{B}}$, размерности которых меньше чем m .

Задача $(\mathcal{D}, f)$ является двойственно вырожденной, если среди гиперплоскостей, опорных к конусу K_0 и пересекающих прямую L в точке, расположенной над конусом K_0 , есть гиперплоскость, содержащая больше чем m векторов из множества $\{\bar{a}_j, j \in 1:n\}$. Поэтому применение метода возмущений для устранения двойственной вырожденности задачи $(\mathcal{D}, f)$ связано с малыми изменениями компонент вектора c .

УПРАЖНЕНИЯ

4.1. См. с. 140. 4.2. См. с. 149. 4.3. См. с. 150.

4.4. См. с. 151. 4.5. См. с. 157. 4.6. См. с. 157.

4.7. См. с. 161. 4.8. См. с. 162. 4.9. См. с. 167.

4.10. См. с. 174. 4.11. См. с. 175. 4.12. См. с. 186.

4.13. См. с. 186.

4.14. Показать, что оптимальный план одной из задач двойственной пары вырожден в том и только том случае, если множество решений другой задачи этой пары состоит более чем из одного элемента.

Глава 5.

ПРИМЕНЕНИЕ ПРОЦЕДУРЫ СИМПЛЕКС-МЕТОДА ДЛЯ РЕШЕНИЯ НЕКОТОРЫХ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ

Экстремальные задачи, методы решения которых обсуждаются в настоящей главе, различны. В первом параграфе рассматривается метод решения задачи линейного программирования с ограниченными сверху переменными, во втором — конечный метод решения задачи квадратичного программирования, а в третьем — метод решения задач целочисленного линейного программирования специального вида. Общей для всех этих задач является возможность их решения применением процедуры симплекс-метода (см. § 3 или 4 гл. 4) с особыми для каждой из них правилами ввода в базис небазисных столбцов матрицы ограничений. Поэтому предполагается, что читатель уже ознакомился с этой процедурой, и в частности с понятиями и обозначениями п. 4.3.1.

§ 1. Метод решения задачи линейного программирования с ограниченными сверху переменными

1.1. Основные обозначения и определения. Рассмотрим задачу

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \mathcal{D} = \{x \in R^n: Ax = b, 0 \leq x \leq h\}, \quad (1)$$

где $A = (a[i, j])$ — $m \times n$ матрица, ранг которой $r(A) = m$, $c \in R^n$, $b \in R^m$. Предполагается, что $h^T = (h[1], \dots, h[n])$, $h > 0$ и $h[j] < +\infty$ для $j \in J'$, $J' \subset 1:n$. Задача (1) называется *задачей линейного программирования с ограниченными сверху переменными* (см. п. 2.4.10). Она может быть приведена к каноническому виду и, следовательно, решена любым из методов гл. 4. Однако при этом значительно возрастает число ограничений, а значит, и объем необходимых вычислений.

Оказывается, в таком преобразовании задачи (1) нет необходимости: ее решение может быть найдено достаточно эффективно применением алгоритма симплекс-метода с несколько измененной схемой вычислений. Описание такого алгоритма составит содержание настоящего параграфа. Для его изложения здесь выбрана вычислительная схема п. 4.4.3, знакомство с которой, так же как и с терминологией п. 4.3.1, предполагается. Кроме того, нам понадобятся некоторые новые понятия и обозначения.

Как и раньше, базисом задачи (1) называется любая система из m линейно независимых столбцов матрицы A . По базису $\mathcal{B}$ можно сформировать множество $N(\mathcal{B})$, вектор $c(\mathcal{B})$, матрицы B , $\bar{B}$ и $\bar{B}^{-1}$. Кроме того, базису $\mathcal{B}$ поставим в соответствие вектор $h(\mathcal{B}) = (h[j_1], h[j_2], \dots, h[j_m])$, $j_k \in N(\mathcal{B})$, $k \in 1:m$.

Рассмотрим базис $\mathcal{B}$ и множество $N^* \subset J' \setminus N(\mathcal{B})$. Базису $\mathcal{B}$ и множеству N^* поставим в соответствие:

1) вектор $b(N^*) = b - \sum_{j \in N^*} h[j] a^j$ и вектор $\bar{b}(N^*)$, который получается из вектора $b(N^*)$ добавлением компоненты $b(N^*)[0] = 0$;

2) векторы $b(\mathcal{B}, N^*) = B^{-1}b(N^*)$ и $\bar{b}(\mathcal{B}, N^*) = \bar{B}^{-1}\bar{b}(N^*)$; i -ю компоненту вектора $\bar{b}(\mathcal{B}, N^*)$ обозначим через $b(\mathcal{B}, N^*)[i]$, $i \in 0:m$;

3) план $x(\mathcal{B}, N^*) = \{x(\mathcal{B}, N^*)[j_i] = b(\mathcal{B}, N^*)[i], i \in 1:m, j_i \in N(\mathcal{B}); x(\mathcal{B}, N^*)[j] = h[j], j \in N^*; x(\mathcal{B}, N^*)[j] = 0, j \in N(\mathcal{B}) \cup N^*\}$.

План $x(\mathcal{B}, N^*)$ будем называть **-базисным*. Обращаем особое внимание читателя на то, что N^* — множество тех индексов l небазисных компонент вектора $x(\mathcal{B}, N^*)$, для которых $x(\mathcal{B}, N^*)[l] = h[l]$. Про такие компоненты вектора $x(\mathcal{B}, N^*)$ будем говорить, что они принимают верхние значения. Очевидно, что **-базисный план* $x(\mathcal{B}, N^*)$ является допустимым в том и только том случае, если $0 \leq b(\mathcal{B}, N^*) \leq h(\mathcal{B})$. Допустимый **-базисный план* называется *невыврожденным*, если $0 < b(\mathcal{B}, N^*) < h(\mathcal{B})$, и *выврожденным* в противном случае. Задача (1) называется *невыврожденной*, если она не имеет вырожденных планов.

1.2. Предварительные замечания. Как уже отмечалось, алгоритм настоящего параграфа основан на идеях алгоритма симплекс-метода. Пусть $\mathcal{B}$ — базис задачи (1), $N^* \subset J' \setminus N(\mathcal{B})$ и $x(\mathcal{B}, N^*)$ — допустимый **-базисный план*. Так же, как и в примере п. 1.5.5, можно показать, что $x(\mathcal{B}, N^*)$ — вершина многогранного выпуклого множества $\mathcal{D}$. — Упражнение 5.1. Условия оптимальности допустимого плана для задачи (1) (см. п. 3.1.11, 2°) позволяют установить оптимальность плана $x(\mathcal{B}, N^*)$. Если эти условия не выполняются, то в случае невырожденности задачи можно указать допустимое прогрессивное направление в точке $x(\mathcal{B}, N^*)$. Тогда, как и в алгоритме симплекс-метода, либо обнаружится неограниченность сверху f на $\mathcal{D}$, либо будет осуществ-

лен переход в другую вершину множества $\mathcal{D}$, в которой целевая функция принимает значение больше $f(x(\mathcal{B}, N^*))$.

Здесь следует отметить два обстоятельства. Во-первых, условия оптимальности плана $x(\mathcal{B}, N^*)$ связаны с множеством N^* индексов компонент вектора $x(\mathcal{B}, N^*)$, принимающих верхние значения. Поэтому проверка оптимальности плана $x(\mathcal{B}, N^*)$, которая осуществляется на этапе 1, отличается от аналогичной проверки оптимальности плана $x(\mathcal{B})$ в алгоритме п. 4.4.3. Во-вторых, в случае неоптимальности плана $x(\mathcal{B}, N^*)$ переход в новую вершину множества $\mathcal{D}$ не всегда связан с изменением базиса — может измениться только множество N^* . Возможно также изменение и базиса $\mathcal{B}$, и множества N^* . Эти обстоятельства отражаются в рекомендациях этапов 2 и 3 алгоритма.

Отметим также, что если $h \in R^n$ ($I' = 1:n$) и $\mathcal{D} \neq \emptyset$, то $\mathcal{D}$ — выпуклый многогранник и задача $(\mathcal{D}, f)$ имеет решение, которое в случае невырожденности может быть получено после конечного числа итераций q .

1.3. Алгоритм. Вычисления начинаются с базиса $\mathcal{B}_1$ и множества $N_1^* \subset I' \setminus N(\mathcal{B}_1)$, которым соответствует допустимый план $x(\mathcal{B}_1, N_1^*)$, — см. п. 1.7. Строится последовательность пар $(\mathcal{B}_1, N_1^*), (\mathcal{B}_2, N_2^*), \dots, (\mathcal{B}_q, N_q^*)$, состоящих из базисов и отмеченных подмножеств. В процессе вычислений заполняется таблица

	h
$d_0(\mathcal{B}_1)$	$\bar{A}$
$\vdots$	
$d_0(\mathcal{B}_q)$	$a_0(\mathcal{B}_1)$
	$\vdots$
	$a_0(\mathcal{B}_q)$

(2)

и строится последовательность таблиц вида

		$h [l]$				
$T(\mathcal{B}, N^*) =$	$N(\mathcal{B})$	$h(\mathcal{B})$	$\bar{b}(\mathcal{B}, N^*)$	$\bar{B}^{-1}$	$\bar{a}^l$	$\bar{a}^l(\mathcal{B})$

Этап 0. В табл. (2) заполняются матрица $\bar{A}$ и строка h . Вычисляются матрица $\bar{B}_1^{-1}$ и вектор $\bar{b}(\mathcal{B}_1, N_1^*) = \bar{B}_1^{-1}(\bar{b} - \sum_{l \in N^*} h[l] \bar{a}^l)$. Полагаем $\mathcal{B} = \mathcal{B}_1$ и $N^* = N_1^*$. В табл. $T(\mathcal{B}, N^*)$ заполняются столбцы $N(\mathcal{B})$, $\bar{b}(\mathcal{B}, N^*)$ и матрица $\bar{B}^{-1}$.

На этапах 1—3 будет описана одна итерация метода: проверяется оптимальность текущего $\ast$ -базисного плана $x(\mathcal{B}, N^*)$, а в случае неоптимальности плана $x(\mathcal{B}, N^*)$ осуществляется переход к новому базису $\mathcal{B}'$, множеству N^* и отвечающему им $\ast$ -базисному плану $x(\mathcal{B}', N^{*'})$.

Этап 1. 1. В верхних углах клеток строки $a_0(\mathcal{B})$ табл. (2), имеющих номера $j \in N^*$, проставляются отметки $\ast$. Из табл. $T(\mathcal{B}, N^*)$ в табл. (2) переписывается верхняя строка $d_0(\mathcal{B})$ матрицы $\bar{B}^{-1}$. В табл. (2) заполняется строка $a_0(\mathcal{B}) = d_0(\mathcal{B})\bar{A}$.

2. Рассматриваются множества индексов

$$J^+(\mathcal{B}, N^*) = \{j \in N^* : a(\mathcal{B})[0, j] > 0\},$$

$$J^-(\mathcal{B}, N^*) = \{j \in 1:n \setminus (N(\mathcal{B}) \cup N^*) : a(\mathcal{B})[0, j] < 0\}$$

и $J(\mathcal{B}, N^*) = J^+(\mathcal{B}, N^*) \cup J^-(\mathcal{B}, N^*)$. Возможны два случая.

2'. $J(\mathcal{B}, N^*) = \emptyset$. Вычислительный процесс закончен: $x(\mathcal{B}, N^*)$ — оптимальный план. В столбце $N(\mathcal{B})$ табл. $T(\mathcal{B}, N^*)$ расположены номера его базисных компонент, рядом, в столбце $\bar{b}(\mathcal{B}, N^*)$, — их значения. Из табл. (2) можно выписать значения небазисных переменных: $x(\mathcal{B}, N^*)[j] = h[j]$ для $j \in N^*$ и $x(\mathcal{B}, N^*)[j] = 0$ для $j \in 1:n \setminus (N(\mathcal{B}) \cup N^*)$. Максимальное значение целевой функции $f(x(\mathcal{B}, N^*)) = b(\mathcal{B}, N^*)[0]$.

2''. $J(\mathcal{B}, N^*) \neq \emptyset$. Выбирается некоторое $l \in J(\mathcal{B}, N^*)$. Из табл. (2) в табл. $T(\mathcal{B}, N^*)$ переписываются $h[l]$ и векторы $\bar{a}^l$ и $h(\mathcal{B})$. Вычисляется вектор $\bar{a}^l(\mathcal{B}) = \bar{B}^{-1}\bar{a}^l$.

Обозначим $I_l^- = \{i \in 1:m : a(\mathcal{B})[i, l] < 0\}$ и $I_l^+ = \{i \in 1:m : a(\mathcal{B})[i, l] > 0\}$. Если l выбрано из множества $J^+(\mathcal{B}, N^*)$, то переходим к этапу 2, если $l \in J^-(\mathcal{B}, N^*)$, переходим к этапу 3.

Этап 2. Вычисляется

$$\lambda(\mathcal{B}, N^*) = \min \left\{ \min_{i \in I_l^-} \frac{b(\mathcal{B}, N^*)[i]}{-a(\mathcal{B})[i, l]}, h[l], \min_{i \in I_l^+} \frac{h[l] - b(\mathcal{B}, N^*)[i]}{a(\mathcal{B})[i, l]} \right\} \quad (3)$$

(если какое-либо из множеств I_l^- или I_l^+ пусто, то соответствующий минимум не вычисляется). Для невырожденной задачи возможен один и только один из следующих трех случаев.

2а) Существует единственное $r \in I_l^-$, для которого

$$\lambda(\mathcal{B}, N^*) = b(\mathcal{B}, N^*)[r] / -a(\mathcal{B})[r, l]. \quad (4)$$

Тогда формируются новый базис $\mathcal{B}'$, для которого $N(\mathcal{B}') = (N(\mathcal{B}) \setminus \{r\}) \cup \{l\}$, и множество $N^{*'} = N^* \setminus \{l\}$. В табл. $T(\mathcal{B}', N^{*'})$ заполняется столбец $N(\mathcal{B}')$. Элементы столбца $\bar{b}(\mathcal{B}', N^{*'})$ и

матрицы $\bar{B}'^{-1}$ заполняются по формулам

$$b(\mathcal{B}', N^*)[r] = h[l] - \lambda(\mathcal{B}, N^*); \quad (5)$$

$$b(\mathcal{B}', N^*)[i] = b(\mathcal{B}, N^*)[i] + \lambda(\mathcal{B}, N^*)a(\mathcal{B})[i, l], \\ i \in 0:m, \quad i \neq r, \quad (6)$$

$$d(\mathcal{B}')[r, j] = d(\mathcal{B})[r, j]/a(\mathcal{B})[r, l], \quad j \in 0:m, \quad (7)$$

$$d(\mathcal{B}')[i, j] = d(\mathcal{B})[i, j] - d(\mathcal{B}')[r, j]a(\mathcal{B})[i, l], \\ i \in 0:m, \quad i \neq r, \quad j \in 0:m. \quad (8)$$

Полагаем $\mathcal{B} = \mathcal{B}'$, $N^* = N^*$ и переходим к этапу 1.

$$2б) \quad \lambda(\mathcal{B}, N^*) = h[l]. \quad (9)$$

Тогда $\mathcal{B}' = \mathcal{B}$, $N(\mathcal{B}') = N(\mathcal{B})$, а $N^* = N^* \setminus \{l\}$. Из табл. $T(\mathcal{B}, N^*)$ в табл. $T(\mathcal{B}', N^*)$ переписываются столбец $N(\mathcal{B})$ и матрица $\bar{B}^{-1}$. Элементы столбца $\bar{b}(\mathcal{B}', N^*)$ вычисляются по формулам

$$b(\mathcal{B}', N^*)[i] = b(\mathcal{B}, N^*)[i] + \lambda(\mathcal{B}, N^*)a(\mathcal{B})[i, l], \quad i \in 0:m. \quad (10)$$

Полагаем $\mathcal{B} = \mathcal{B}'$, $N^* = N^*$ и переходим к этапу 1.

2в) Существует единственное $r \in I_l^+$, для которого

$$\lambda(\mathcal{B}, N^*) = (h[j_r] - b(\mathcal{B}, N^*)[r]) / a(\mathcal{B})[r, l]. \quad (11)$$

Тогда формируются новый базис $\mathcal{B}'$, для которого $N(\mathcal{B}') = (N(\mathcal{B}) \setminus \{j_r\}) \cup \{l\}$, и множество $N^* = (N^* \setminus \{l\}) \cup \{j_r\}$. В табл. $T(\mathcal{B}', N^*)$ заполняется столбец $N(\mathcal{B}')$. Элементы столбца $\bar{b}(\mathcal{B}', N^*)$ и матрицы $\bar{B}'^{-1}$ заполняются по формулам (5) — (8). Полагаем $\mathcal{B} = \mathcal{B}'$, $N^* = N^*$ и переходим к этапу 1.

Этап 3. Вычисляется

$$\lambda'(\mathcal{B}, N^*) = \min \left\{ \min_{i \in I_l^+} \frac{b(\mathcal{B}, N^*)[i]}{a(\mathcal{B})[i, l]}, \quad h[l], \right. \\ \left. \min_{i \in I_l^-} \frac{h[j_i] - b(\mathcal{B}, N^*)[i]}{-a(\mathcal{B})[i, l]} \right\} \quad (12)$$

(если какое-либо из множеств I_l^- или I_l^+ пусто, то соответствующий минимум не вычисляется).

Если $a(\mathcal{B})[i, l] \leq 0$, $i \in 1:m$, $h[l] = +\infty$ и $h[j_i] = +\infty$ для $i \in I_l^-$, то $\lambda'(\mathcal{B}, N^*) = +\infty$ и целевая функция f не ограничена сверху на $\mathcal{D}$. Естественно, что в этом случае вычисления прекращаются.

Если $\lambda'(\mathcal{B}, N^*) < +\infty$, то для невырожденной задачи возможен один и только один из следующих трех случаев.

3а) Существует единственное $r \in I_1^+$, для которого

$$\lambda'(\mathcal{B}, N^*) = b(\mathcal{B}, N^*)[r]/a(\mathcal{B})[r, l]. \quad (13)$$

Тогда формируются новый базис $\mathcal{B}'$, для которого $N(\mathcal{B}') = (N(\mathcal{B}) \setminus \{j_r\}) \cup \{l\}$, и множество $N^{*'} = N^*$. В табл. $T(\mathcal{B}', N^{*'})$ заполняется столбец $N(\mathcal{B}')$. Элементы столбца $\bar{b}(\mathcal{B}', N^{*'})$ заполняются по формулам

$$b(\mathcal{B}', N^{*'})[r] = \lambda'(\mathcal{B}, N^*), \quad (14)$$

$$b(\mathcal{B}', N^{*'})[i] = b(\mathcal{B}, N^*)[i] - \lambda'(\mathcal{B}, N^*)a(\mathcal{B})[i, l], \\ i \in 0:m, \quad i \neq r. \quad (15)$$

Элементы матрицы $\bar{B}'^{-1}$ заполняются по формулам (7), (8). Полагаем $\mathcal{B} = \mathcal{B}'$, $N^* = N^{*'}$ и переходим к этапу 1.

$$3б) \quad \lambda'(\mathcal{B}, N^*) = h[l]. \quad (16)$$

Тогда $\mathcal{B}' = \mathcal{B}$, $N(\mathcal{B}') = N(\mathcal{B})$, а $N^{*' = N^* \cup \{l\}$. Из табл. $T(\mathcal{B}, N^*)$ в табл. $T(\mathcal{B}', N^{*'})$ переписываются столбец $N(\mathcal{B})$ и матрица $\bar{B}^1$. Элементы столбца $\bar{b}(\mathcal{B}', N^{*'})$ вычисляются по формулам

$$b(\mathcal{B}', N^{*'})[i] = b(\mathcal{B}, N^*)[i] - \lambda'(\mathcal{B}, N^*)a(\mathcal{B})[i, l], \quad i \in 0:m. \quad (17)$$

Полагаем $\mathcal{B} = \mathcal{B}'$, $N^* = N^{*'}$ и переходим к этапу 1.

3в) Существует единственное $r \in I_1^-$, для которого

$$\lambda'(\mathcal{B}, N^*) = (h[j_r] - b(\mathcal{B}, N^*)[r])/(-a(\mathcal{B})[r, l]). \quad (18)$$

Тогда формируются новый базис $\mathcal{B}'$, для которого $N(\mathcal{B}') = (N(\mathcal{B}) \setminus \{j_r\}) \cup \{l\}$, и множество $N^{*' = N^* \cup \{j_r\}$. В табл. $T(\mathcal{B}', N^{*'})$ заполняется столбец $N(\mathcal{B}')$. Элементы столбца $\bar{b}(\mathcal{B}', N^{*'})$ заполняются по формулам (14), (15), а элементы матрицы $\bar{B}'^{-1}$ — по формулам (7), (8). Полагаем $\mathcal{B} = \mathcal{B}'$, $N^* = N^{*'}$ и переходим к этапу 1. ■

Предлагаем читателю самостоятельно продумать рациональную организацию вычислений по алгоритму. — Упражнение 5.2. В первую очередь это относится к вычислениям, связанным со случаями 2б) и 3б).

1.4. Обсуждение. Как и в п. 4.3.4, обращаем внимание читателя на следующие две проблемы. Первая из них — рациональный выбор l из множества $J(\mathcal{B}, N^*)$. Обычно рекомендуют выбирать l так, чтобы

$$|a(\mathcal{B})[0, l]| = \max \{|a(\mathcal{B})[0, j]|, \quad j \in J(\mathcal{B}, N^*)\}.$$

Вторая — решение вырожденных задач. Здесь можно было бы лишь повторить сказанное в п. 4.3.4 и § 7 гл. 4,

1.5. Пример. В табл. (19) и (20) приведено решение задачи

$$(\mathcal{D}, f): f(x) = 2x_1 + x_2 + 2x_3 + 3x_4 - 12x_5 \rightarrow \max,$$

$$\mathcal{D} = \{x \in R^5: x_1 + 2x_2 + 3x_3 - x_4 + 2x_5 = 5,$$

$$2x_1 + 3x_2 + 4x_3 + x_4 = 10,$$

$$0 \leq x_1 \leq 4, 0 \leq x_2 \leq 3, 0 \leq x_3 \leq 1, 0 \leq x_4 \leq 2, 0 \leq x_5 \leq 1\}.$$

Оно начинается с базиса $\mathcal{B}_1 = \{a^1, a^2\}$ и множества $N_1^* = \{4, 5\}$, которым отвечает допустимый $\#$ -базисный план $x(\mathcal{B}_1, N_1^*) = (1, 2, 0, 2, 1)$. Число итераций $q = 4$.

			4	3	1	2	1
-1	-4	3	2	1	2	3	-12
-1	-8/5	7/5	1	2	3	-1	2
-1	-8/5	7/5	2	3	4	1	0
-1	-10/7	11/7	0	0	-2	4*	4*
			-4/5*	0	-6/5	0	44/5*
			-4/5*	0	-6/5	0	44/5
			-2/7*	6/7	0	0	64/7

(19)

		2	-2	-1	-4	3	3	4
$T(\mathcal{B}_1, N_1^*) =$	1	4	1	0	-3	2	-1	5
	2	3	2	0	2	-1	1	-3
		1	2/5	-1	-8/5	7/5	-12	44/5
$T(\mathcal{B}_2, N_2^*) =$	4	2	7/5	0	-3/5	2/5	2	-6/5
	2	3	1/5	0	1/5	1/5	0	2/5
		1	46/5	-1	-8/5	7/5	2	-6/5
$T(\mathcal{B}_3, N_3^*) =$	4	2	1/5	0	-3/5	2/5	3	-1/5
	2	3	3/5	0	1/5	1/5	4	7/5
			68/7	-1	-10/7	11/7		
$T(\mathcal{B}_4, N_4^*) =$	4		2/7	0	-4/7	3/7		
	3		3/7	0	1/7	1/7		

(20)

В табл. (19) заполняются матрица $\bar{A}$ и строка h . Вычисляются матрица $\bar{B}_1^{-1}$, векторы $\bar{b}(N_1^*) = b - 2\bar{a}^4 - \bar{a}^5 = (6, 5, 8)^T$ и $\bar{b}(\mathcal{B}_1, N_1^*) = \bar{B}_1^{-1} \bar{b}(N_1^*) = (-2, 1, 2)^T$. В табл. $T(\mathcal{B}_1, N_1^*)$ заполняются матрица $\bar{B}_1^{-1}$, столбцы $N(\mathcal{B}_1)$ и $\bar{b}(\mathcal{B}_1, N_1^*)$.

В табл. (19) из табл. $T(\mathcal{B}_1, N_1^*)$ переписывается строка $d_0(\mathcal{B}_1) = (-1, -4, 3)$. Вычисляется строка $a_0(\mathcal{B}_1) = d_0(\mathcal{B}_1) \bar{A}$. У четвертого и пятого элементов этой строки проставляются отметки *. Тогда $J^-(\mathcal{B}_1, N_1^*) = \{3\}$, $J^+(\mathcal{B}_1, N_1^*) = \{4, 5\}$, $J(\mathcal{B}_1, N_1^*) = \{3, 4, 5\} \neq \emptyset$. Следовательно, план $x(\mathcal{B}_1, N_1^*)$ неоптимален. В соответствии с приведенной в п. 1.4 рекомендацией полагаем $l=4$. В табл. $T(\mathcal{B}_1, N_1^*)$ заполняются столбец $h(\mathcal{B}_1) = (4, 3)^T$ и элемент $h[4] = 2$. Из табл. 19 переписывается $\bar{a}^4$ и вычисляется $\bar{a}^4(\mathcal{B}_1) = (4, 5, -3)^T$. Так как $l \in J^+(\mathcal{B}_1, N_1^*)$, то по формуле (3) вычисляется $\lambda(\mathcal{B}_1, N_1^*) = \min\{2/3, 2, 3/5\} = 3/5$. Осуществился случай 2в), $r=1$. Следовательно, $N(\mathcal{B}_2) = \{4, 2\}$, $N_2^* = \{1, 5\}$. Элементы столбца $\bar{b}(\mathcal{B}_2, N_2^*)$ и матрицы $\bar{B}_2^{-1}$ в табл. $T(\mathcal{B}_2, N_2^*)$ заполняются по формулам (5) — (8).

Дальнейшие вычисления, проведенные в соответствии с рекомендациями алгоритма, также отражены в табл. (19), (20). На второй итерации $\lambda(\mathcal{B}_2, N_2^*) = h[5]$ (случай 2б)), так что $\mathcal{B}_3 = \mathcal{B}_2$, $N_3^* = \{1\}$. На третьей итерации $\lambda(\mathcal{B}_3, N_3^*) = 3/7$ (случай 3в)). На четвертой установлена оптимальность плана $x(\mathcal{B}_4, N_4^*) = (4, 0, 3/7, 2/7, 0)$, отвечающего базису $\mathcal{B}_4 = \{a^4, a^3\}$ и множеству $N_4^* = \{1\}$. Максимальное значение целевой функции $z^*/7$.

1.6. Обоснование алгоритма. Как уже отмечалось, алгоритм настоящего параграфа является модификацией симплекс-метода. Точнее, этот алгоритм является реализацией идеи и принципов конечных методов допустимых направлений (см. § 2 гл. 4), основанной на применении вычислительной схемы симплекс-метода. Чтобы убедиться в этом, достаточно проверить утверждения 1°—12°, формулировки которых приведены ниже. Некоторые из этих утверждений фактически были доказаны в п. 4.3.7, доказательства других лишь технически отличаются от доказательств аналогичных утверждений п. 4.3.6. Проверку утверждений 1°—12° — громоздкое, но, по-видимому, полезное упражнение — предоставляем читателю.

Пусть $\mathcal{B}$ — базис задачи (1), $N(\mathcal{B}) = \{j_1, \dots, j_m\}$, $N^* \subset J' \setminus N(\mathcal{B})$ и $x(\mathcal{B}, N^*)$ (см. 3) в п. 1.1) — допустимый π -базисный план. Для каждого $l \in 1:n \setminus N(\mathcal{B})$ определим вектор $s^l(\mathcal{B}) \in R^n$, компоненты которого вычисляются по формулам

$$\begin{aligned} s^l(\mathcal{B})[j_i] &= -a(\mathcal{B})[i, l], \quad i \in 1:m; \quad s^l(\mathcal{B})[l] = 1; \\ s^l(\mathcal{B})[j] &= 0, \quad j \in 1:n \setminus (N(\mathcal{B}) \cup \{l\}) \end{aligned}$$

(см. (13) в п. 4.3.6).

1°. Если задача $(\mathcal{D}, f)$ невырождена, то для любого $l \in 1:n \setminus (N(\mathcal{B}) \cup N^*)$ вектор $s^l(\mathcal{B})$ и для любого $l \in N^*$ вектор $-s^l(\mathcal{B})$ являются допустимыми направлениями в точке $x(\mathcal{B}, N^*)$.

2°. Пусть $l \in 1:n \setminus (N(\mathcal{B}) \cup N^*)$, $a^l(\mathcal{B}) \leq 0$, $h[l] = +\infty$ и $h[j_i] = +\infty$, если $a(\mathcal{B})[i, l] < 0$. Тогда $s^l(\mathcal{B})$ — допустимое направление в точке $x(\mathcal{B}, N^*)$ и $x(\mathcal{B}, N^*) + \lambda s^l(\mathcal{B}) \in \mathcal{D}$ при любом $\lambda \geq 0$.

3°. Для любого $l \in 1:n \setminus N(\mathcal{B})$ справедливо равенство $cs^l(\mathcal{B}) = -a(\mathcal{B})[0, l]$ и вектор $s^l(\mathcal{B})$ (вектор $-s^l(\mathcal{B})$) в том и только том случае является прогрессирующим направлением в точке $x(\mathcal{B}, N^*)$, если $a(\mathcal{B})[0, l] < 0$ ($a(\mathcal{B})[0, l] > 0$).

4°. Если $a(\mathcal{B})[0, l] \geq 0$ для $l \in 1:n \setminus N^*$ и $a(\mathcal{B})[0, l] \leq 0$ для $l \in N^*$, то $x(\mathcal{B}, N^*)$ — решение задачи $(\mathcal{D}, f)$.

5°. $f(x(\mathcal{B}, N^*)) = b(\mathcal{B}, N^*)[0]$.

6°. Пусть $l \in 1:n \setminus (N(\mathcal{B}) \cup N^*)$, $a^l(\mathcal{B}) \leq 0$, $h[l] = +\infty$, $h[j_i] = +\infty$, если $a^l(\mathcal{B})[i, l] < 0$, и $a(\mathcal{B})[0, l] < 0$. Тогда f не ограничена сверху на $\mathcal{D}$.

7°. Пусть $l \in N^*$ и число $\lambda(\mathcal{B}, N^*)$ определяется формулой (3). Тогда $\lambda(\mathcal{B}, N^*) = \max\{\lambda : x(\mathcal{B}, N^*) - \lambda s^l(\mathcal{B}) \in \mathcal{D}, \lambda \geq 0\}$ и $\lambda(\mathcal{B}, N^*) > 0$ для невырожденной задачи. Если к тому же $a(\mathcal{B})[0, l] > 0$, то $f(x(\mathcal{B}, N^*) - \lambda(\mathcal{B}, N^*)s^l(\mathcal{B})) > f(x(\mathcal{B}, N^*))$.

Пусть $l \in 1:n \setminus (N(\mathcal{B}) \cup N^*)$ и $\lambda'(\mathcal{B}, N^*) < +\infty$ (см. (12), а также 2°). Тогда $\lambda'(\mathcal{B}, N^*) = \max\{\lambda : x(\mathcal{B}, N^*) + \lambda s^l(\mathcal{B}) \in \mathcal{D}, \lambda \geq 0\}$ и $\lambda'(\mathcal{B}, N^*) > 0$ для невырожденной задачи. Если к тому же $a(\mathcal{B})[0, l] < 0$, то $f(x(\mathcal{B}, N^*) + \lambda'(\mathcal{B}, N^*)s^l(\mathcal{B})) > f(x(\mathcal{B}, N^*))$.

8°. Пусть $l \in 1:n \setminus N(\mathcal{B})$, $r \in 1:m$ и $a(\mathcal{B})[r, l] \neq 0$. Тогда система векторов $\mathcal{B}' = (\mathcal{B} \setminus \{a^l_r\}) \cup \{a^l\}$ — базис задачи $(\mathcal{D}, f)$.

9°. В каждом из случаев 2а) — 2в) вектор $x(\mathcal{B}', N^*)$ — допустимый $*$ -базисный план и

$$x(\mathcal{B}', N^*) = x(\mathcal{B}, N^*) - \lambda(\mathcal{B}, N^*)s^l(\mathcal{B}).$$

В каждом из случаев 3а) — 3в) вектор $x(\mathcal{B}', N^*)$ — допустимый $*$ -базисный план и

$$x(\mathcal{B}', N^*) = x(\mathcal{B}, N^*) + \lambda'(\mathcal{B}, N^*)s^l(\mathcal{B}).$$

10°. В каждом из случаев этапов 2 и 3 элементы столбца $\bar{b}(\mathcal{B}', N^*)$ и матрицы $\bar{B}'^{-1}$ вычисляются в соответствии с приведенными в этих этапах рекомендациями.

11°. Пусть задача (1) невырождена. Если $l \in N^*$, то имеет место один и только один из случаев 2а) — 2в) и в каждом случае 2а) и 2в) формулы (4) и (11) определяют r однозначно. Если $l \in 1:n \setminus (N(\mathcal{B}) \cup N^*)$ и $\lambda'(\mathcal{B}, N^*) < +\infty$, то имеет место один и только один из случаев 3а) — 3в) и в каждом случае 3а) и 3в) формулы (13) и (18) определяют r однозначно.

12°. Для невырожденной задачи число итераций q конечно.

1.7. Построение начального допустимого $\ast$ -базисного плана. Для нахождения базиса $\mathcal{B}_1$ и множества N_1^* , которым отвечает допустимый $\ast$ -базисный план $x(\mathcal{B}_1, N_1^*)$, воспользуемся методом минимизации невязок (см. п. 4.3.8). Не ограничивая общности, можно считать, что $b \geq 0$. Обозначим через $A' \in m \times (m+n)$ матрицу $[A']$ и рассмотрим задачу

$$\begin{aligned} (\mathcal{D}, \psi): \psi(z) &= - \sum_{j=n+1}^{n+m} z[j] \rightarrow \min, \\ \mathcal{D} &= \{z \in R^{n+m}: A'z = b; 0 \leq z[j] \leq h[j], j \in 1:n; \\ &\quad z[j] \geq 0, j \in n+1:n+m\}. \end{aligned} \quad (21)$$

Очевидно, что базису $\mathcal{B}_1$ этой задачи, для которого $N(\mathcal{B}_1) = \{n+1, \dots, n+m\}$, и множеству $N_1^* = \emptyset$ отвечает допустимый $\ast$ -базисный план $z(\mathcal{B}_1, N_1^*) = (0, \dots, 0, b[1], \dots, b[m])$. Пусть задача (1) невырождена. Тогда применение алгоритма п. 1.3 позволит после конечного числа итераций получить оптимальный $\ast$ -базисный план $\bar{z} = z(\bar{\mathcal{B}}, \bar{N}^*)$. Если $\psi(\bar{z}) = 0$, то $\bar{\mathcal{B}}$ — базис задачи $(\mathcal{D}, f)$, $\bar{N}^* \subset 1:n$ и $\bar{x} = (\bar{z}[1], \dots, \bar{z}[n])$ — допустимый $\ast$ -базисный план задачи $(\mathcal{D}, f)$. Если $\psi(\bar{z}) < 0$, то $\mathcal{D} = \emptyset$.

Определенные трудности могут возникнуть, когда задача (21) вырождена. Может случиться, например, что $\psi(\bar{z}) = 0$, а $N(\bar{\mathcal{B}})$ не содержится в множестве $1:n$. В этом случае очевидно $x \in \mathcal{D}$ и система столбцов матрицы A , отвечающих тем компонентам $\bar{z}[j]$ вектора $\bar{x}$, для которых $0 < \bar{z}[j] < h[j]$, линейно независима. Поэтому для построения базиса $\mathcal{B}_1$ достаточно любым способом дополнить эту систему столбцов до базиса задачи $(\mathcal{D}, f)$.

§ 2. О задачах квадратичного программирования и одном методе их решения

В классе нелинейных задач выпуклого программирования задачи квадратичного программирования занимают особое место. Некоторая специфика этих задач: область определения — многогранное выпуклое множество, сравнительная простота задания целевой функции, возможность прямого вычисления ее значений и значений вектора градиента этой функции позволяют строить довольно простые алгоритмы решения таких задач. В настоящее время разработано большое число различных алгоритмов решения задач квадратичного программирования, достаточно полный обзор которых приведен в [29]. Здесь мы рассмотрим только один из них — метод Вульфа. Наш выбор связан не с какими-либо преимуществами этого метода перед другими, а прежде всего с возможностью иллюстрации прямого использования симплекс-метода для решения нелинейных задач.

2.1. Обозначения и некоторые свойства. Определение задачи квадратичного программирования было дано в п. 2.2.3. Здесь будем рассматривать каноническую форму задачи квадратичного программирования (см. упражнение 2.11), т. е. экстремальную задачу

$$(\mathcal{D}, f) : f(x) = cx + x^T D x \rightarrow \max, \mathcal{D} = \{x \in R^n : Ax = b, x \geq 0\}, \quad (1)$$

где $D = (d[i, j])$ — неотрицательно определенная матрица порядка n , $A = (a[i, j])$ — $m \times n$ матрица ограничений, ранг которой $r(A) = n$.

В общем случае задача $(\mathcal{D}, f)$ может не иметь решения: если $\mathcal{D}$ — неограниченное множество и матрица D не является отрицательно определенной, то функция f может быть неограниченной на $\mathcal{D}$. — Упражнение 5.3. Однако справедлива следующая теорема.

Теорема. Если $\mathcal{D} \neq \emptyset$ и матрица D отрицательно определена, то задача (1) имеет единственное решение.

Доказательство. Из отрицательной определенности матрицы D следует строгая вогнутость функции f . — Упражнение 5.4. Поэтому если задача $(\mathcal{D}, f)$ имеет решение, то оно единственно (см. п. 1.8.2, 2°). Докажем существование решения задачи.

Так как матрица D отрицательно определена, то функция $|cx/x^T Dx|$ непрерывна на компактном множестве $S = \{x \in R^n : \|x\| = 1\}$. Положим $m_1 = \max_{x \in S} x^T Dx$, $m_2 = \max_{x \in S} |cx/x^T Dx|$. Тогда $m_1 < 0$ и $x^T Dx(1 + cx/rx^T Dx) \leq m_1(1 - m_2/r)$ для любого $x \in S$, если $m_2/r < 1$.

Пусть $x_0 \in \mathcal{D}$. Выберем такое число $r_0 > 0$, чтобы выполнялись неравенства $\|x_0\| \leq r_0$, $m_2/r_0 < 1$ и $r_0^2 m_1(1 - m_2/r_0) \leq f(x_0)$. Функция f непрерывна на непустом компактном множестве $\bar{\mathcal{D}}(0, r_0) \cap \mathcal{D}$ ($\bar{\mathcal{D}}(0, r)$ — замкнутый шар с центром в нуле и радиусом r_0) и, следовательно, принимает наибольшее на этом множестве значение в некоторой точке $\bar{x} \in \bar{\mathcal{D}}(0, r_0) \cap \mathcal{D}$. Покажем, что $f(\bar{x}) = \max_{x \in \mathcal{D}} f(x)$. Очевидно, что $f(z) \leq f(\bar{x})$, если $z \in \bar{\mathcal{D}}(0, r_0) \cap \mathcal{D}$. Пусть $z \in \mathcal{D} \setminus \bar{\mathcal{D}}(0, r_0)$. Тогда $\|z\| > r_0$. Положим $z' = z/\|z\|$. Так как $z' \in S$, то в силу выбора числа r_0 имеем

$$\begin{aligned} f(z) &= cz + z^T Dz = \|z\|^2 z'^T Dz'(1 + cz'/\|z\| z'^T Dz') \leq \\ &\leq \|z\|^2 m_1(1 - m_2/\|z\|) \leq r_0^2 m_1(1 - m_2/r_0) \leq f(x_0) \leq f(\bar{x}). \quad \blacksquare \end{aligned}$$

2.2. Преобразование задачи. Покажем теперь, что решение задачи квадратичного программирования $(\mathcal{D}, f)$ сводится к поиску решения некоторой системы, состоящей из линейных уравнений и неравенств и одного нелинейного ограничения. Для

этого воспользуемся условиями оптимальности, сформулированными в п. 3.1.9: для оптимальности $x \in \mathcal{D}$ необходимо и достаточно существование таких $u \in R^m$ и $v \in R^n$, что

$$2Dx - A^T u + I_n v = -c, \quad (2)$$

$$v \geq 0, \quad (3)$$

$$vx = 0, \quad (4)$$

где I_n — единичная матрица порядка n . Поэтому задача $(\mathcal{D}, f)$ имеет решение в том и только том случае, если совместна система

$$Ax = b, \quad (5)$$

$$2Dx - A^T u + I_n v = -c, \quad (6)$$

$$x \geq 0, \quad (7)$$

$$v \geq 0, \quad (8)$$

$$vx = 0 \quad (9)$$

и для каждого решения (x, u, v) системы (5) — (9) вектор x является решением задачи $(\mathcal{D}, f)$.

Действительно, если x — решение задачи $(\mathcal{D}, f)$, то $x \in \mathcal{D}$ и найдутся такие u и v , что для x, u, v выполняются соотношения (2) — (4), а значит, и все соотношения (5) — (9). Обратно, если (x, u, v) — решение системы (5) — (9), то для x выполняются условия (5), (7). Следовательно, $x \in \mathcal{D}$ и найдены такие u и v , что выполняются условия оптимальности (2) — (4). ■

Таким образом, поиск решения задачи $(\mathcal{D}, f)$ сводится к поиску решения системы (5) — (9). Отметим, что если $\mathcal{D} \neq \emptyset$ и матрица D отрицательно определена, то в силу теоремы п. 2.1 задача $(\mathcal{D}, f)$ имеет решение, и, следовательно, система (5) — (9) совместна.

2.3. Применение процедуры симплекс-метода для решения задачи квадратичного программирования (метод Вульфа). Вычислительная процедура, приведенная в настоящем пункте, позволяет находить оптимальный план задачи квадратичного программирования $(\mathcal{D}, f)$ в случае отрицательной определенности матрицы D и непустоты множества допустимых планов $\mathcal{D}$. Если матрица D неположительно определена, но не является отрицательно определенной, то можно решать задачу — и во многих случаях это приводит к желаемому результату — так же, как и в случае отрицательной определенности матрицы D .

Рассмотрим систему

$$Ax = b, \quad (10)$$

$$2Dx - A^T y' + A^T y'' + I_n v = -c, \quad (11)$$

$$x \geq 0, \quad y' \geq 0, \quad y'' \geq 0, \quad v \geq 0, \quad (12)$$

$$xv = 0. \quad (13)$$

Очевидно, что если (x, y', y'', v) — ее решение, то $(x, y' - y'', v)$ — решение системы (5) — (9), и обратно, по любому решению системы (5) — (9) можно построить решение системы (10) — (13). Таким образом, проблема нахождения оптимального плана задачи $(\mathcal{D}, f)$ свелась к поиску вектора (x, y', y'', v) , удовлетворяющего условиям (10) — (13), или, другими словами, к поиску решения системы линейных уравнений и неравенств (10) — (12), удовлетворяющего условию (13).

Вулф предложил метод решения системы (10) — (13), основанный на применении процедуры симплекс-метода. В дальнейшем при описании метода Вулфа используются понятия и обозначения п. 4.3.1.

Метод состоит из двух этапов. На первом этапе строится базисный план $x_0 \in \mathcal{D}$. Если $\mathcal{D} \neq \emptyset$, то план x_0 можно найти методом минимизации невязок (см. пп. 4.3.8 или 4.4.8).

Пусть $(-c - 2Dx_0)[k]$ — k -я компонента вектора $-c - 2Dx_0$. Для $k \in 1:n$ положим $z_0[k] = |(-c - 2Dx_0)[k]|$. Обозначим через $\mathcal{I}$ диагональную матрицу, k -й диагональный элемент которой равен 1, если $(-c - 2Dx_0)[k] \geq 0$, и равен -1 , если $(-c - 2Dx_0)[k] < 0$. Рассмотрим следующую каноническую задачу линейного программирования:

$$\begin{aligned} (\mathcal{G}, g) : g(w) &= - \sum_{k=1}^n z[k] \rightarrow \min, \\ \mathcal{G} &= \{w = (x, y', y'', v, z) \in R^{2m+3n} : \\ Ax &= b, \\ 2Dx - A^T y' + A^T y'' + Iv + \mathcal{I}z &= -c, \\ x \geq 0, \quad y' \geq 0, \quad y'' \geq 0, \quad v \geq 0, \quad z \geq 0\}. \end{aligned}$$

Матрицу ограничений задачи $(\mathcal{G}, g)$ обозначим через G . Очевидно, что $w_0 = (x_0, 0, 0, 0, z_0)$ — допустимый базисный план $w(\mathcal{B}_0)$ задачи $(\mathcal{G}, g)$, где $\mathcal{B}_0$ — допустимый базис задачи $(\mathcal{G}, g)$, составленный из столбцов матрицы G , отвечающих базисным компонентам вектора x_0 и всем компонентам вектора z .

На втором этапе решается задача $(\mathcal{G}, g)$. Вычисления начинаются с допустимого базиса $\mathcal{B}_1$ и допустимого базисного плана $w(\mathcal{B}_1) = (x(\mathcal{B}_1), y'(\mathcal{B}_1), y''(\mathcal{B}_1), v(\mathcal{B}_1), z(\mathcal{B}_1))$ задачи $(\mathcal{G}, g)$, обладающих следующими свойствами: *столбцы матрицы ограничений G задачи $(\mathcal{G}, g)$, отвечающие j -й компоненте вектора x и j -й компоненте вектора v , не могут одновременно входить в базис $\mathcal{B}_1$* . Отметим, что допустимый базис $\mathcal{B}_0$, отвечающий вектору w_0 , этим свойством очевидно обладает, поскольку он не содержит столбцов, отвечающих компонентам вектора v . Дальнейшие вычисления производятся в соответствии с процедурой алгоритма симплекс-метода с учетом следующего правила ввода:

если $\mathcal{B}$ — допустимый базис задачи, полученный на некоторой итерации алгоритма, и $J^-(\mathcal{B}) = \{k: a(\mathcal{B}) [0, k] < 0\} \neq \emptyset$, то столбец матрицы G , отвечающий l -й компоненте вектора x , индекс которого входит в множество $J^-(\mathcal{B})$, разрешается вводить в базис только в том случае, если в базис не входит столбец, отвечающий l -й компоненте вектора v ;

столбец матрицы G , отвечающий l -й компоненте вектора v , индекс которой входит в множество $J^-(\mathcal{B})$, разрешается вводить в базис только в том случае, если в него не входит столбец, отвечающий l -й компоненте вектора x .

Очевидно, что это правило ввода вместе с правилом выбора начального базиса обеспечивает на каждой итерации выполнение равенства $x(\mathcal{B})v(\mathcal{B}) = 0$, где $x(\mathcal{B})$ и $v(\mathcal{B})$ — векторы, составленные из компонент вектора $w(\mathcal{B})$, отвечающих индексам векторов x и v . Поэтому если на некоторой итерации алгоритма симплекс-метода, в котором учтены эти правила, будет построен такой базис $\overline{\mathcal{B}}$, что $g(w(\overline{\mathcal{B}})) = 0$, то $(x(\overline{\mathcal{B}}), y'(\overline{\mathcal{B}}), y''(\overline{\mathcal{B}}), v(\overline{\mathcal{B}}))$ — решение системы (10) — (13), и, следовательно, $x(\overline{\mathcal{B}})$ — решение задачи $(\mathcal{D}, f)$.

Обоснованием сходимости предложенного здесь итеративного процесса служит следующая теорема.

Теорема. Пусть $\mathcal{D} \neq \emptyset$, матрица D отрицательно определена и для решения задачи $(\mathcal{Z}, g)$ применяется алгоритм симплекс-метода, в котором учтены правила выбора начального базиса и правило ввода. Если на некоторой итерации $J^-(\mathcal{B}) \neq \emptyset$, но дальнейшие вычисления невозможны без нарушения правила ввода, то $g(w(\overline{\mathcal{B}})) = 0$.

Доказательство этой теоремы будет проведено в следующем пункте. Заметим, что число ограничений в задаче $(\mathcal{Z}, g)$ обычно значительно превышает число ограничений в задаче $(\mathcal{D}, f)$. Это обстоятельство является одной из причин медленной сходимости алгоритма. Дополнительные вычислительные трудности возникают и в связи с вырожденностью задачи $(\mathcal{Z}, g)$.

2.4. Доказательство теоремы п. 2.3. Предположим, что на некоторой итерации алгоритма построен такой допустимый базис $\overline{\mathcal{B}}$, что $J^-(\overline{\mathcal{B}}) \neq \emptyset$, но дальнейшие итерации невозможны без нарушения правила ввода. Докажем, что $g(w(\overline{\mathcal{B}})) = 0$.

Не умаляя общности, можно считать, что $w(\overline{\mathcal{B}}) = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{y}', \bar{y}'', \bar{v}_1, \bar{v}_2, \bar{v}_3, \bar{z})$, где вектор $\bar{x}_1$ составлен из базисных компонент вектора x , а вектор $\bar{v}_2$ — из базисных компонент вектора v . В силу правила выбора начального базиса и правила ввода множество индексов компонент вектора $\bar{x}$, входящих в $\bar{x}_1$, не пересекается с множеством индексов компонент вектора $\bar{v}$, входящих в $\bar{v}_1$. Поэтому $\bar{x}_2 = 0$, $\bar{v}_1 = 0$, $\bar{x}_3 = \bar{v}_3 = 0$ и систему условий задачи $(\mathcal{Z}, g)$ можно схематически записать в виде следующей таблицы:

	x_1	x_2	x_3	y'	y''	v_1	v_2	v_3	z	
s	A_1	A_2	A_3							b
t_1	$2D_{11}$	$2D_{12}$	$2D_{13}$	$-A_1^T$	A_1^T	I_1			$\mathcal{J}_1$	$-c_1$
t_2	$2D_{21}$	$2D_{22}$	$2D_{23}$	$-A_2^T$	A_2^T		I_2		$\mathcal{J}_2$	$-c_2$
t_3	$2D_{31}$	$2D_{32}$	$2D_{33}$	$-A_3^T$	A_3^T			I_3	$\mathcal{J}_3$	$-c_3$
									-1	

$$(14)$$

Все элементы незаполненных клеток табл. (14) равны нулю, $2D_{11}$, $2D_{22}$ и $2D_{33}$ — квадратные матрицы. Слева от таблицы записан вектор $u(\bar{\mathcal{B}}) = (s, t_1, t_2, t_3)$ (см. п. 4.3.1, 4)). Так как столбцы матрицы G , отвечающие компонентам векторов x_1 и v_2 , входят в базис, то

$$sA_1 + 2t_1D_{11} + 2t_2D_{21} + 2t_3D_{31} = 0, \quad (15)$$

$$t_2 = 0. \quad (16)$$

В силу сделанных предположений среди индексов столбцов матрицы G , отвечающих компонентам векторов x_3 , y' , y'' и v_3 , нет элементов множества $J^-(\bar{\mathcal{B}})$. Следовательно,

$$sA_3 + 2t_1D_{13} + 2t_2D_{23} + 2t_3D_{33} \geq 0, \quad (17)$$

$$t_1A_1^T + t_2A_2^T + t_3A_3^T = 0, \quad (18)$$

$$t_3 \geq 0. \quad (19)$$

Рассмотрим следующую каноническую задачу линейного программирования: $(\mathcal{G}', g') : g'(w') = -\sum_{k=1}^n z[k] \rightarrow \max$,

$$\mathcal{G}' = \{w' = (x_1, x_3, y', y'', v_2, v_3, z) : w' \geq 0,$$

$$A_1x_1 + A_3x_3 = b,$$

$$2D_{11}x_1 + 2D_{13}x_3 - A_1^Ty' + A_1^Ty'' + \mathcal{J}_1z = -c_1,$$

$$2D_{21}x_1 + 2D_{23}x_3 - A_2^Ty' + A_2^Ty'' + I_2v_2 + \mathcal{J}_2z = -c_2,$$

$$2D_{31}x_1 + 2D_{33}x_3 - A_3^Ty' + A_3^Ty'' + I_3v_3 + \mathcal{J}_3z = -c_3\}.$$

Так как $\bar{x}_2 = 0$ и $\bar{v}_1 = 0$, то вектор $\bar{w}' = (\bar{x}_1, \bar{x}_3, \bar{y}', \bar{y}'', \bar{v}_2, \bar{v}_3, \bar{z}) \in \mathcal{G}'$ и $\bar{w}' = w'(\bar{\mathcal{B}})$. В множестве $J^-(\bar{\mathcal{B}})$ нет индексов, отвечающих компонентам вектора w' , поэтому $\bar{w}'$ — решение задачи $(\mathcal{G}', g')$ (см. п. 4.3.6, 4°), а $u(\bar{\mathcal{B}}) = (s, t_1, t_2, t_3)$ — решение задачи, двойственной к задаче $(\mathcal{G}', g')$ (см. упражнение 4.8). Следовательно

(см. теорему 2 п. 3.3.5), $g(w(\mathcal{B})) = g'(\bar{w}') = sb - t_1c_1 - t_2c_2 - t_3c_3$ и для завершения доказательства теоремы достаточно проверить, что $sb - t_1c_1 - t_2c_2 - t_3c_3 = 0$.

Умножим обе части равенства (15) на вектор t_1^T , а обе части неравенства (17) — на t_3^T . Тогда

$$sA_1t_1^T + 2t_1D_{11}t_1^T + 2t_2D_{12}t_1^T + 2t_3D_{13}t_1^T = 0, \quad (20)$$

$$sA_3t_3^T + 2t_1D_{13}t_3^T + 2t_2D_{23}t_3^T + 2t_3D_{33}t_3^T \geq 0 \quad (21)$$

см. (19)). Очевидно также (см. (16)), что

$$sA_2t_2^T + 2t_1D_{12}t_2^T + 2t_2D_{22}t_2^T + 2t_3D_{23}t_2^T = 0. \quad (22)$$

Складывая (20), (22) и (21), получаем $sAt^T + 2tDt^T = 0$, где $t = (t_1, t_2, t_3)$. Так как (см. (18)) $tA^T = 0$, то $sAt^T = 0$. Поэтому $tDt^T \geq 0$ и $t = 0$ в силу отрицательной определенности матрицы D .

Осталось доказать, что $sb = 0$. Так как $t = 0$, то из (15) следует, что $sA_1 = 0$, и поскольку $A\bar{x} = b$, $\bar{x}_2 = 0$, $\bar{x}_3 = 0$, то $b = A\bar{x} = A_1\bar{x}_1$ и $sb = sA_1\bar{x}_1 = 0$. ■

Замечание. Если все ограничения задачи квадратичного программирования заданы в форме неравенств, т. е.

$$(\mathcal{D}, f) : f(x) = cx + x^T Dx \rightarrow \max, \quad \mathcal{D} = \{x \in R^n : Ax \leq b, x \geq 0\},$$

то решение задачи $(\mathcal{D}, f)$ сводится к поиску решения системы линейных уравнений и неравенств

$$Ax + I_m z = b, \quad (23)$$

$$2Dx - A^T y + I_n v = -c, \quad (24)$$

$$x \geq 0, \quad y \geq 0, \quad v \geq 0, \quad z \geq 0, \quad (25)$$

удовлетворяющего условиям

$$vx = 0, \quad yz = 0. \quad (26)$$

Если $\mathcal{D} \neq \emptyset$ и матрица D отрицательно определена, то решение системы (23) — (26) можно найти, применив для поиска решения системы (23) — (25) алгоритм симплекс-метода с учетом правила ввода одновременно для условий $vx = 0$ и $yz = 0$. — Упражнение 5.5.

§ 3. Метод Трубина решения задачи целочисленного линейного программирования специального вида

Здесь рассматривается один класс задач целочисленного линейного программирования (см. п. 2.2.4), для которого процедура симплекс-метода со специальными правилами ввода оказывается эффективной. Основанием для включения этого метода в настоящую главу явилось желание еще раз проиллюстрировать возможность такого его применения для решения

нелинейных экстремальных задач. Условия прямого использования процедуры симплекс-метода для решения задач целочисленного линейного программирования обсуждаются в § 1 гл. 8.

3.1. Обозначения и некоторые свойства. Рассмотрим задачу линейного программирования с ограниченными сверху переменными:

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max,$$

$$\mathcal{D} = \{x \in R^n: Ax = 1_m, 0 \leq x[j] \leq 1, j \in 1:n\},$$

где $A = (a[i, j]) - m \times n$ матрица, ранг которой $r(A) = m$, элементы $a[i, j]$ равны 0 или 1, а $1_m^T = (1, 1, \dots, 1) \in R^m$. Задаче $(\mathcal{D}, f)$ отвечает задача целочисленного линейного программирования

$$(\mathcal{D}_c, f): f(x) = cx \rightarrow \max, \mathcal{D}_c = \{x \in R^n: Ax = 1_m, x[j] = 0 \text{ или } x[j] = 1, j \in 1:n\}.$$

Интерес к таким задачам обусловлен тем обстоятельством, что к ним сводятся некоторые экстремальные задачи, связанные с экономико-математическими моделями (например, задача коммивояжера и некоторые специальные целочисленные задачи транспортного типа).

Прежде всего отметим, что множество $\mathcal{D}$ является выпуклым многогранником (если только $\mathcal{D} \neq \emptyset$), поскольку $\mathcal{D} \subset \subset P[0, 1_n]$. Очевидно также, что каждая точка из $\mathcal{D}_c$ является вершиной выпуклого многогранника $\mathcal{D}$, так как она является вершиной куба $P[0, 1_n]$, содержащего $\mathcal{D}$. Обратное, вообще говоря, неверно: выпуклый многогранник $\mathcal{D}$ может иметь нецелочисленные вершины.

Приведем пример. Задача

$$(\mathcal{D}_c, f): f(x) = -x_1 + 7x_2 + 6x_3 + 2x_4 - x_5 \rightarrow \max,$$

$$\mathcal{D}_c = \{x \in R^5: x_1 + x_2 + x_3 = 1,$$

$$x_2 + x_3 + x_4 + x_5 = 1,$$

$$x_1 + x_2 + x_4 = 1,$$

$$x_j = 0 \text{ или } x_j = 1, j \in 1:5\}$$

является задачей с булевыми переменными. С ней связана задача линейного программирования с ограниченными сверху переменными $(\mathcal{D}, f)$. Выпуклый многогранник

$$\mathcal{D} = \{x \in R^5: x_1 + x_2 + x_3 = 1, x_2 + x_3 + x_4 + x_5 = 1,$$

$$x_1 + x_2 + x_4 = 1, 0 \leq x_j \leq 1, j \in 1:5\}$$

имеет нецелочисленные вершины: базису $\mathcal{B}(N(\mathcal{B}) = \{1, 3, 4\})$ отвечает допустимый нецелочисленный базисный план $x(\mathcal{B}) = (1/2, 0, 1/2, 1/2, 0)$ — нецелочисленная вершина выпуклого многогранника $\mathcal{D}$ (см. пример в п. 1.5.5).

Обоснованием алгоритма решения задачи $(\mathcal{D}_n, f)$ служит следующая теорема [42], которую приводим без доказательства.

Теорема. Если x_1 и x_2 — целочисленные вершины выпуклого многогранника $\mathcal{D}$, а $\mathcal{D}_1$ — его грань минимальной размерности (см. п. 1.7.6), содержащая точки x_1 и x_2 , то все вершины грани $\mathcal{D}_1$ целочисленны.

Из теоремы следует, что из любой неоптимальной целочисленной вершины выпуклого многогранника $\mathcal{D}$ всегда можно перейти вдоль некоторого ребра в смежную целочисленную вершину, в которой целевая функция принимает большее значение (см. п. 4.3.6). Это обстоятельство используется в предлагаемом ниже алгоритме.

3.2. Алгоритм. Предполагается, что читатель знаком с методом § 1. В вычислительную процедуру этого метода вносятся следующие изменения.

1. В процессе вычислений осуществляется перебор только целочисленных допустимых $*$ -базисных планов задачи $(\mathcal{D}, f)$. Проблемы, связанные с построением начального целочисленного допустимого $*$ -базисного плана, обсуждаются ниже (см. 2).

Пусть на некоторой итерации построены такой базис $\mathcal{B}$ и множество N^* , что $J^-(\mathcal{B}, N^*) \neq \emptyset$. Тогда разрешается выбирать только такое l из множества $J^-(\mathcal{B}, N^*)$, для которого вычисления, рекомендуемые в этапах 2 и 4 или этапах 3 и 4, приведут к построению нового допустимого целочисленного $*$ -базисного плана $x(\mathcal{B}', N^*)$. Если такого l не существует, то $x(\mathcal{B}, N^*)$ — решение задачи $(\mathcal{D}_n, f)$.

2. Нахождение начального целочисленного $*$ -базисного плана задачи $(\mathcal{D}, f)$ предлагается осуществлять методом минимизации невязок. Для этого строится вспомогательная задача

$$(\mathcal{D}', \psi): \psi(x, z) = - \sum_{k=1}^m z[k] \rightarrow \max, \mathcal{D}' = \{(x, z) \in \mathbb{R}^{n+m}:$$

$$Ax + Iz = 1_m; 0 \leq x[j] \leq 1, j = 1:n; 0 \leq z[k] \leq 1, k = 1:m\}.$$

Задача $(\mathcal{D}', \psi)$ очевидно принадлежит к рассматриваемому здесь классу задач целочисленного линейного программирования. Вектор $(\underbrace{0, \dots, 0}_n, \underbrace{1, \dots, 1}_m)$ — целочисленный допустимый $*$ -базисный план задачи $(\mathcal{D}', \psi)$.

Поэтому для ее решения можно применить алгоритм § 1 с учетом сформулированного выше правила ввода. Пусть $(\bar{x}, \bar{z})$ — целочисленный оптимальный план задачи $(\mathcal{D}', \psi)$, построенный в соответствии с такими рекомендациями. Если $\psi(\bar{x}, \bar{z}) = 0$, то $\bar{x}$ — целочисленный допустимый $*$ -базисный план задачи $(\mathcal{D}, f)$, а если $\psi(\bar{x}, \bar{z}) < 0$, то $\mathcal{D}_n = \emptyset$.

3.3. О вычислительных трудностях. Применение рассмотренного выше алгоритма связано с двумя основными трудностями.

Вырожденными (см. п. 1.1) являются все целочисленные $\ast$ -базисные планы, последовательный перебор которых необходимо осуществлять в процессе вычислений. Как следствие вырожденности практически на каждой итерации возникают возможность неоднозначного выбора вектора, выводимого из базиса, неоднозначность формирования множества N^* и т. п., что приводит как к теоретическим осложнениям, так и к чисто практическим затруднениям, связанным прежде всего с рациональной организацией вычислений на ЭВМ. Вторая проблема также достаточно очевидна: почти на каждой итерации выбор кандидата для ввода в базис, вообще говоря, связан с дополнительными вычислениями и, следовательно, с затратами времени на проверку того, будет ли следующий $\ast$ -базисный план целочисленным. При неблагоприятном исходе этой проверки необходим дополнительный пересчет и т. д., что, понятно, не ускоряет сходимости метода.

УПРАЖНЕНИЯ

- 5.1. См. с. 194. 5.2. См. с. 198. 5.3. См. с. 203. 5.4. См. с. 203.
5.5. См. с. 208.

Глава 6

СПЕЦИАЛЬНЫЕ КЛАССЫ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ И МЕТОДЫ ИХ РЕШЕНИЯ

§ 1. Специальные классы задач линейного программирования. Проблемы генерирования их условий и вычислительной информации. Симуляция допустимых планов

1.1. Предварительные замечания. В настоящей главе рассматриваются некоторые специальные классы задач линейного программирования. Решение любой из них можно очевидным образом свести к нахождению оптимального плана некоторой канонической задачи линейного программирования. Методы ее решения рассматривались в четвертой главе.

Однако в некоторых случаях применение универсальных алгоритмов нерационально или сведение к канонической задаче линейного программирования приводит к задаче, у которой матрица ограничений имеет слишком большую размерность (в первую очередь это касается числа строк), что не позволяет осуществлять ее решение даже на современных ЭВМ. В то же время специфика блочной (см. § 2) и транспортных (см. § 3—5) задач и их сводимость к канонической задаче линейного программирования позволяют в результате модификации методов четвертой главы существенно упростить вычислительные процедуры либо, что особенно важно, преодолеть в той или иной степени трудности, связанные с большой размерностью решаемой задачи, т. е. разработать вычислительные методы ее решения, допускающие их реализацию на ЭВМ.

Специфическая структура позволяет не только строить удобные вычислительные процедуры решения задач, но и формулировать в довольно неожиданных терминах некоторые их свойства. Например, для транспортной задачи в п. 4.3 приведены изящные условия ее вырожденности (невыврожденности). Итеративный алгоритм решения блочной задачи (см. § 2) послужил основой любопытной экономической интерпретации как самой задачи, так и метода ее решения.

Выбирая материал для настоящей главы, авторы ставили целью не только ознакомить читателя с некоторыми интересными классами задач линейного программирования, но и при-

влечь его внимание к важности и целесообразности учета специфики задач при решении их известными методами.

1.2. Проблемы генерирования условий задач и вычислительной информации. Вычислительные методы данной главы, как и предыдущей, иллюстрируются условными примерами, решение которых легко осуществляется. Поэтому нам хотелось бы предостеречь читателя и обратить его внимание на некоторые вычислительные аспекты алгоритмов, которых мы ранее не касались.

Задачи, возникающие на практике, как правило, имеют большую размерность. Их решение даже на современных ЭВМ связано с определенными трудностями. К ним следует отнести: необходимость сокращения объема вводимой в ЭВМ информации, так как скорость ее ввода не соответствует быстродействию машины;

приемлемое размещение в памяти ЭВМ массивов исходной информации;

эффективное использование текущей информации;

определение оптимальных размеров рабочего поля в оперативной памяти ЭВМ и т. д.

Все это делает целесообразным формирование алгоритмическим путем — *генерирование* условий задач и информации, необходимой в процессе вычислений.

Прежде всего остановимся на проблеме генерирования условий задач. Успешное ее решение может существенно сократить время ввода исходной информации. Вместе с тем при решении некоторых задач большой размерности эффективные алгоритмы генерирования открывают единственную возможность реализации на ЭВМ вычислительных методов. Приведем несколько примеров.

1°. Анализ некоторых проблем сводится к решению задач вида

$$(\mathcal{D}, f): f(x) \rightarrow \max, \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\}, \quad (1)$$

где матрицу ограничений A можно представить в виде произведения двух матриц: $A = GP$, где $G = (g[i, k]) - m \times r$ и $P = (p[k, j]) - r \times n$ матрицы; g_i — i -я строка матрицы G , а p^j — j -й столбец матрицы P . Для ввода в ЭВМ элементов матрицы $A = (a[i, j])$ потребовалось бы (если невозможна упаковка) mn ячеек памяти. При больших m и n объем вводимой информации может превосходить даже возможности внешней памяти машины. Если же r мало по сравнению с m и n , то вместо ввода в память элементов матрицы A можно ограничиться вводом элементов матриц G и P (в этом случае число $mr + rn$ может оказаться значительно меньше mn), а элементы матрицы A генерировать по правилу $a[i, j] = g_i p^j$.

2°. В п. 3.1.11, 3° транспортная задача записана в виде (1). Понятно, что не следует вводить в ЭВМ все элементы мат-

рицы A , так как для каждой пары индексов i, j элементы столбца a^j генерируются по правилу: i -я и $(m + j)$ -я компоненты равны единице, остальные — нулю.

3°. Предлагаем самостоятельно сформулировать алгоритмы генерирования столбцов матрицы A для станковой задачи (см. п. 2.3.5), задачи о назначениях (см. п. 2.3.8) и задачи коммивояжера (см. п. 2.3.9). — Упражнение 6.1. Обращаем внимание на то, что при решении последних двух задач в память ЭВМ достаточно вводить только элементы матрицы C , т. е. коэффициенты целевой функции задачи.

4°. Довольно часто в матрицах ограничений решаемых на ЭВМ задач содержится много нулей (например, такая задача рассматривается в следующем параграфе). Как правило, сами нули в память не вводятся. Фиксируются и запоминаются их число и расположение. Столбцы, содержащие нули, генерируются по соответствующим программам.

5°. Идея генерирования широко используется и для формирования информации, необходимой в процессе вычислений. В общем плане эта идея проводится во всех алгоритмах гл. 4. Например, в симплекс-методе (см. § 3 гл. 4) вся информация, необходимая для дальнейших вычислений, генерируется по базису $\mathcal{B}$ (формулы п. 4.3.1), и если он оказывается оптимальным, то аналогично генерируются решение $x(\mathcal{B})$ и максимальное значение целевой функции $f(x(\mathcal{B}))$.

6°. В машинных алгоритмах генерирование позволяет осуществлять экономию памяти либо времени счета. Доминирует принцип: в каждый момент вычислений на рабочем поле должна находиться только нужная информация. Например, при проверке оптимальности плана $x(\mathcal{B})$ (п. 1 этапа 1 алгоритма п. 4.3.3) необходимо знать лишь элементы строки $a_0(\mathcal{B})$. Последние можно генерировать по правилу $a_0(\mathcal{B}) = d_0(\mathcal{B})A$. В некоторых случаях рационально не хранить в памяти вектор $d_0(\mathcal{B})$, а генерировать его по базису $\mathcal{B}$.

В памяти не хранятся столбцы матрицы $\bar{A}(\mathcal{B})$ с номерами $j \in N(\mathcal{B})$. Их элементы генерируются очевидным образом.

7°. В методах § 4, 5 гл. 4 аналогичные идеи находят применение не только в машинных алгоритмах. Действительно, в отличие от алгоритма § 3 гл. 4, где на каждой итерации вычисляются все элементы матрицы $\bar{A}(\mathcal{B})$, в алгоритмах § 4, 5 той же главы вычисляется лишь вектор $a_0(\mathcal{B})$ — верхняя строка матрицы $A(\mathcal{B})$, — необходимый для проверки оптимальности плана $x(\mathcal{B})$ (см. 5°). Если план $x(\mathcal{B})$ неоптимален, то по строке $a_0(\mathcal{B})$ можно генерировать (по правилам, указанным в соответствующих алгоритмах) номер l столбца матрицы A , который вводится в базис (или является кандидатом). Столбец $\bar{a}^l$ генерируется из внешней памяти, а столбцы $\bar{a}^i(\mathcal{B})$, $\bar{b}(\mathcal{B})$ — по базису $\mathcal{B}$: $\bar{a}^i(\mathcal{B}) = B^{-1}\bar{a}^i$, $\bar{b}(\mathcal{B}) = B^{-1}\bar{b}$, т. е. знание обратной матрицы позволяет однозначно генерировать всю информацию, не-

обходимую для дальнейших вычислений. Матрицу B^{-1} можно хранить в оперативной памяти или даже генерировать в нужный момент, используя стандартные программы обращения матриц. В таком случае элементы матрицы B генерируются из внешних запоминающих устройств ЭВМ, где хранится вся неоперативная информация: упакованный (или неупакованный) массив элементов матрицы A с информацией о нулевых элементах, компоненты векторов c и b и т. п.

Машинная и ручная реализация методов настоящей главы невозможна без эффективных алгоритмов генерирования вычислительной информации. На эту особенность следует обратить особое внимание, поскольку, как и прежде, вычислительные аспекты детально не обсуждаются.

1.3. Симуляция допустимых планов. Возвращаясь к п. 2. 3.1, отметим, что, вообще говоря, можно построить математическую модель любого функционирующего объекта, процесса или явления. Трудности (иногда непреодолимые) появляются при анализе моделей и экстремальных задач, возникающих на их основе. До сих пор рассматривались достаточно исследованные задачи, предлагались точные методы их решения. В общем случае нахождение оптимального плана задачи $(\mathcal{D}, f)$ является чрезвычайно сложной проблемой. Существуют классы экстремальных задач, для которых еще не удалось предложить точные или хорошие приближенные методы их решения. В этом случае пытаются осуществить разумное упрощение соответствующей модели. Если же оно невозможно, прибегают к некоторой процедуре построения элементов множества $\mathcal{D}$. Для этой цели и применяются симуляторы.

Симулятором называют алгоритм (или совокупность алгоритмов), позволяющий в случае непустоты множества $\mathcal{D}$ получать план $x \in \mathcal{D}$ и вычислять значение $f(x)$. Процесс однократного применения симулятора называется *симуляцией*.

Примерами симуляторов могут служить рассмотренные в гл. 4 методы построения допустимых планов. В этом случае симуляторы являются сложными: они сводятся к применению универсальных методов для решения вспомогательных задач или связаны с громоздкими вычислительными процедурами (см. § 6). В то же время простота — наиболее важное требование к симулятору.

По этой причине бесполезно обсуждать принципы формирования симулятора для произвольной экстремальной задачи, так как возможность эффективного вычисления точки множества $\mathcal{D}$ зависит от структуры этого множества, свойств допустимых планов и т. п. Даже для ограниченных классов экстремальных задач построение симуляторов может оказаться достаточно сложным.

Однако для некоторых классов задач специального вида и некоторых дискретных задач можно предложить простые про-

цедуры симуляции. Приведем несколько примеров. Рассмотрим транспортную задачу в матричной постановке (см. п. 2.3.2):

$$(\mathcal{D}, f): f(x) = \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j] \rightarrow \min,$$

$$\mathcal{D} = \{x \in R^{mn} : \sum_{j=1}^n x[i, j] = a[i], i \in 1:m; \sum_{i=1}^m x[i, j] = b[j], j \in 1:n; x \geq 0\}. \quad (1)$$

Элементы множества $\mathcal{D}$, если выполняется условие баланса $\sum_{i=1}^m a[i] = \sum_{j=1}^n b[j]$, можно получать с помощью приводимых ниже симуляторов.

1°. Симуляция без учета стоимости перевозки.* Симулятор позволяет определять положительные компоненты плана $x \in \mathcal{D}$. Для их вычисления используются только массивы $a = (a[1], a[2], \dots, a[m])$ и $b = (b[1], b[2], \dots, b[n])$ ($a > 0, b > 0$).

Алгоритм. 1. Полагаем $I = 1:m, J = 1:n, l = 1, r = 1, x[i, j] = 0$ для $i \in I, j \in J$ и переходим к рекомендации 2.

2. Если $I = \emptyset, J = \emptyset$, то вычисляем

$$f(x) = \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j]$$

и заканчиваем вычислительный процесс. В противном случае переходим к рекомендации 3.

3. Вычисляем $u = \min\{a[l], b[r]\}$ и полагаем $x[l, r] = u$. Возможны три случая.

3'. $a[l] - u = 0, b[r] - u > 0$. Переходим к рекомендации 4.

3''. $a[l] - u > 0, b[r] - u = 0$. Переходим к рекомендации 5.

3'''. $a[l] - u = 0, b[r] - u = 0$. Переходим к рекомендации 6.

4. Полагаем $I = I \setminus \{l\}, l = l + 1, b[r] = b[r] - u$. Переходим к рекомендации 3.

5. Полагаем $J = J \setminus \{r\}, r = r + 1, a[l] = a[l] - u$. Переходим к рекомендации 3.

6. Полагаем $I = I \setminus \{l\}, J = J \setminus \{r\}, l = l + 1, r = r + 1$. Переходим к рекомендации 2. ■

2°. Симуляция с учетом стоимости перевозки.** Данный симулятор, как и предыдущий, позволяет вычислять положительные компоненты плана $x \in \mathcal{D}$. Для расчетов используются массивы $a = (a[1], a[2], \dots, a[m])$, $b = (b[1], b[2], \dots, b[n])$ ($a > 0, b > 0$) и массив элементов матрицы $C = (c[i, j], i \in 1:m, j \in 1:n)$.

Алгоритм. 1. Полагаем $I = 1:m, J = 1:n, x[i, j] = 0$ для $i \in I, j \in J$ и переходим к рекомендации 2.

* Этот метод называется *методом северо-западного угла*.

** Обычно этот метод называют *методом минимального элемента*.

2. Если $I = \emptyset$, $J = \emptyset$, то вычисляем

$$f(x) = \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j]$$

и заканчиваем вычислительный процесс. В противном случае переходим к рекомендации 3.

3. Вычисляем $d = \min_{i \in I, j \in J} \{c[i, j]\}$; полагаем l равным наименьшему из индексов s , для которых $c[s, j] = d$, а r — наименьшему из индексов s , для которых $c[l, s] = d$. Дальнейшие вычисления осуществляются так же, как в предыдущем алгоритме, начиная с рекомендации 3, с последующими переходами к рекомендациям 2, 3 настоящего алгоритма. ■

Конечность алгоритмов 1°, 2° очевидна. Допустимость вычисленного вектора x следует из построения.

Пусть в транспортной задаче $(\mathcal{D}, f)$ план x' получен с помощью первого симулятора, а x'' — с помощью второго. Можно было бы предполагать, что всегда $f(x'') \leq f(x')$. Однако нетрудно привести пример такой транспортной задачи, в которой $f(x'') > f(x')$. Объяснение причин и построение соответствующего примера читателю предлагается провести самостоятельно. — Упражнение 6.2.

В транспортной задаче $(\mathcal{D}, f)$ с нелинейной функцией f (разрывной, невыпуклой и т. п.), генерируя допустимые планы и значения функции f с помощью симуляторов, можно (за приемлемое время) перебрать некоторое количество планов и среди них выбрать тот, на котором достигается наименьшее значение функции f . Для этой цели симуляторы должны генерировать различные планы. Последнее можно осуществить, например, последовательной перестановкой чисел $a[i]$, $i \in 1:m$, и $b[j]$, $j \in 1:n$, или случайным генерированием индексов l и r . Соответствующие процедуры нетрудно построить самостоятельно и включить в симуляторы.

Предлагаем самостоятельно построить симуляторы для задачи о назначениях (см. п. 2.3.8) и задачи коммивояжера (см. п. 2.3.9). — Упражнение 6.3.

§ 2. Блочная задача.

Метод декомпозиции (разложения) Данцига—Вулфа

2.1. Определения и комментарии. Как уже отмечалось, трудности, связанные с решением задач линейного программирования большой размерности, могут быть преодолены, если матрица ограничений задачи имеет специальную структуру: большое число нулевых элементов и «хорошее» их взаимное расположение. Одним из наиболее характерных примеров матриц с такой специальной структурой являются блочные матрицы.

Матрица A порядка $m \times n$ называется *блочной*, если некоторыми перестановками строк и столбцов она может быть приведена к виду

$$\begin{array}{|c|} \hline A_0 \\ \hline \begin{array}{|c|c|} \hline A_1 & \\ \hline & A_2 \\ \hline & & \ddots \\ \hline & & & A_r \\ \hline \end{array} \\ \hline \end{array} \quad (1)$$

Здесь A_0 — матрица порядка $m_0 \times n$; A_s , $s \in 1:r$, — матрица порядка $m_s \times n_s$ ($\sum_{s=0}^r m_s = m$, $\sum_{s=1}^r n_s \leq n$). Все остальные элементы табл. (1) равны нулю. Задача линейного программирования

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\}$$

с блочной матрицей A называется *блочной*, или *разложимой, задачей*.

Мы не останавливаемся здесь на проблемах, связанных с проверкой блочной структуры матрицы и приведением блочных матриц к виду (1). Обычно такие матрицы возникают в задачах оптимального распределения ресурсов в системах, имеющих иерархическую структуру: матрицы A_s , $s \in 1:r$, «ответственны» за распределение ресурсов подразделений, а блок A_0 — за распределение ресурсов, находящихся в распоряжении управляющего центра системы. В таких задачах естественная нумерация ресурсов и интенсивностей их потребления (производства) приводит к матрицам вида (1). В дальнейшем будем называть блочными только такие матрицы.

Агрегирование ограничений и переменных блочной задачи позволяет переформулировать ее следующим образом:

$$\begin{aligned} (\mathcal{D}, f): f(x) = \sum_{s=0}^r c_s x^s \rightarrow \max, \mathcal{D} = \{x = (x^0, x^1, \dots, x^r): \\ x^s \in R^{n_s}, s \in 0:r; A_s x^s = b^s, s \in 1:r; \\ \sum_{s=0}^r A_0 x^s = b^0; x^s \geq 0, s \in 0:r\}, \end{aligned} \quad (2)$$

где $A_0 s$ и c_s , $s \in 0:r$, — блоки матрицы A_0 и вектора c , отвечающие переменным, входящим в вектор x^s (столбцам матрицы A_s), а b^s , $s \in 0:r$, — блоки вектора b , отвечающие строкам, входящим в A_s . Задача (2) может быть представлена схемой (2').

x^0	x^1	x^2	$\dots$	x^r	
A_{00}	A_{01}	A_{02}	$\dots$	A_{0r}	b^0
	A_1				b^1
		A_2			b^2
			$\ddots$		$\vdots$
				A_r	b^r
c_0	c_1	c_2	$\dots$	c_r	

(2')

2.2. Угловая задача. Если $\mathcal{D} \neq \emptyset$, то при каждом $s \in 1:r$ множество $\mathcal{D}^s = \{x^s \in R^{n_s} : A_s x^s = b^s, x^s \geq 0\}$ является многогранным выпуклым множеством в R^{n_s} , не содержащим прямых, поскольку оно содержится в R^{n_s+} . Поэтому в силу теоремы 1 п. 1.7.5 множество $\mathcal{D}^s$ представимо в виде алгебраической суммы выпуклого многогранника V^s , являющегося выпуклой оболочкой конечного множества экстремальных точек — вершин множества $\mathcal{D}^s$, и многогранного выпуклого конуса K_0^s с вершиной в нуле. При этом если множество $\mathcal{D}^s$ ограничено, то $\mathcal{D}^s = V^s$.

Множество $\mathcal{D}^s$ называется *блоком задачи* $(\mathcal{D}, f)$.

Обозначим экстремальные точки множества $\mathcal{D}^s$ через $u_1^s, u_2^s, \dots, u_{k_s}^s$. Образующие векторы конуса K_0^s , если $\mathcal{D}^s$ не является выпуклым многогранником, обозначим через $z_1^s, z_2^s, \dots, z_{h_s}^s$. Положим

$$a_t^s = A_{0s} u_t^s, \quad t \in 1:k_s, \quad s \in 1:r, \quad (3)$$

$$\tilde{a}_t^s = A_{0s} z_t^s, \quad t \in 1:h_s, \quad s \in 1:r, \quad (4)$$

$$c[s, t] = c_s u_t^s, \quad t \in 1:k_s, \quad s \in 1:r, \quad (5)$$

$$\tilde{c}[s, t] = c_s z_t^s, \quad t \in 1:h_s, \quad s \in 1:r \quad (6)$$

(естественно, если $h_s = 0$, то векторы $\tilde{a}_t^s$ и числа $\tilde{c}[s, t]$ не определяются). Рассмотрим экстремальную задачу $(\mathcal{G}, g)$:

$$g(x^0, \alpha, \tilde{\alpha}) = c_0 x^0 + \sum_{s=1}^r \sum_{t=1}^{k_s} c[s, t] \alpha[s, t] + \sum_{s=1}^r \sum_{t=1}^{h_s} \tilde{c}[s, t] \tilde{\alpha}[s, t] \rightarrow \max,$$

$$\mathcal{G} = \{(x^0, \alpha, \tilde{\alpha}) : x^0 \in R^{n_0}, \alpha \in R^{n'}, \tilde{\alpha} \in R^{\tilde{n}'},$$

$$A_{00} x^0 + \sum_{s=1}^r \sum_{t=1}^{k_s} a_t^s \alpha[s, t] + \sum_{s=1}^r \sum_{t=1}^{h_s} \tilde{a}_t^s \tilde{\alpha}[s, t] = b^0, \quad (7)$$

$$\sum_{t=1}^{k_s} \alpha[s, t] = 1, \quad s \in 1:r, \quad (8)$$

$$\alpha \geq 0, \quad \tilde{\alpha} \geq 0, \quad x^0 \geq 0, \quad (9)$$

где $n' = \sum_{s=1}^r k_s$, $\tilde{n}' = \sum_{s=1}^r h_s$. Она называется *угловой задачей*, порожденной блочной задачей $(\mathcal{D}, f)$.

Справедливы следующие утверждения.

1°. Задачи $(\mathcal{D}, f)$ и $(\mathcal{G}, g)$ одновременно имеют или не имеют решения.

2°. Если $(\bar{x}^0, \bar{\alpha}, \bar{a})$ — решение угловой задачи, то вектор $\bar{x} = (\bar{x}^0, \bar{x}^1, \dots, \bar{x}^r)$, где

$$\bar{x}^s = \sum_{t=1}^{k_s} \bar{\alpha}[s, t] u_t^s + \sum_{t=1}^{h_s} \bar{a}[s, t] z_t^s$$

— решение задачи $(\mathcal{D}, f)$.

Доказательство. Пусть $x = (x^0, x^1, \dots, x^r) \in \mathcal{D}$. Тогда для каждого $s \in 1:r$ вектор x^s принадлежит многогранному выпуклому множеству $\mathcal{D}^s$ и в силу теоремы 1 п. 1.7.5 существуют такие числа $\alpha[s, t]$, $t \in 1:k_s$, и $\bar{a}[s, t]$, $t \in 1:h_s$, что

$$x^s = \sum_{t=1}^{k_s} \alpha[s, t] u_t^s + \sum_{t=1}^{h_s} \bar{a}[s, t] z_t^s, \quad (10)$$

$$\alpha[s, t] \geq 0, t \in 1:k_s, \sum_{t=1}^{k_s} \alpha[s, t] = 1, \bar{a}[s, t] \geq 0, t \in 1:h_s. \quad (11)$$

Определим векторы $\alpha \in R^{n'}$ ($n' = \sum_{s=1}^r k_s$) и $\bar{a} \in R^{\tilde{n}'}$ ($\tilde{n}' = \sum_{s=1}^r h_s$) с компонентами $\alpha[s, t]$ и $\bar{a}[s, t]$ соответственно. Очевидно, что для вектора $\varphi(x) = (x^0, \alpha, \bar{a})$ выполняются соотношения (8) и (9), а также в силу (10), (3) и (4) — равенство (7). Следовательно, $\varphi(x) \in \mathcal{G}$. Кроме того, в силу (10), (5) и (6) выполняется равенство $f(x) = g(\varphi(x))$.

Обратно, если задан вектор $(x^0, \alpha, \bar{a}) \in \mathcal{G}$, то, вычислив при каждом $s \in 1:r$ вектор x^s по формуле (10), определим вектор $\psi(x^0, \alpha, \bar{a}) = (x^0, x^1, \dots, x^r)$, который в силу тех же соотношений (3) — (6) принадлежит $\mathcal{D}$, и $g(x^0, \alpha, \bar{a}) = f(\psi(x^0, \alpha, \bar{a}))$.

Итак, для экстремальных задач $(\mathcal{D}, f)$ и $(\mathcal{G}, g)$ существуют эквивалентные преобразования (см. п. 2.4.2), и утверждения 1° и 2° доказываются аналогично свойствам а) — в) эквивалентных экстремальных задач (см. п. 2.4.3). Заметим, что прямая ссылка на свойства эквивалентных экстремальных задач не оправдана, поскольку отображения φ и ψ не определены алгоритически. ■

Особо отметим, что если известны вектор x^0 , положительные компоненты векторов α и $\bar{a}$, а также отвечающие этим положительным компонентам вершины u_t^s и образующие z_t^s , то формулы (10) при $s \in 1:r$ позволяют вычислить вектор $x = \psi(x^0, \alpha, \bar{a})$. При этом если $(x^0, \alpha, \bar{a}) \in \mathcal{G}$, то $\psi(x^0, \alpha, \bar{a}) \in \mathcal{D}$, и если $(x^0, \alpha, \bar{a})$ — решение задачи $(\mathcal{G}, g)$, то $\psi(x^0, \alpha, \bar{a})$ — решение задачи $(\mathcal{D}, f)$.

2.3. Матрица угловой задачи. В дальнейшем предполагается, что $\sum_{s=1}^r n_s = n$.

Обозначим $m' \times (n' + \tilde{n}')$ матрицу ограничений задачи $(\mathcal{G}, g)$ через A' . Столбцы матрицы A' формируются следующим образом:

1) вершине $u_i^s \in \mathcal{D}^s$ соответствует столбец a^{st} , у которого первые m_0 компонент вычисляются по формуле (3), $(m_0 + s)$ -я компонента равна единице, а остальные — нулю;

2) образуемому вектору z_i^s отвечает столбец $\bar{a}^{st}$, первые m_0 компонент которого вычисляются по формуле (4), а остальные равны нулю.

Матрица $\bar{A}'$ получается расширением матрицы A' : к последней присоединяется сверху строка a_0 , составленная из коэффициентов целевой функции g , вычисленных по формулам (5) и (6) соответственно. Через $\bar{a}^{st}$ и $\bar{a}^{st}$ обозначим столбцы матрицы $\bar{A}'$.

Поскольку задача $(\mathcal{G}, g)$ является обычной канонической задачей линейного программирования, то в дальнейшем будут использоваться понятия и обозначения п. 4.3.1. В частности, как и раньше, систему $\mathcal{B}$ из m' линейно независимых столбцов матрицы A' назовем базисом задачи $(\mathcal{G}, g)$.

По произвольному базису $\mathcal{B}$ можно построить матрицы B , $\bar{B}$ и $\bar{B}^{-1}$. Обозначим через b' вектор ограничений задачи $(\mathcal{G}, g)$. По построению $b'^T = (b^0, 1, \dots, 1) \in R^{m_0+r}$. Базису $\mathcal{B}$ отвечает базисный план $\alpha(\mathcal{B})$ задачи $(\mathcal{G}, g)$, базисные компоненты которого равны соответствующим компонентам вектора $b(\mathcal{B}) = B^{-1}b'$, а остальные компоненты — нули. Верхняя строка матрицы $\bar{B}^{-1}$ обозначается через $d_0(\mathcal{B}) = (-1, u(\mathcal{B}))$, где $u(\mathcal{B}) = c(\mathcal{B})B^{-1}$. В дальнейшем верхнюю строку матрицы $\bar{B}^{-1}$ будем записывать в виде

$$d_0(\mathcal{B}) = (-1, v(\mathcal{B}), w(\mathcal{B})), \quad (12)$$

где вектор $v(\mathcal{B})$ составлен из m_0 первых компонент вектора $u(\mathcal{B})$, а $w(\mathcal{B})$ — из r последних компонент вектора $u(\mathcal{B})$; $w(\mathcal{B})[s]$ — s -я компонента вектора $w(\mathcal{B})$ ($(m_0 + s)$ -я компонента вектора $u(\mathcal{B})$).

Пусть $\mathcal{B}$ — базис задачи $(\mathcal{G}, g)$. Для каждого $s \in 1:r$ введем в рассмотрение каноническую задачу линейного программирования

$$(\mathcal{D}^s, y^s, \mathcal{B}) : y^s(x^s) = [-c_s + v(\mathcal{B}) A_{0s}] x^s \rightarrow \min,$$

$$\mathcal{D}^s = \{x^s \in R^{n_s} : A_s x^s = b^s, x^s \geq 0\}. \quad (13)$$

Целесообразность рассмотрения задач (13) будет обоснована ниже.

2.4. Предварительные замечания. Метод декомпозиции используется как одна из крайних мер в ситуации, когда обычные вычислительные алгоритмы не работают и необходимо уменьшить размерность задачи, точнее, уменьшить число ее ог-

раничений, поскольку именно эта характеристика является основным препятствием при машинной реализации. С этой целью в п. 2.2 введена в рассмотрение угловая задача. Действительно, переход от блочной задачи, у которой $m_0 + \sum_{s=1}^r m_s$ нетривиальных ограничений, к соответствующей угловой задаче с $m_0 + r$ ограничениями в общем случае значительно уменьшает число ограничений.

Тем не менее, может показаться, что ситуация остается бесперспективной, так как для построения задачи $(\mathcal{G}, g)$ необходимо при каждом s знать все (1) вершины выпуклого многогранника V^s и образующие векторы конуса K_0^s . Вообще говоря, нетрудно, например на основе симплекс-метода, предложить алгоритм формирования множества всех вершин и образующих векторов. Однако выполнить весь объем подобных вычислений для реальных задач практически невозможно.

Выход из этого положения, предложенный Дж. Данцигом и Ф. Вулфом, весьма изящен и состоит в следующем.

Если известен допустимый базис $\mathcal{B}$ угловой задачи, то по нему можно сформировать допустимый базисный план $\alpha(\mathcal{B})$ и вектор $d_0(\mathcal{B})$. Для оптимальности плана $\alpha(\mathcal{B})$ достаточно, а в случае невырожденности задачи $(\mathcal{G}, g)$ и необходимо (см. п. 4.3.6, 4°) выполнение неравенства $a_0(\mathcal{B}) = d_0(\mathcal{B})\bar{A}' \geq 0$. Оказывается, что проверка этого неравенства сводится к решению r задач линейного программирования (13) при $s \in 1:r$. Если каждая из задач $(\mathcal{D}^s, y^s, \mathcal{B})$ имеет решение (пусть x^s — решение задачи $(\mathcal{D}^s, y^s, \mathcal{B})$) и $y^s(x^s) + w(\mathcal{B})[s] \geq 0$, $s \in 1:r$, то $\alpha(\mathcal{B})$ — решение задачи $(\mathcal{G}, g)$. В противном случае определяется вершина u_i^s или образующий вектор z_i^s , а следовательно, и столбец $\bar{a}^{st}$ или $\bar{a}^{st}$ матрицы $\bar{A}'$, введение которого в базис приводит к возрастанию значения целевой функции g .

Итак, решение блочной задачи большой размерности сводится к последовательности итераций, на каждой из которых решается ряд вспомогательных задач относительно небольшой размерности и выполняется одна итерация алгоритма решения угловой задачи. Следует отметить, что подобный подход имеет существенный недостаток — крайне медленную сходимость. Мы не приводим рекомендаций, позволяющих в той или иной степени сократить объем вычислений, поскольку считаем своей целью изложение идеи метода, а не тонкостей и трудностей его практической реализации. Хотелось бы также обратить внимание читателя на то, что идея метода декомпозиции допускает любопытные экономические интерпретации и является основой многочисленных исследований иерархических систем в экономике.

2.5. Описание и обоснование вычислительной процедуры. В основу вычислительной процедуры положен алгоритм симплекс-метода (см. § 4 гл. 4). Однако для аналогичной цели можно

было бы использовать, например, и процедуру двойственного симплекс-метода (см. § 5 гл. 4).

Если угловая задача $(\mathcal{Z}, g)$ и все множества $\mathcal{D}^s$, $s \in 1:r$, невырождены, то метод декомпозиции Данцига — Вульфа позволяет за конечное число итераций найти оптимальный план блочной задачи $(\mathcal{D}, f)$ или обнаружить неограниченность сверху f на $\mathcal{D}$. Число итераций неизвестно до окончания вычислений. Оно зависит от выбора начального допустимого базиса $\mathcal{B}_1$ задачи $(\mathcal{Z}, g)$. Проблемы, связанные с построением такого базиса, обсуждаются ниже.

Опишем одну итерацию метода. Предположим, что задана система из $m_0 + r$ вершин выпуклых многогранников V^s и образующих многогранных выпуклых конусов K_0^s , $s \in 1:r$, которой отвечает допустимый базис $\mathcal{B}$ задачи $(\mathcal{Z}, g)$ (см. п. 2.3). Предположим также, что в табл. $T^2(\mathcal{B})$ (см. п. 4.4.3) заполнены матрица $\bar{B}^{-1}$ и столбцы $b(\mathcal{B})$ и $N(\mathcal{B})$.

О заполнении столбца $N(\mathcal{B})$ следует сказать особо. Для того чтобы по допустимому базисному плану $\alpha(\mathcal{B})$ задачи $(\mathcal{Z}, g)$ (или по столбцу $b(\mathcal{B})$ табл. $T^2(\mathcal{B})$) можно было по формуле (10) построить план задачи $(\mathcal{D}, f)$, необходимо помнить вершины многогранников V^s и образующие векторы конусов K_0^s , которым отвечают векторы базиса $\mathcal{B}$. Кроме того, необходимо знать, какому блоку $\mathcal{D}^s$ принадлежат эти вершины многогранников и образующие векторы конусов; соответствующая информация должна содержаться в столбце $N(\mathcal{B})$.

Для проверки оптимальности плана $\alpha(\mathcal{B})$ последовательно решают симплекс-методом задачи (13) для $s = 1, 2, \dots, r$. Так как алгоритм п. 4.4.3 сформулирован для задачи максимизации, то вместо задач минимизации $(\mathcal{D}^s, y^s, \mathcal{B})$ удобнее рассматривать задачи максимизации $(\mathcal{D}^s, -y^s, \mathcal{B})$.

Предположим сначала, что каждая из задач $(\mathcal{D}^s, y^s, \mathcal{B})$, $s \in 1:r$, имеет решение. Тогда при каждом $s \in 1:r$ применение алгоритма симплекс-метода к задаче $(\mathcal{D}^s, -y^s, \mathcal{B})$ приведет к определению вершины $\hat{u}_i^s \in \mathcal{D}^s$, являющейся решением этой задачи. Возможны следующие случаи.

1. Для каждого $s \in 1:r$ выполняется неравенство $y^s(\hat{u}_i^s) + w(\mathcal{B})[s] \geq 0$. Так как при каждом $s \in 1:r$ для любой вершины $u_i^s \in \mathcal{D}^s$

$$\begin{aligned} d_0(\mathcal{B}) \bar{a}^{st} &= -c[s, t] + v(\mathcal{B}) a_i^s + w(\mathcal{B})[s] = \\ &= [-c_s + v(\mathcal{B}) A_{0s}] u_i^s + w(\mathcal{B})[s] = y^s(u_i^s) + w(\mathcal{B})[s] \end{aligned} \quad (14)$$

(см. (12) и формулы (3) и (5)), то

$$d_0(\mathcal{B}) \bar{a}^{st} \geq y^s(\hat{u}_i^s) + w(\mathcal{B})[s] \geq 0.$$

Кроме того, при каждом $s \in 1:r$ для любого образующего вектора z_i^s конуса K_0^s

$$d_0(\mathcal{B}) \bar{a}^{st} = -\bar{c}[s, t] + v(\mathcal{B}) \bar{a}_i^s = [-c_s + v(\mathcal{B}) A_{0s}] z_i^s = y^s(z_i^s) \quad (15)$$

(см. (12) и формулы (4) и (6)) и в силу свойства 6° п. 1.8.2 $d_0(\mathcal{B}) \bar{a}^{st} \geq 0$. Следовательно, $a_0(\mathcal{B}) = d_0(\mathcal{B}) \bar{A}' \geq 0$, и план $\alpha(\mathcal{B})$ — решение задачи $(\mathcal{G}, g)$. Остается сформировать решение задачи $(\mathcal{D}, f)$ по формуле (10).

2. Найдется такое s , для которого $y^s(\hat{a}_i^s) + w(\mathcal{B})[s] < 0$. Обозначим через $\hat{a}^t = \hat{a}^{st}$ столбец матрицы $\bar{A}'$, отвечающий вершине $\hat{a}_i^s$. Тогда в силу соотношения (14) $d_0(\mathcal{B}) \hat{a}^t < 0$, и введение столбца $\hat{a}^t$ в базис приведет к увеличению значения целевой функции задачи $(\mathcal{G}, g)$. При этом поскольку $\hat{a}^t[m_0 + s] = \hat{a}^{st}[m_0 + s] = 1$, то, как нетрудно проверить, *по крайней мере одна компонента вектора $\hat{a}^t(\mathcal{B}) = B^{-1} \hat{a}^t$ положительна.* — Упражнение 6.4. Так что применение рекомендаций алгоритма п. 4.4.3 приведет к построению нового базиса $\mathcal{B}'$. Отметим, что правила формирования (генерирования) столбца $\bar{a}^t$ сформулированы в п. 2.3, а правила формирования нового базиса $\mathcal{B}'$ и табл. $T^2(\mathcal{B}')$ — в п. 4.4.3.

Предположим теперь, что при некотором s в процессе решения задачи $(\mathcal{D}^s, -y^s, \mathcal{B})$ методом п. 4.4.3 (или п. 4.3.3) обнаружена неограниченность сверху целевой функции $-y^s$ на $\mathcal{D}^s$. Тогда формула (13) п. 4.3.6 определяет направление s^t неограниченности многогранного выпуклого множества $\mathcal{D}^s$, и $y^s(s^t) < 0$ (см. 2° и 6° п. 4.3.6). Как уже отмечалось в п. 4.3.6 (см. 2°), в случае невырожденной задачи вектор s^t направлен вдоль ребра множества $\mathcal{D}^s$. Читателю предлагается самостоятельно доказать, что для каждого направления неограниченности s^t многогранного выпуклого множества $\mathcal{D}^s$, направленного вдоль его ребра, найдется такой образующий вектор $\hat{z}_i^s$ конуса K_0^s и такое число $\lambda > 0$, что $s^t = \lambda \hat{z}_i^s$. — Упражнение 6.5. Можно считать, что $\lambda = 1$, т. е. определить $\hat{z}_i^s$ равным s^t , и сформировать столбец $\hat{a}^t = \hat{a}^{st}$ матрицы $\bar{A}'$, отвечающий образующему вектору $\hat{z}_i^s$ конуса K_0^s в соответствии с рекомендациями п. 2.3. Из формулы (15) следует, что $d_0(\mathcal{B}) \hat{a}^t = y^s(s^t) < 0$. Поэтому введение вектора $\hat{a}^t$ в базис угловой задачи либо приведет к новому базису $\mathcal{B}'$ и допустимому базисному плану $\alpha(\mathcal{B}')$, для которого $g(\alpha(\mathcal{B}')) > g(\alpha(\mathcal{B}))$, либо, если $B^{-1} \hat{a}^{st} \leq 0$, будет обнаружена неограниченность сверху g на $\mathcal{G}$, а значит, и неограниченность сверху f на $\mathcal{D}$. В первом случае формирование базиса $\mathcal{B}'$ и заполнение табл. $T^2(\mathcal{B}')$ следует осуществлять в соответствии с рекомендациями п. 4.4.3.

Подведем итоги. Описанная вычислительная процедура позволяет на каждой итерации алгоритма решения угловой задачи проверить выполнение неравенства $a_0(\mathcal{B}) = d_0(\mathcal{B})\bar{A}' \geq 0$. Если оно выполняется, то $\alpha(\mathcal{B})$ — решение задачи $(\mathcal{G}, g)$, и можно по формулам (10) при $s \in 1:r$ сформировать оптимальный план задачи $(\mathcal{D}, f)$. В противном случае генерируется столбец матрицы A' (и соответствующий столбец матрицы $\bar{A}'$), введение которого в базис можно осуществить в соответствии со стандартной процедурой симплекс-метода. Конечность алгоритма в случае невырожденности задачи $(\mathcal{G}, g)$ и каждого из блоков $\mathcal{D}^s$, $s \in 1:r$, следует из конечности алгоритма симплекс-метода для невырожденных задач. Проблемы, возникающие в случае вырожденности задачи $(\mathcal{G}, g)$ или какого-либо из блоков $\mathcal{D}^s$, те же, что и в алгоритме симплекс-метода (см. п. 4.3.4 и § 7 гл. 4).

Остается выяснить, как строятся начальный допустимый базис $\mathcal{B}_1$ задачи $(\mathcal{G}, g)$ и порождающие такой базис система вершин многогранников V^s и образующих векторов конусов K_0^s . Решение этой задачи обычно проводится методом минимизации невязок. Рассматривается вспомогательная задача

$$(\mathcal{G}', \psi) : \psi(z') = -\sum_{i=1}^{m'} z'[n' + i] \rightarrow \max, \\ \mathcal{G}' = \{z' \in R^{n'+n'+m'} : A''z' = b', z' \geq 0\}$$

($n' = \sum_{s=1}^r k_s$, $\bar{n}' = \sum_{s=1}^r h_s$, $m' = m_0 + r$), где матрица $A'' = [A', I]$, I — единичная матрица порядка m' . Для этой задачи (в предположении $b^0 \geq 0$) система из m' последних столбцов матрицы A'' очевидно составляет допустимый базис. Решая ее методом декомпозиции, в случае невырожденности задачи $(\mathcal{G}', \psi)$ и блоков $\mathcal{D}^s$, $s \in 1:r$, после конечного числа итераций либо сформируем допустимый базис задачи $(\mathcal{G}, g)$, либо установим, что $\mathcal{G} = \emptyset$ ($\mathcal{D} = \emptyset$). Обращаем внимание на то, что коэффициенты $c[s, t]$ целевой функции ψ задачи $(\mathcal{G}', \psi)$ равны нулю для $s \in 1:r$ и $c[n' + i] = -1$ при $i \in 1:m'$.

2.6. Пример. Проиллюстрируем алгоритм метода декомпозиции. Для этого рассмотрим следующую блочную задачу:

$$(\mathcal{D}, f) : f(x) = 2x_1 + 6x_2 - 3x_3 - x_4 \rightarrow \max, \\ \mathcal{D} = \{x \in R^4 : 3x_1 + 2x_2 + 8x_3 + 6x_4 = 23, \\ x_1 + 2x_2 = 3, \\ 3x_3 + x_4 = 3, \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0\}, \quad (16)$$

где $m_0 = 1$, $r = 2$. Легко видеть, что блоки задачи (16) — выпуклые многогранники в R^2 :

$$\mathcal{D}^1 = \{x^1 = (x_1, x_2) \in R^2 : x_1 + 2x_2 = 3, x_1 \geq 0, x_2 \geq 0\}, \\ \mathcal{D}^2 = \{x^2 = (x_3, x_4) \in R^2 : 3x_3 + x_4 = 3, x_3 \geq 0, x_4 \geq 0\}.$$

Оба они являются отрезками. Первый соединяет базисные планы $u_1^1 = (3, 0)$ и $u_2^1 = (0, 3/2)$, второй — $u_1^2 = (1, 0)$ и $u_2^2 = (0, 3)$.

Поскольку известны все вершины, то для задачи (16) легко построить угловую задачу. Ее размерность 3×4 , т. е. $m' = 3$, $n' = 4$. Таким образом, в данном примере угловая задача имеет такую же размерность, как и блочная. Поэтому текст настоящего пункта следует рассматривать только как иллюстрацию идей метода разложения.

Для нахождения начального допустимого базиса $\mathcal{B}_1$ угловой задачи используется метод минимизации невязок. Решение вспомогательной задачи приведено в табл. (17). Она состоит

$T^2(\mathcal{B}_0) =$		-25	-1	-1	-1	-1	0	-19
	$n' + 1$	23	0	1	0	0	18	18
	$n' + 2$	1	0	0	1	0	0	0
	$n' + 3$	1	0	0	0	1	1	1*
$T^2(\mathcal{B}'_0) =$		-6	-1	-1	-1	18	0	-10
	$n' + 1$	5	0	1	0	-18	9	9*
	$n' + 2$	1	0	0	1	0	1	1
	[1, 2]	1	0	0	0	1	0	0
$T^2(\mathcal{B}''_0) =$		-4/9	-1	1/9	-1	-2	0	-10/9
	[2, 1]	5/9	0	1/9	0	-2	8	-10/9
	$n' + 2$	4/9	0	-1/9	1	2	0	10/9*
	[1, 2]	1	0	0	0	1	1	1
$T^2(\mathcal{B}'''_0) =$		0	-1	0	0	0		
	[2, 1]	1	0	0	1	0		
	[3, 2]	2/5	0	-1/10	9/10	9/5		
	[1, 2]	3/5	0	1/10	-9/10	-4/5		

(17)

из табл. $T^2(\mathcal{B}_0)$, $T^2(\mathcal{B}'_0)$, $T^2(\mathcal{B}''_0)$ и $T^2(\mathcal{B}'''_0)$, где $N(\mathcal{B}_0) = \{n' + 1, n' + 2, n' + 3\}$, а $\mathcal{B}'''_0 = \mathcal{B}_1$ — допустимый базис угловой задачи.

Поскольку $m_0 = 1$, $r = 2$, то из верхней строки табл. $T^2(\mathcal{B}_0)$ можно выписать векторы $v(\mathcal{B}_0) = (-1)$ и $w(\mathcal{B}_0) = (-1, -1)$. Для проверки оптимальности плана $z'(\mathcal{B}_0)$ вспомогательной

задачи решаются следующие задачи:

$$\begin{aligned}(\mathcal{D}^1, -y^1, \mathcal{B}_0): -y^1(x^1) &= -v(\mathcal{B}_0) A_{01}x^1 = 3x_1 + 2x_2 \rightarrow \max, \\ \mathcal{D}^1 &= \{x^1 \in R^2: x_1 + 2x_2 = 3, x_1 \geq 0, x_2 \geq 0\}; \\ (\mathcal{D}^2, -y^2, \mathcal{B}_0): -y^2(x^2) &= -v(\mathcal{B}_0) A_{02}x^2 = 8x_3 + 6x_4 \rightarrow \max, \\ \mathcal{D}^2 &= \{x^2 \in R^2: 3x_3 + x_4 = 3, x_3 \geq 0, x_4 \geq 0\}.\end{aligned}$$

Легко видеть, что решением первой из них является вершина $u_1^1 = (3, 0)$, решением второй — вершина $u_2^2 = (0, 3)$. При этом $-y^1(u_1^1) = 9$, $-y^2(u_2^2) = 18$. Так как $y^1(u_1^1) + w(\mathcal{B}_0)[1] = -10 < 0$, $y^2(u_2^2) + w(\mathcal{B}_0)[2] = -19 < 0$, то в базис можно вводить векторы a^{11} или a^{22} . Полагаем $l = [2, 2]$. Вектор $\bar{a}^l = \bar{a}^{22}$, отвечающий вершине u_2^2 , формируется в соответствии с рекомендациями п. 2.3. Переход к базису $\mathcal{B}'_0$ и заполнение табл. $T^2(\mathcal{B}'_0)$ осуществляются в соответствии с рекомендациями алгоритма п. 4.4.3.

На второй итерации для проверки оптимальности плана $z'(\mathcal{B}'_0)$ решаются задачи

$$\begin{aligned}(\mathcal{D}^1, -y^1, \mathcal{B}'_0): -y^1(x^1) &= 3x_1 + 2x_2 \rightarrow \max, \\ \mathcal{D}^1 &= \{x^1 \in R^2: x_1 + 2x_2 = 3, x_1 \geq 0, x_2 \geq 0\}; \\ (\mathcal{D}^2, -y^2, \mathcal{B}'_0): -y^2(x^2) &= 8x_3 + 6x_4 \rightarrow \max, \\ \mathcal{D}^2 &= \{x^2 \in R^2: 3x_3 + x_4 = 3, x_3 \geq 0, x_4 \geq 0\}.\end{aligned}$$

Их решения соответственно $u_1^1 = (3, 0)$ и $u_2^2 = (0, 3)$. Так как $y^1(u_1^1) + w(\mathcal{B}'_0)[1] = -10 < 0$ (естественно, $y^2(u_2^2) + w(\mathcal{B}'_0)[2] = 0$, поскольку вектор a^{22} , который формировался по вершине u_2^2 , содержится в базисе), то план $z'(\mathcal{B}'_0)$ неоптимален. Полагаем $l = [1, 1]$. Вектор $\bar{a}^l$ строится по вершине u_1^1 .

Аналогично для проверки оптимальности плана $z'(\mathcal{B}''_0)$ решаются задачи

$$\begin{aligned}(\mathcal{D}^1, -y^1, \mathcal{B}''_0): -y^1(x^1) &= -\frac{1}{3}x_1 - \frac{2}{9}x_2 \rightarrow \max, \\ \mathcal{D}^1 &= \{x^1 \in R^2: x_1 + 2x_2 = 3, x_1 \geq 0, x_2 \geq 0\}; \\ (\mathcal{D}^2, -y^2, \mathcal{B}''_0): -y^2(x^2) &= -\frac{8}{9}x_3 - \frac{2}{3}x_4 \rightarrow \max, \\ \mathcal{D}^2 &= \{x^2 \in R^2: 3x_3 + x_4 = 3, x_3 \geq 0, x_4 \geq 0\}.\end{aligned}$$

Их решения соответственно $u_1^1 = (0, \frac{3}{2})$ и $u_2^2 = (1, 0)$. Поскольку $y^1(u_1^1) + w(\mathcal{B}''_0)[1] = -\frac{2}{3} < 0$, $y^2(u_2^2) + w(\mathcal{B}''_0)[2] = -\frac{10}{9} < 0$, то план $z'(\mathcal{B}''_0)$ неоптимален. Полагаем $l = [2, 1]$, т. е. вектор $\bar{a}^l$ формируется по вершине u_2^2 . После перехода к базису $\mathcal{B}'''_0$ (табл. $T^2(\mathcal{B}'''_0)$) очевидно получаем оптимальный план $z'(\mathcal{B}'''_0)$, так как $v(\mathcal{B}'''_0) = (0)$ и $w(\mathcal{B}'''_0) = (0, 0)$. Базис $\mathcal{B}_1 = \mathcal{B}'''_0$ является

допустимым базисом угловой задачи, поскольку все векторы базиса $\mathcal{B}_0''$ отвечают вершинам многогранников $\mathcal{D}^1$ и $\mathcal{D}^2$.

Переход к решению угловой задачи осуществляется так же, как это рекомендовано в п. 4.4.8. В частности, поскольку необходимо вычислить вектор $u(\mathcal{B}_1) = c(\mathcal{B}_1) B_1^{-1}$, а все вершины, по которым сформировано базисное множество, известны, то элементы вектора $c(\mathcal{B}_1)$ можно определить по формуле (5): $c(\mathcal{B}_1) = (6, -3, -3)$, $u(\mathcal{B}_1) = (0, 6, -3)$.

В табл. (18), состоящей из табл. $T^2(\mathcal{B}_1)$ и $T^2(\mathcal{B}_2)$, приведено

$$T^2(\mathcal{B}_1) = \begin{array}{c|cccccc|cc} & 3 & -1 & 0 & 6 & -3 & 9 & -3 \\ \hline [2, 1] & 1 & 0 & 0 & 1 & 0 & 3 & 1 \\ [3, 2] & 2/5 & 0 & -1/10 & 9/10 & 9/5 & 1 & 3/5^* \\ [1, 2] & 3/5 & 0 & 1/10 & -9/10 & -4/5 & 0 & -3/5 \end{array}$$

$$T^2(\mathcal{B}_2) = \begin{array}{c|ccccc|} & 5 & -1 & -1/2 & 21/2 & 6 \\ \hline [2, 1] & 1/3 & 0 & 1/6 & -1/2 & -3 \\ [4, 1] & 2/3 & 0 & -1/6 & 3/2 & 3 \\ [1, 2] & 1 & 0 & 0 & 0 & 1 \end{array}$$

(18)

(18)

решение угловой задачи. Для проверки оптимальности плана $\alpha(\mathcal{B}_1)$ решаются задачи

$$(\mathcal{D}^1, -y^1, \mathcal{B}_1): -y^1(x^1) = [-v(\mathcal{B}_1) A_{01} + c_1] x^1 = 2x_1 + 6x_2 \rightarrow \max,$$

$$\mathcal{D}^1 = \{x^1 \in R^2: x_1 + 2x_2 = 3, x_1 \geq 0, x_2 \geq 0\};$$

$$(\mathcal{D}^2, -y^2, \mathcal{B}_1): -y^2(x^2) = [-v(\mathcal{B}_1) A_{02} + c_2] x^2 = -3x_3 - x_4 \rightarrow \max,$$

$$\mathcal{D}^2 = \{x^2 \in R^2: 3x_3 + x_4 = 3, x_3 \geq 0, x_4 \geq 0\}.$$

Их решения соответственно $u_1^1 = (0, 3/2)$ и $u_2^2 = (0, 3)$ (или $u_1^1 = (1, 0)$). Так как $y^1(u_1^1) + w(\mathcal{B}_1)[1] = -3 < 0$, то план $\alpha(\mathcal{B}_1)$ неоптимален. Положим $l = [1, 2]$, т. е. вектор $\bar{a}^l$ строится по вершине u_1^1 . Осуществляется переход к базису $\mathcal{B}_2$, $N(\mathcal{B}_2) = \{[1, 1], [1, 2], [2, 2]\}$.

Решениями задач

$$(\mathcal{D}^1, -y^1, \mathcal{B}_2): -y^1(x^1) = 7/2 x_1 + 7x_2 \rightarrow \max, x^1 \in \mathcal{D}^1,$$

$$(\mathcal{D}^2, -y^2, \mathcal{B}_2): -y^2(x^2) = x_3 + 2x_4 \rightarrow \max, x^2 \in \mathcal{D}^2$$

являются вершины $u_1^1 = (3, 0)$ (или $u_2^1 = (0, 3/2)$) и $u_2^2 = (0, 3)$. План $\alpha(\mathcal{B}_2)$ оптимален, поскольку $y^1(u_1^1) + w(\mathcal{B}_2)[1] = 0$ и $y^2(u_2^2) + w(\mathcal{B}_2)[2] = 0$. Решение блочной задачи вычисляется по формулам (10) при $s = 1, 2$. При этом используются базисные компоненты плана $\alpha(\mathcal{B}_2)$ и вершины, которым отвечают столбцы базиса: $\bar{x}^1(\alpha(\mathcal{B}_2)) = 1/3 \cdot (3, 0) + 2/3 \cdot (0, 3/2) = (1, 1)$; $\bar{x}_2(\alpha(\mathcal{B}_2)) =$

$= 1 \cdot (0, 3) = (0, 3)$; $\bar{x}(\alpha(\mathcal{B}_2)) = (\bar{x}^1(\alpha(\mathcal{B}_2)), \bar{x}^2(\alpha(\mathcal{B}_2))) = (1, 1, 0, 3)$. Максимальное значение целевой функции $f(\bar{x}(\alpha(\mathcal{B}_2))) = 5$.

§ 3. Графы. Транспортные сети и транспортные задачи

3.1. Графы. В настоящем и следующих пунктах определены только понятия, необходимые для дальнейшего изложения. Для более углубленного знакомства с теорией графов рекомендуем, например, [3, 34].

Будем говорить, что задан *граф* G , если определены непустое множество V *вершин графа*, множество D *дуг* или *ребер графа* и отображение Γ , ставящее в соответствие каждому d из D упорядоченную (v', v'') или неупорядоченную $\{v', v''\}$ пару элементов из V . Обычно используется обозначение $G = (V, D, \Gamma)$, или просто (V, D, Γ) .

Если $Y \subset D$ и отображение Γ_Y является сужением отображения Γ на множество Y , то граф (V, Y, Γ_Y) называется *частичным графом* графа (V, D, Γ) .

В дальнейшем будут рассматриваться графы с конечным множеством D и с конечным множеством вершин V . Такие графы называются *конечными*. Число $n = |V|$ называется *порядком графа* G .

Если отображение Γ каждому элементу d из D ставит в соответствие упорядоченную пару $\Gamma(d) = (v', v'')$ вершин из V , то элементы множества D будем называть *дугами*, а граф G — *ориентированным графом*, или *орграфом*. Вершины v' и v'' называются *граничными*, или *смежными*, вершинами дуги d , при этом v' называют *началом дуги* d , а v'' — ее *концом*. Дугу d назовем *инцидентной* вершинам v' и v'' .

Если отображение Γ каждому элементу $d \in D$ ставит в соответствие неупорядоченную пару $\Gamma(d) = \{v', v''\}$, то элементы множества D называют *ребрами*, а граф G — *неориентированным графом*, или *неорграфом*. Вершины v' и v'' (их порядок не существует) называют *граничными*, или *смежными*, вершинами ребра d .

Две разные дуги (два ребра) называют *смежными*, если они имеют общую граничную вершину. Иногда, если не существует двух разных дуг (ребер) d_1 и d_2 , для которых $\Gamma(d_1) = \Gamma(d_2)$, дугу (ребро) d обозначают парой ее (его) смежных вершин: $d = (v', v'')$ (или $d = \{v', v''\}$). Ниже, если это не может вызвать недоразумений, будем пользоваться такими обозначениями дуги и ребра.

Если число n вершин графа $G = (V, D, \Gamma)$ не очень велико, то с ним можно связать некоторую иллюстрацию. Для этого на листе бумаги расставляются n точек — их отождествляют с вершинами графа. Если $d \in D$, $\Gamma(d) = (v', v'')$, то точки, отве-

чающие вершинам v' и v'' , соединяют непрерывной линией, на которой указывают направление от v' к v'' , если G — оргграф. Эти линии отождествляют с дугами (ребрами) графа G . Линии стараются проводить так, чтобы они пересекались только в точках, отвечающих вершинам графа. Иногда это оказывается невозможным (граф G не является *планарным*). Тогда специальным способом отмечают точки пересечения дуг (ребер), не отвечающие вершинам графа.

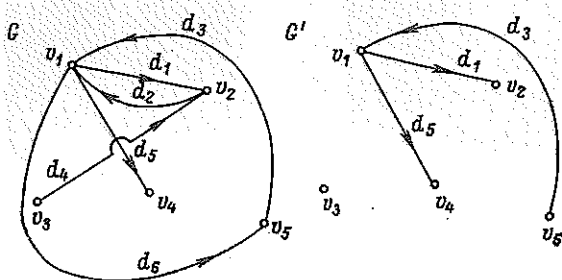


Рис. 26.

На рис. 26 изображены оргграф $G = (V, D, \Gamma)$ и его частичный граф $G' = (V, Y, \Gamma_Y)$, у которых множество вершин $V = \{v_1, v_2, v_3, v_4, v_5\}$, множества дуг $D = \{d_1, d_2, d_3, d_4, d_5, d_6\}$ и $Y = \{d_1, d_3, d_5\}$. Направления на дугах указаны стрелками, v_1 и v_2 — начало и конец дуги d_1 , v_2 и v_1 — начало и конец дуги d_2 , v_3 и v_2 — начало и конец дуги d_4 и т. д. Смежными являются, например, дуги d_5 и d_6 .

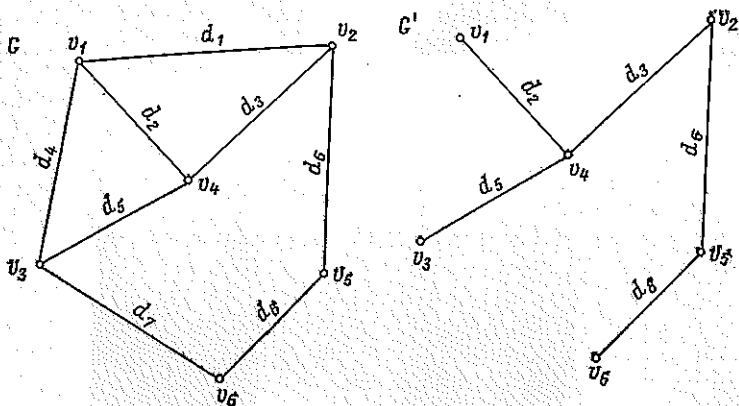


Рис. 27.

На рис. 27 изображены неоргграф $G = (V, D, \Gamma)$ и его частичный граф $G' = (V, Y, \Gamma_Y)$. Множество их вершин

$V = \{v_1, v_2, v_3, v_4, v_5, v_6\}$, множество ребер $D = \{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8\}$ и $Y = \{d_2, d_3, d_5, d_6, d_8\}$, $\Gamma(d_1) = \{v_1, v_2\}$, $\Gamma(d_2) = \{v_1, v_4\}$ и т. д.

3.2. Пути и контуры. В этом пункте будем рассматривать только оргграфы. Говорят, что дуга $d = (v', v'')$ оргграфа G *исходит* из вершины v' и *входит* в вершину v'' . Последовательность дуг $d_1, d_2, d_3, \dots, d_k$ образует *путь* в графе G , если каждая из дуг $d_2, d_3, \dots, d_k$ исходит из вершины, в которую входит предыдущая. Такой путь будем обозначать $\mu = [d_1, d_2, \dots, d_k]$ или $\mu = [v_1, v_2, \dots, v_{k+1}]$, где $(v_1, v_2) = d_1$, $(v_2, v_3) = d_2$ и т. д. Будем называть путь μ *элементарным*, если все его вершины разные, *простым*, если он составлен из разных дуг, и *составным* в противном случае. *Контуром* назовем путь $\mu = [v_1, v_2, \dots, v_k]$, у которого $v_1 = v_k$.

Назовем граф $G = (V, D, \Gamma)$ *симметрическим*, если для любой дуги $(v', v'') \in D$ существует дуга $(v'', v') \in D$. В таком графе каждая пара смежных вершин соединена двумя противоположно ориентированными дугами, поэтому симметрический граф удобно рассматривать как неорграф, предполагая ориентацию определенной на каждом ребре в двух направлениях. Граф G назовем *антисимметрическим*, если для любой дуги $(v', v'') \in D$ дуга $(v'', v') \notin D$.

Наконец, назовем граф (V, D, Γ) *полным*, если для любой пары вершин v', v'' из условия $(v', v'') \in D$ следует $(v'', v') \in D$.

3.3. Цепи и циклы. Пусть G — неорграф. *Цепью* в G будем называть такую последовательность ребер $d_1, d_2, \dots, d_k$, в которой у каждого из ребер d_i , $i \in 2: (k-1)$, одна из граничных вершин совпадает с граничной вершиной ребра d_{i-1} , а другая — с граничной вершиной ребра d_{i+1} . *Элементарные* и *простые цепи* определяются аналогично соответствующим понятиям предыдущего пункта.

Цепь $\{v_1, v_2\}, \{v_2, v_3\}, \dots, \{v_k, v_{k+1}\}$, состоящая не менее чем из двух ребер, в которой $v_1 = v_{k+1}$, называется *циклом*. Цикл называется *простым*, если все ребра цикла различны, и *составным* в противном случае. Цикл называется *элементарным*, если все его вершины (кроме начальной и конечной) разные.

Неорграф (V, D, Γ) , у которого любая пара различных вершин может быть соединена цепью, называется *связным*. Легко проверить, что графы G и G' , изображенные на рис. 27, являются связными. Любой оргграф можно однозначно превратить в неорграф, заменив дуги ребрами (обратное преобразование неоднозначно!). Оргграфу G' , изображенному на рис. 26, отвечает несвязный неорграф. Будем говорить, что оргграф *связен*, если ему отвечает связный неорграф.

В неорграфе $G = (V, D, \Gamma)$ для любого $x \in V$ можно выделить множество вершин V_x , состоящее из вершины x и всех вершин, соединенных с ней цепью, и построить граф G_x , получаю-

щийся удалением из графа G вершин $V \setminus V_x$ и инцидентных им ребер. Граф G_x называется *компонентой связности*, или просто *компонентой графа G* . Легко доказать, что *граф связан тогда и только тогда, когда он состоит из единственной компоненты.* — Упражнение 6.6.

Графы, изображенные на рис. 27, — *связные* ($V_{v_1} = V_{v_2} = \dots = V_{v_n} = V$). Орграфу G' на рис. 26 отвечает неорграф, состоящий из двух компонент связности.

3.4. Деревья. Поскольку в дальнейшем нас будут интересовать только простые циклы, для краткости будем называть их *циклами*.

Деревом назовем конечный связный неорграф без циклов. Непосредственно проверяется, что граф G (рис. 27) деревом не является, а граф G' (рис. 27) — *дерево*. Дерево обладает рядом важных характеристических свойств, каждое из которых можно использовать как его определение. Так что если $G = (V, D, \Gamma)$ — неорграф порядка $n = |V|$, то следующие характеристические свойства дерева эквивалентны.

1°. G *связен и не содержит циклов.*

2°. G *связен и имеет $n - 1$ ребер.*

3°. G *не содержит циклов и имеет $n - 1$ ребер.*

4°. G *не содержит циклов, но добавление ребра между любыми двумя вершинами приводит к их появлению, причем только одного — элементарного.*

5°. G *связен, но теряет это свойство после удаления любого ребра.*

6°. В G *любая пара различных вершин соединена цепью, и только одной.*

Приведем указания к проверке эквивалентности 1°–6°.

Из 1° следует 2° ($1^\circ \Rightarrow 2^\circ$). Доказывается по индукции. Она проводится по числу вершин. — Упражнение 6.7.

Из 2° следует 3°. Если в G есть цикл, то удаление его ребра не нарушает связности графа. Последовательно удаляя ребра из оставшихся циклов, получаем связный граф без циклов, с числом ребер меньше $n - 1$, что противоречит $1^\circ \Rightarrow 2^\circ$.

Из 3° следует 4°. Отсутствие цикла после добавления ребра противоречит $1^\circ \Rightarrow 2^\circ$. Появление двух элементарных циклов возможно лишь тогда, когда G содержит цикл.

Из 4° следует 5°. Действительно, добавление ребра, соединяющего вершины двух разных компонент связности, не приводит к образованию нового цикла. Не нарушая связности, можно удалить только ребро цикла.

Из 5° следует 6°, так как в связном графе любые две вершины соединены цепью, а если две вершины связного графа соединены двумя цепями, то удаление любого ребра в любой из этих цепей не нарушает связности графа.

Наконец, из 6° следует 1°: связность графа очевидна, а любые две вершины цикла соединены двумя разными цепями. ■

Пусть $G = (V, D, \Gamma)$ — конечный неорграф. Частичный граф графа G , являющийся деревом, называется *частичным деревом* графа G .

7°. Граф G содержит частичное дерево тогда и только тогда, когда он *связен*.

Действительно, если G несвязен, то несвязны и все его частичные графы, и, следовательно, никакой частичный граф графа G не является деревом. Если G *связен*, то либо удаление любого его ребра приводит к нарушению связности, и в силу 5° G — дерево, либо можно удалить ребро, не нарушив связности графа. Продолжая такой процесс удаления ребер, через конечное число шагов получим частичное дерево. ■

Вершину v графа $G = (V, D, \Gamma)$ назовем *висячей*, если множество ребер, инцидентных вершине v , состоит из одного элемента.

8°. В каждом дереве, у которого $|V| \geq 2$, есть по крайней мере две висячие вершины. — Упражнение 6.8.

3.5. Транспортные сети. Поток. Транспортной сетью назовем конечный связный оргграф $G = (V, D, \Gamma)$ без дуг вида (v, v) , у которого каждой вершине $v \in V$ поставлено в соответствие число $b[v]$ — *потребность* в этой вершине, каждой дуге $d \in D$ поставлены в соответствие число $c[d]$ — *стоимость* транспортировки единицы груза по дуге d и число $h[d]$ — *пропускная способность дуги* d . Предполагается, что для некоторых или даже для всех дуг $h[d]$ может быть равно $+\infty$.

Транспортная сеть называется *сбалансированной*, если выполняется условие

$$\sum_{v \in V} b[v] = 0, \quad (1)$$

и *несбалансированной* (или *открытой*), если условие (1) не выполняется.

Каждой дуге d транспортной сети поставим в соответствие число $x[d]$. Набор чисел $x(D) = \{x[d], d \in D\}$ назовем *поток*ом транспортной сети. Поток является *допустимым*, если

$$0 \leq x[d] \leq h[d], \quad d \in D, \quad (2)$$

$$\sum_{d \in D_v^+} x[d] - \sum_{d \in D_v^-} x[d] = b[v], \quad v \in V, \quad (3)$$

где через D_v^+ обозначено множество дуг, входящих в вершину v , а через D_v^- — множество дуг, исходящих из нее, и *недопустимым*, если нарушается хотя бы одно из соотношений (2) или (3). Число $f(x(D)) = \sum_{d \in D} c[d] x[d]$ называется *стоимостью* потока $x(D)$.

3.6. Транспортные задачи. Пусть задана сбалансированная транспортная сеть (см п. 3.5). Экстремальная задача

$$(\mathcal{D}, f): f(x) = \sum_{d \in D} c[d] x[d] \rightarrow \min, \quad \mathcal{D} = \left\{ x = \right. \\ \left. = \{x[d], d \in D\} : \sum_{d \in D_v^+} x[d] - \sum_{d \in D_v^-} x[d] = b[v], v \in V; \quad (4) \right. \\ \left. 0 \leq x[d] \leq h[d], d \in D \right\},$$

называется *транспортной задачей в сетевой постановке с ограниченными сверху переменными* (см. п. 2.3.3), связанной с данной транспортной сетью.

В п. 2.3.2 рассматривалась транспортная задача в матричной постановке. Легко убедиться, что она является частным случаем

задачи (4). Действительно, пусть V^- — множество пунктов производства и V^+ — множество пунктов потребления ($|V^-| = m$, $|V^+| = n$) транспортной задачи в матричной постановке. Рассмотрим антисимметрический граф $G = (V, D, \Gamma)$, в котором $V = V^- \cup V^+$ и дуга $d = (v, w)$ в том и только том случае принадлежит

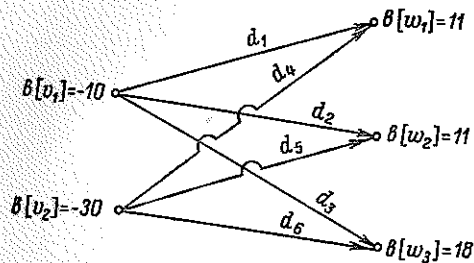


Рис. 28.

тогда D , если $v \in V^-$, а $w \in V^+$. Определим $b[v] = -a[i]$, если $v \in V^-$ и i — индекс пункта производства v , и $b[w] = a[j]$, если $w \in V^+$ и j — индекс пункта потребления w . Пропускную способность каждой дуги из D будем считать равной $+\infty$, а стоимость транспортировки определим в соответствии с ее значениями в матрице C транспортной задачи в матричной постановке. Экстремальная задача (4) на построенной таким образом транспортной сети очевидно эквивалентна исходной транспортной задаче в матричной постановке.

Приведем пример. Рассмотрим транспортную задачу в матричной постановке с $m=2$ пунктами производства и $n=3$ пунктами потребления, с объемами производства $a[1]=10$, $a[2]=30$ и потребления $b[1]=11$, $b[2]=11$, $b[3]=18$, и

пусть $C = \begin{pmatrix} 2 & 4 & 5 \\ 1 & 10 & 6 \end{pmatrix}$ — матрица стоимостей перевозок. Этой задаче соответствует сбалансированная транспортная сеть, изображенная на рис. 28. Здесь $V^- = \{v_1, v_2\}$, $V^+ = \{w_1, w_2, w_3\}$; $b[v_1] = -10$, $b[v_2] = -30$, $b[w_1] = 11$, $b[w_2] = 11$, $b[w_3] = 18$; $D = \{d_1 = (v_1, w_1), d_2 = (v_1, w_2), \dots\}$; $c[d_1] = 2$, $c[d_2] = 4$

и т. д. Транспортная задача в сетевой постановке $(\mathcal{D}, f)$:

$$\begin{aligned} f(x) &= 2x[d_1] + 4x[d_2] + 5x[d_3] + x[d_4] + 10x[d_5] + 6x[d_6] \rightarrow \min, \\ \mathcal{D} &= \{x = \{x[d], d \in D\} : x[d_1] + x[d_4] = 11, x[d_2] + x[d_5] = 11, \\ &\quad x[d_3] + x[d_6] = 18, -x[d_1] - x[d_2] - x[d_3] = -10, \\ &\quad -x[d_4] - x[d_5] - x[d_6] = -30, x[d_i] \geq 0, i \in 1:6\} \end{aligned}$$

очевидно эквивалентна исходной транспортной задаче в матричной постановке.

3.7. Структура и свойства транспортной задачи. Предположим, что вершины транспортной сети (см. п. 3.5) занумерованы натуральными числами $v \in 1:m$, а дуги — числами $d \in 1:n$. Очевидно, что задача (4) является задачей линейного программирования с ограниченными сверху переменными (см. п. 5.1.1), т. е. ее можно записать в виде

$$(\mathcal{D}, f) : f(x) = -cx \rightarrow \max, \mathcal{D} = \{x \in R^n : Ax = b, 0 \leq x \leq h\},$$

положив $c = (c[1], c[2], \dots, c[n])$, $b^T = (b[1], b[2], \dots, b[m])$, $h^T = (h[1], \dots, h[n])$, $x^T = (x[1], \dots, x[n])$. Столбцы матрицы A соответствуют дугам транспортной сети, строки — ее вершинам. Каждый столбец матрицы A содержит ровно два ненулевых элемента. В столбце, записанном для дуги $d = (v', v'')$, стоит -1 в строке, отвечающей вершине v' , и $+1$ в строке, отвечающей вершине v'' . Таким образом, $A = (a[v, d])$, $v \in 1:m$, $d \in 1:n$, где $a[v, d] = -1$, если v — начало дуги d , $a[v, d] = 1$, если v — ее конец, и $a[v, d] = 0$, если дуга d не инцидентна вершине v . Последние формулы можно использовать для генерирования матрицы A . Очевидно, что строки матрицы A линейно зависимы (их сумма равна нулевой строке), так что ее ранг $r(A) \leq m - 1$. Ниже будет установлено, что $r(A) = m - 1$.

Граф G , с которым связана транспортная задача $(\mathcal{D}, f)$, по определению является орграфом. Как уже отмечалось, снятие ориентации с дуг графа G определяет неорграф, который обозначается той же буквой G . Именно в таком смысле мы иногда будем называть дуги графа G ребрами, говорить о циклах в графе G или его частичных графах и т. п.

Рассмотрим некоторое множество $\mathcal{A}$ столбцов матрицы A и частичный граф $G_{\mathcal{A}}$ графа G , содержащий дуги, отвечающие столбцам из множества $\mathcal{A}$.

1°. Множество $\mathcal{A}$ столбцов матрицы A линейно независимо в том и только том случае, если граф $G_{\mathcal{A}}$ не содержит циклов.

Доказательство. Необходимость. Пусть Z — какой-нибудь цикл в графе $G_{\mathcal{A}}$ и $a^1, a^2, \dots, a^k$ — столбцы матрицы A , отвечающие ребрам цикла Z . Определим на ребрах цикла Z такую ориентацию, чтобы начало каждой следующей дуги совпадало с концом предыдущей, и положим $a_{\alpha} = 1$, если

первоначальная ориентация дуги d совпадает с новой, и $\alpha_d = -1$ в противном случае. Тогда очевидно $\sum_{d=1}^k \alpha_d a^d = 0$.

Достаточность. Предположим, что система столбцов $\mathcal{B} = \{a^1, a^2, \dots, a^q\}$ матрицы A линейно зависима, так что $\sum_{d=1}^q \alpha_d a^d = 0$ и $\alpha_d \neq 0$, $d = 1:q$. Очевидно, что $q > 1$. Предположим, что -1 в столбце a^1 стоит на месте v_1 , а 1 — на месте v'_1 . Тогда среди столбцов $a^2, \dots, a^q$ найдется столбец a^{d_1} , у которого на месте v'_1 стоит 1 или -1 . Пусть в этом столбце -1 или 1 стоит на месте v'_2 . Тогда в системе $\mathcal{B}$ найдется столбец, отличный от a^{d_1} , в котором на месте v'_2 стоит 1 или -1 . Продолжая этот процесс, получаем такую последовательность столбцов $a^1, a^{d_1}, a^{d_2}, \dots$, в которой каждая пара соседних столбцов имеет на одном и том же месте 1 или -1 , а последовательность ребер $1, d_2, d_3, \dots$ графа $G_{\mathcal{B}}$, отвечающая последовательности столбцов $a^1, a^{d_1}, a^{d_2}, \dots$, образует цепь. Поскольку множество $\mathcal{B}$ конечно, какое-то ребро d_k впервые встретится в этой последовательности повторно. Тогда последовательность ребер $d_k, d_{k+1}, \dots, d_k$ образует цикл в графе $G_{\mathcal{B}}$. ■

2°. Ранг $r(A)$ матрицы A равен $m - 1$.

Доказательство. По определению транспортной сети граф G связан и в силу утверждения 7° п. 3.4 содержит частичное дерево G' . Так как число вершин графа G равно m , то G' содержит $m - 1$ ребер (см. утверждение 2° п. 3.4), и система столбцов $\mathcal{B}'$ матрицы A , отвечающая ребрам графа G' , линейно независима (см. 1°). Таким образом, в матрице A есть система из $m - 1$ линейно независимых столбцов, и поскольку $r(A) \leq m - 1$, то $r(A) = m - 1$. ■

Из утверждения 2° следует, что для сбалансированной транспортной сети (см. (1)) система уравнений в (4) совместна, и исключение из нее любого уравнения приводит к равносильной системе уравнений.

Частичное дерево графа G будем называть *остовом транспортной сети*, или просто *остовом*. Из утверждений 3° п. 3.4 и 1° настоящего пункта следует, что каждой системе $\mathcal{B}$ из $m - 1$ линейно независимых столбцов матрицы A , которые можно отождествить с соответствующими дугами графа G , отвечает остов, и обратно, каждому остову отвечает система из $m - 1$ линейно независимых столбцов матрицы A (дуг графа G). Существование остова следует из утверждения 2°.

Систему $\mathcal{B}$ из $m - 1$ линейно независимых столбцов матрицы A (или соответствующих дуг графа G) назовем *базисом задачи (4)*. Ей отвечает базис задачи, которая получается из (4) после исключения одного (любого) уравнения. Остов, порождаемый базисом $\mathcal{B}$, обозначим через $G_{\mathcal{B}} = (V, \mathcal{B}, \Gamma_{\mathcal{B}})$. Сказанное выше можно переформулировать в виде следующего утверждения.

3°. Каждому базису $\mathcal{B}$ задачи (4) отвечает остов $G_{\mathcal{B}} = (V, \mathcal{B}, \Gamma_{\mathcal{B}})$, и соответствие $\mathcal{B} \mapsto G_{\mathcal{B}}$ взаимно однозначно. Множество базисов задачи (4) (остовов транспортной сети) непусто.

Итак, остовы транспортной сети можно отождествить с отвечающими им базисами задачи линейного программирования с ограниченными сверху переменными.

Рассмотрим теперь транспортную сеть на графе $G = (V, D, \Gamma)$, ее остов $G_{\mathcal{B}} = (V, \mathcal{B}, \Gamma_{\mathcal{B}})$ и некоторое множество $N^* \subset D \setminus \mathcal{B}$. Поток $x(\mathcal{B}, N^*)$ назовем **-базисным*, если

$$x(\mathcal{B}, N^*)[d] = h[d], \quad d \in N^*, \quad (5)$$

$$x(\mathcal{B}, N^*)[d] = 0, \quad d \in D \setminus (\mathcal{B} \cup N^*). \quad (6)$$

Компоненту $x(\mathcal{B}, N^*)[d]$ потока $x(\mathcal{B}, N^*)$ назовем *базисной*, если $d \in \mathcal{B}$, и *небазисной*, если $d \in N^*$.

Допустимый **-базисный* поток назовем *невырожденным*, если $0 < x(\mathcal{B}, N^*) < h[d]$ для $d \in \mathcal{B}$, и *вырожденным* в противном случае. Транспортная задача (4) называется *невырожденной*, если множество $\mathcal{D}$ не содержит вырожденных **-базисных* потоков, и *вырожденной* в противном случае.

В транспортной задаче в матричной постановке (см. п. 2.3.2)

$$\begin{aligned} (\mathcal{D}, f): f(x) &= \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j] \rightarrow \min, \\ \mathcal{D} &= \{x \in R^{mn} : \sum_{j=1}^n x[i, j] = a[i], \quad i \in 1:m; \\ &\quad \sum_{i=1}^m x[i, j] = b[j], \quad j \in 1:n; \quad x \geq 0\} \end{aligned} \quad (7)$$

$(\sum_{i=1}^m a[i] = \sum_{j=1}^n b[j])$ остову $G_{\mathcal{B}}$, $\mathcal{B} = \{(i_1, j_1), (i_2, j_2), \dots, (i_{m'-1}, j_{m'-1})\}$, $m' = m + n$, $i_k \in 1:m$, $j_k \in 1:n$ для $k \in 1:m'$, естественным образом ставится в соответствие базисный поток $x(\mathcal{B})$. Допустимый базисный поток $x(\mathcal{B})$ называется *невырожденным*, если $x(\mathcal{B})[i, j] > 0$ для $(i, j) \in \mathcal{B}$. Задача (7) называется *невырожденной*, если она не имеет вырожденных допустимых базисных потоков, и *вырожденной* в противном случае.

При анализе и решении экстремальных задач важно иметь эффективные критерии проверки их свойств. Например, в канонической задаче линейного программирования, методы решения которой изучались в гл. 4, проверка линейной независимости системы столбцов матрицы ограничений задачи в общем случае связана с вычислением ранга некоторых подматриц. В транспортных задачах такую проверку можно осуществить проще: необходимо лишь проверить наличие циклов в соответствующем графе. Аналогично допустимый план x задачи (7) является базисным в том и только том случае, если его ненулевым компонентам отвечает частичный граф без циклов.

Для транспортных задач можно ответить просто и на некоторые другие вопросы, в общем случае требующие специальных исследований. Приведем лишь следующие утверждения.

4°. Множество $\mathcal{D}$ допустимых планов транспортной задачи в матричной постановке непусто. Действительно, симуляторы 1° и 2° п. 1.3 устанавливают (алгоритмически) непустоту множества $\mathcal{D}$. ■

5°. В задаче (4) условие $\sum_{d \in D} h[d] \geq \sum_{v: b[v] > 0} b[v]$ необходимо для непустоты множества $\mathcal{D}$. Утверждение очевидно. ■

6°. Задача (7) имеет оптимальный план. — Упражнение 6.9.

7°. Задача (4) имеет оптимальный план, если $\mathcal{D} \neq \emptyset$. — Упражнение 6.10.

О вырожденности. Для транспортной задачи в матричной постановке приведем следующее условие ее вырожденности (невырожденности).

Совокупность, состоящую из множества пунктов производства $I \subset 1:m$, $I \neq \emptyset$, $I \neq 1:m$, и множества пунктов потребления $J \subset 1:n$, $J \neq \emptyset$, $J \neq 1:n$, назовем автономным районом, если $\sum_{i \in I} a[i] = \sum_{j \in J} b[j]$.

8°. Задача (7) вырождена в том и только том случае, если существует автономный район.

Доказательство. Предположим, что задача (7) имеет автономный район. Не умаляя общности, можно считать, что $I = 1:k$, $1 \leq k < m$, $J = 1:r$, $1 \leq r < n$. Тогда симулятор 1° п. 1.3 очевидно формирует вырожденный поток.

Если $x(\mathcal{B})$ — вырожденный поток, то в остове $G_{\mathcal{B}}$ существует такая дуга (i_k, j_k) , что $x(\mathcal{B})[i_k, j_k] = 0$. В силу свойства 5° п. 3.4 удаление этой дуги нарушает связность графа $G_{\mathcal{B}}$. Каждая из компонент связности графа, полученного из графа $G_{\mathcal{B}}$ после удаления дуги (i_k, j_k) , очевидно является сбалансированной транспортной сетью, и, следовательно, в задаче существует автономный район. ■

3.8. Генерирование потоков. В машинных алгоритмах довольно часто применяют генерирование допустимого потока $x(\mathcal{B}, N^*)$ задачи (4) по остову $G_{\mathcal{B}}$ и множеству $N^* \subset D \setminus \mathcal{B}$. Для этого используют формулы (5), (6) и свойство 8° п. 3.4.

Алгоритм. 1. Для каждой дуги $d = (v', v'') \in N^*$ полагаем $x(\mathcal{B}, N^*) = h[d]$ и изменяем (соответствующим образом) на величину $h[d]$ потребности $b[v']$ и $b[v'']$. Аналогично для каждой дуги $d = (v', v'') \in D \setminus (\mathcal{B} \cup N^*)$ полагаем $x(\mathcal{B}, N^*)[d] = 0$.

Формирование базисных компонент потока $x(\mathcal{B}, N^*)$ осуществляется при уже преобразованных потребностях $b[v]$, $v \in V$.

2. В остове $G_{\mathcal{B}}$ выбираем произвольную висячую вершину $\bar{v} \in V$ (см. 8° п. 3.4) и инцидентную ей дугу $\bar{d} = (\bar{v}, \bar{v})$ (или $\bar{d} = (\bar{v}, \bar{v})$). Полагаем $x(\mathcal{B}, N^*)[\bar{d}] = |b[\bar{v}]|$ и изменяем потребность $b[\bar{v}]$ на величину $b[\bar{v}]$ ($b[\bar{v}]$ полагаем равной $b[\bar{v}] + b[\bar{v}]$).

3. Полагаем $\mathcal{B} = \mathcal{B} \setminus \{\bar{d}\}$, $V = V \setminus \{\bar{v}\}$. Переходим к рекомендации 2. ■

Очевидно, что если остову $G_{\mathcal{B}}$ и множеству N^* отвечает допустимый поток $x(\mathcal{B}, N^*)$, то его можно генерировать с помощью приведенной процедуры.

Аналогичным образом осуществляется формирование допустимого потока $x(\mathcal{B})$ задачи (7) по остову $G_{\mathcal{B}}$.

§ 4. Метод потенциалов

4.1. Предварительные замечания. В настоящем параграфе рассматривается алгоритм решения транспортной задачи в сетевой постановке с ограниченными сверху переменными $(\mathcal{D}, f)$:

$$\begin{aligned} f(x) = \sum_{d \in D} c[d] x[d] \rightarrow \min, \quad \mathcal{D} = \{x = \{x[d], d \in D\} : \\ \sum_{d \in D_v^+} x[d] - \sum_{d \in D_v^-} x[d] = b[v], v \in V; \\ 0 \leq x[d] \leq h[d], d \in D\} \end{aligned} \quad (1)$$

(см. п. 3.6). Если $\mathcal{D} \neq \emptyset$ и задача (1) невырождена, то алгоритм п. 4.2 позволяет за конечное число итераций q найти ее решение. Вычисления начинаются после построения базиса $\mathcal{B}_1$ и множества дуг N_1^* , которым отвечает допустимый $\ast$ -базисный поток $x(\mathcal{B}_1, N_1^*)$. В общем случае построение таких $\mathcal{B}_1$ и N_1^* — самостоятельная задача. Метод ее решения обсуждается в п. 4.5.

Принципиально алгоритм п. 4.2 ничем не отличается от алгоритма п. 5.1.3. Однако с вычислительной точки зрения он значительно проще последнего в силу определенных свойств матрицы ограничений задачи (1) (см. § 1 гл. 8).

Для небольшой транспортной сети метод допускает реализацию непосредственно на графе. В п. 4.3 приведены соответствующий пример и комментарий. Машинная реализация алгоритма связана с рядом специфических проблем: организация и упорядочение массивов исходной и текущей информации, ее экономный поиск в памяти ЭВМ, генерирование вычислительной информации, $\ast$ -базисных потоков, циклов и т. п. Их детальное решение зависит от конкретной ЭВМ, назначения программы, опыта и даже вкуса программиста. Здесь будет приведена лишь принципиальная вычислительная схема алгоритма.

Название метода обусловлено тем, что на каждой итерации для проверки оптимальности потока вычисляются определенные числа — потенциалы вершин графа (см. 3° и 4° п. 3.1.11).

4.2. Алгоритм. Предполагаются заданными остов $G_{\mathcal{B}_1} = (V, \mathcal{B}_1, \Gamma_{\mathcal{B}_1})$ транспортной сети и множество дуг $N_1^* \subset D \setminus \mathcal{B}_1$, $h[d] < +\infty$ для $d \in N_1^*$, которым соответствует допустимый $\ast$ -базисный поток $x(\mathcal{B}_1, N_1^*)$.

Этап 0. По остову $G_{\mathcal{A}}$ и множеству N^* формируется допустимый $\ast$ -базисный поток $x(\mathcal{B}_1, N^*)$ (см. п. 3.8). Полагаем $\mathcal{B} = \mathcal{B}_1$, $N^* = N^*$. На этапах 1—4 будет описана одна итерация метода: вычисление массива потенциалов $\{p[v], v \in V\}$, проверка оптимальности текущего $\ast$ -базисного потока $x(\mathcal{B}, N^*)$ и в случае неоптимальности потока $x(\mathcal{B}, N^*)$ — переход к новому остову $G_{\mathcal{B}'}$, множеству $N^{*'}$ и отвечающему им допустимому $\ast$ -базисному потоку $x(\mathcal{B}', N^{*'})$.

Этап 1. 1. Последовательно вычисляется массив потенциалов $\{p[v], v \in V\}$: а) для некоторого произвольного $v_0 \in V$ (например, для $v_0 = 1$) произвольно задается число $p[v_0]$ (например, $p[1] = 0$); б) если уже определен потенциал одной из смежных вершин v' или v'' , связанных дугой $d = (v', v'') \in \mathcal{B}$, то потенциал другой вершины определяется так, чтобы выполнялось равенство

$$p[v''] - p[v'] = c[d]. \quad (2)$$

Возможность вычисления массива потенциалов и однозначное их определение после выбора вершины v_0 и числа $p[v_0]$ следует из того, что любая пара вершин графа $G_{\mathcal{B}}$ связана единственной цепью (см. п. 3.4, 6°). Следовательно, потенциалы определяются с точностью до постоянного слагаемого, поскольку изменение числа $p[v_0]$ на некоторую величину приводит к изменению на ту же величину каждого из потенциалов $p[v]$, $v \in V$.

2. Для каждой дуги $d = (v', v'')$ положим

$$u[d] = c[d] - p[v''] + p[v']. \quad (3)$$

Очевидно, что $u[d] = 0$ для $d \in \mathcal{B}$. Рассматриваются множества $D_{\mathcal{B}}^- = \{d \in D \setminus (\mathcal{B} \cup N^*) : u[d] < 0\}$, $D_{\mathcal{B}}^+ = \{d \in N^* : u[d] > 0\}$ и $D_{\mathcal{B}} = D_{\mathcal{B}}^- \cup D_{\mathcal{B}}^+$. Возможны следующие случаи.

2'. Множество $D_{\mathcal{B}} = \emptyset$. Вычислительный процесс закончен. Поток $x(\mathcal{B}, N^*)$ является оптимальным. Минимальное значение целевой функции $f(x(\mathcal{B}, N^*)) = \sum_{d \in \mathcal{B} \cup N^*} c[d] x(\mathcal{B}, N^*)[d]$.

2''. Множество $D_{\mathcal{B}} \neq \emptyset$. Переходим к рекомендации 3.

3. Выбирается некоторая дуга $l = (\bar{v}, \bar{v}) \in D_{\mathcal{B}}$. Рекомендуется выбирать l так, чтобы $|u[l]| = \max\{|u[d]|, d \in D_{\mathcal{B}}\}$ (см. п. 5.1.4). В силу свойства 6° п. 3.4 вершины $\bar{v}$ и $\bar{v}$ соединены единственной цепью в остове $G_{\mathcal{B}}$. Вместе с ребром l она образует цикл графа G . Обозначим этот цикл через Z . Отметим, что здесь утверждается существование такого цикла, но не приводится алгоритм его формирования. Разработку такого алгоритма предлагается осуществить самостоятельно. — Упражнение 6.11.

Возможны два случая.

3'. $l = (\bar{v}, \bar{v}) \in N^*$. Переходим к этапу 2.

3''. $l = (\bar{v}, \bar{v}) \in D \setminus (\mathcal{B} \cup N^*)$. Переходим к этапу 3.

Этап 2. В цикле Z вводим ориентацию, противоположную ориентации дуги $l = (\bar{v}, \bar{v})$. Все дуги цикла, первоначальная ориентация которых совпадает с ориентацией цикла, отмечаются знаком «+», а дуги цикла, первоначальная ориентация которых противоположна ориентации цикла, — знаком «-». Множество дуг, отмеченных знаком «+», обозначим через Z^+ , а множество дуг, отмеченных знаком «-», — через Z^- . Вычисляется

$$\lambda(\mathcal{B}, N^*) = \min \{ \min_{d \in Z^-} x(\mathcal{B}, N^*)[d], h[l], \min_{d \in Z^+} (h[d] - x(\mathcal{B}, N^*)[d]) \}. \quad (4)$$

Как и в алгоритме п. 5.1.3, для невырожденной задачи (1) возможен один и только один из следующих трех случаев.

2а) Найдется единственная дуга $r \in Z^-$, для которой

$$\lambda(\mathcal{B}, N^*) = x(\mathcal{B}, N^*)[r]. \quad (5)$$

Полагаем $\mathcal{B}' = (\mathcal{B} \setminus \{r\}) \cup \{l\}$, $N^{*'} = N^* \setminus \{l\}$ и переходим к этапу 4.

$$2б) \quad \lambda(\mathcal{B}, N^*) = h[l]. \quad (6)$$

Полагаем $\mathcal{B}' = \mathcal{B}$, $N^{*'} = N^* \setminus \{l\}$ и переходим к этапу 4.

2в) Найдется единственная дуга $r \in Z^+$, для которой

$$\lambda(\mathcal{B}, N^*) = h[r] - x(\mathcal{B}, N^*)[r]. \quad (7)$$

Полагаем $\mathcal{B}' = (\mathcal{B} \setminus \{r\}) \cup \{l\}$, $N^{*'} = (N^* \setminus \{l\}) \cup \{r\}$ и переходим к этапу 4.

Этап 3. В цикле Z вводим ориентацию, совпадающую с ориентацией дуги $l = (\bar{v}, \bar{v})$. Все дуги цикла, первоначальная ориентация которых совпадает с ориентацией цикла, отмечаются знаком «+», а дуги цикла, первоначальная ориентация которых противоположна ориентации цикла, — знаком «-». Множество дуг, отмеченных знаком «+», обозначим через Z^+ , а множество дуг, отмеченных знаком «-», — через Z^- . Вычисляется

$$\lambda(\mathcal{B}, N^*) = \min \{ \min_{d \in Z^-} x(\mathcal{B}, N^*)[d], h[l], \min_{d \in Z^+} (h[d] - x(\mathcal{B}, N^*)[d]) \}. \quad (8)$$

Как и в алгоритме п. 5.1.3, для невырожденной задачи возможен один и только один из следующих трех случаев.

3а) Найдется единственная дуга $r \in Z^-$, для которой

$$\lambda(\mathcal{B}, N^*) = x(\mathcal{B}, N^*)[r]. \quad (9)$$

Полагаем $\mathcal{B}' = (\mathcal{B} \setminus \{r\}) \cup \{l\}$, $N^{*'} = N^*$ и переходим к этапу 4.

$$36) \quad \lambda(\mathcal{B}, N^*) = h[l]. \quad (10)$$

Полагаем $\mathcal{B}' = \mathcal{B}$, $N^{*'} = N^* \cup \{l\}$ и переходим к этапу 4.

3в) Найдется единственная дуга $r \in Z^+$, для которой

$$\lambda(\mathcal{B}, N^*) = h[r] - x(\mathcal{B}, N^*)[r]. \quad (11)$$

Полагаем $\mathcal{B}' = (\mathcal{B} \setminus \{r\}) \cup \{l\}$, $N^{*'} = N^* \cup \{r\}$ и переходим к этапу 4.

Этап 4. На этом этапе осуществляется преобразование потока. Поскольку уже сформированы остов $G_{\mathcal{B}'}$ и множество $N^{*'}$, то по ним можно генерировать поток $x(\mathcal{B}', N^{*'})$. Для экономии памяти именно так и поступают при реализации алгоритма на ЭВМ. Однако можно указать и правила вычисления компонент потока $x(\mathcal{B}', N^{*'})$, использующие известные компоненты потока $x(\mathcal{B}, N^*)$.

Если $d \in Z^+$, полагаем

$$x(\mathcal{B}', N^{*'})[d] = x(\mathcal{B}, N^*)[d] + \lambda(\mathcal{B}, N^*); \quad (12)$$

если $d \in Z^-$, полагаем

$$x(\mathcal{B}', N^{*'})[d] = x(\mathcal{B}, N^*)[d] - \lambda(\mathcal{B}, N^*); \quad (13)$$

для $d \in Z$ компоненты потока не изменяются:

$$x(\mathcal{B}', N^{*'})[d] = x(\mathcal{B}, N^*)[d]. \quad (14)$$

Полагаем $\mathcal{B} = \mathcal{B}'$, $N^* = N^{*'}$ и переходим к этапу 1. ■

4.3. Пример. Как уже отмечалось в п. 4.1, для небольшой транспортной сети метод допускает реализацию непосредственно на графе. Ниже рассматривается соответствующий пример.

На рис. 29 приведена сбалансированная транспортная сеть с пятью вершинами и восьмью дугами. Каждая вершина графа изображена квадратом. Номера вершин указаны в верхних клетках, значения $b[\cdot]$ — в левых нижних, свободные клетки оставлены для значений потенциалов. На дугах сети (они не занумерованы) проставлены значения $c[\cdot]$ и $h[\cdot]$ — стоимости транспортировки единицы груза по дуге и пропускные способности дуг.

На рис. 30—32 приводится решение задачи. Начальный остов $G_{\mathcal{B}}$ отвечает множеству дуг $\mathcal{B}_1 = \{(4, 3), (3, 5), (2, 3), (2, 1)\}$, а начальное отмеченное множество $N_1^* = \{(4, 5), (1, 5)\}$. Положительные компоненты потока $x(\mathcal{B}_1, N_1^*)$ обведены кружком с указанием направления перевозки (см. рис. 30). Небазисные компоненты, находящиеся на верхних границах, отмечены звездочкой. Нулевые компоненты не проставлены.

Потенциал первой вершины положен равным нулю (правая нижняя клетка квадрата первой вершины). Для остальных потенциалов использовались дуги остова $G_{\mathcal{B}_1}$ (см. рекомендацию 1 этапа 1): $p^1[2] = -6$, $p^1[3] = 3$, $p^1[4] = -4$, $p^1[5] = 11$.

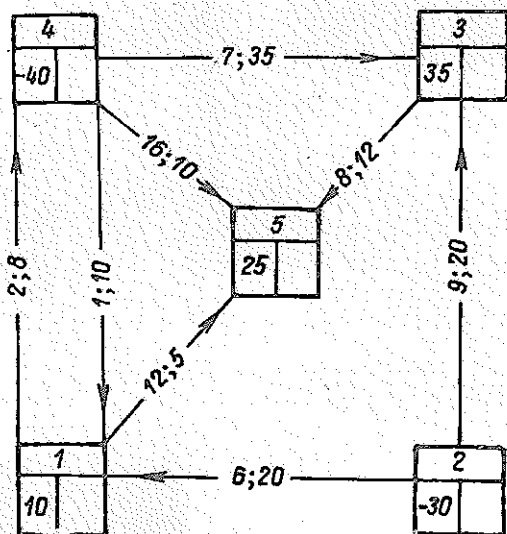


Рис. 29.

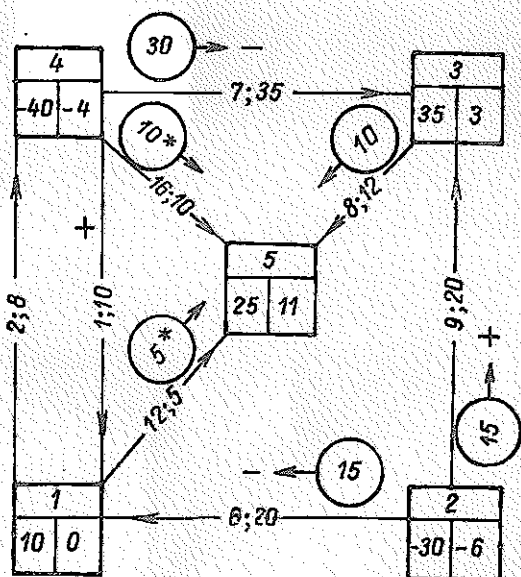
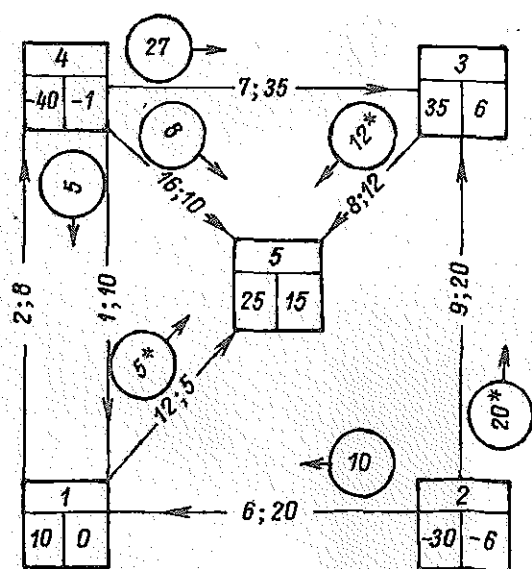
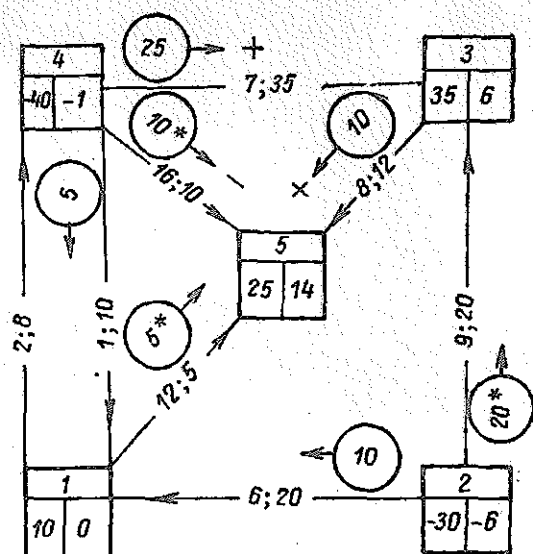


Рис. 30.



Для дуги $(4,1)$ величина $u(4,1) = -3$ (см. (3)). Поток $x(\mathcal{B}_1, N_1^*)$ не является оптимальным. На рис. 30 знаки «+» и «-» на дугах цикла Z_1 проставлены в соответствии с рекомендацией этапа 3. Величина $\lambda(\mathcal{B}_1, N_1^*) = 5$, множество $N_2^* = \{(1, 5), (4, 5), (2, 3)\}$, $\mathcal{B}_2 = \{(4, 1), (2, 1), (4, 3), (3, 5)\}$. Поток $x(\mathcal{B}_2, N_2^*)$ указан на рис. 31. Он также не является оптимальным, так как для дуги $(4, 5)$ величина $u(4, 5) = 1$. На дугах цикла Z_2 (см. рис. 31) проставлены знаки «+» и «-». Величина $\lambda(\mathcal{B}_2, N_2^*) = 2$, $\mathcal{B}_3 = \{(4, 1), (2, 1), (4, 3), (4, 5)\}$, $N_3^* = \{(1, 5), (3, 5), (2, 3)\}$ (см. рекомендации этапа 2). Поток $x(\mathcal{B}_3, N_3^*)$ указан на рис. 32. Он является оптимальным, поскольку множество $D_{\mathcal{B}_3} = \emptyset$; $f(x(\mathcal{B}_3, N_3^*)) = 718$.

4.4. Обоснование алгоритма п. 4.2 сводится к установлению идентичности его с алгоритмом п. 5.1.3. Для этого достаточно последовательно сравнить алгоритмы и выяснить соответствие их рекомендаций. Особое внимание обращаем на следующие обстоятельства:

вычисление массива потенциалов и величин $u[\cdot]$ (см. (3)) эквивалентно вычислению компонент вектора $u(\mathcal{B})$ (см. п. 4.3.1) и строки $a_0(\mathcal{B}) = d_0(\mathcal{B})\bar{A}$, где $d_0(\mathcal{B}) = (-1, u(\mathcal{B}))$;

формирование цикла Z и разметка его дуг равнозначны разложению вектора a^l (он соответствует дуге l) по базису $\mathcal{B}$. Коэффициенты этого разложения равны 1, -1 или 0 (см. п. 8.1.5), так что формулы (4) — (11) являются частными случаями формул (3), (4), (9), (11) — (13), (16), (18) п. 5.1.3;

преобразование потока по формулам (12) — (14) равносильно пересчету компонент π -базисного плана в алгоритме п. 5.1.3.

Детальное обоснование алгоритма п. 4.2 предоставляем читателю. — Упражнение 6.12.

4.5. Нахождение допустимого потока. Для формирования остова $G_{\mathcal{B}}$ и множества N^* , которым отвечает допустимый поток $x(\mathcal{B}, N^*)$ транспортной сети, можно использовать следующий подход, являющийся аналогом метода минимизации невязок п. 5.1.7.

Рассматривается вспомогательная транспортная сеть на графе $G' = (V', D', \Gamma')$. Он получается из графа G добавлением одной новой вершины v_0 , дуг, идущих из вершины v_0 во все вершины $v \in V$, для которых $b[v] > 0$, и дуг, идущих в вершину v_0 из всех вершин $v \in V$, для которых $b[v] < 0$. Определим потребности $b'[v]$ в вершинах графа G' формулами $b'[v] = b[v]$, $v \in V$, $b'[v_0] = 0$, стоимости перевозок $c'[d]$ положим равными нулю для $d \in D$ и единице для $d \in D' \setminus D$. Пропускные способности дуг графа G' определим формулами $h'[d] = h[d]$, $d \in D$, $h'[d] = +\infty$, $d \in D' \setminus D$.

Частичный граф $G'_{D' \setminus D} = (V', D' \setminus D, \Gamma'_{D' \setminus D})$ очевидно является остовом построенной таким образом транспортной сети,

и остову $G_{D' \setminus D}^*$ и множеству $N_1^* = \emptyset$ отвечает допустимый поток $x(D' \setminus D, N_1^*)$. Если вспомогательная задача невырождена, то метод п. 4.2 позволяет найти оптимальный поток $x'(\mathcal{D}, N^*)$ транспортной сети на графе G' . Возможны два случая.

а) Множество дуг остова $G_{\mathcal{D}}^*$ содержится в D . Тогда полагаем $G_{\mathcal{D}} = (V, \mathcal{D}, \Gamma_{\mathcal{D}})$, $N^* = \bar{N}^*$.

б) Множество дуг остова $G_{\mathcal{D}}^*$ не содержится в D . Тогда если вспомогательная задача невырождена, то $\mathcal{D} = \emptyset$. Вырожденность вспомогательной задачи не вызывает принципиальных осложнений (см. п. 5.1.7).

4.6. О транспортных задачах с нарушенным балансом. На несбалансированной транспортной сети (см. п. 3.5) можно определить следующие экстремальные задачи.

1. В случае $\sum_{v \in V} b[v] < 0$ рассматривается задача

$$(\mathcal{D}, f) : f(x) = \sum_{d \in D} c[d] x[d] \rightarrow \min, \\ \mathcal{D} = \{x = \{x[d], d \in D\} : \sum_{d \in D_v^+} x[d] - \sum_{d \in D_v^-} x[d] = b[v], \\ v \in \{u \in V : b[u] \geq 0\}; \\ \sum_{d \in D_v^+} x[d] - \sum_{d \in D_v^-} x[d] \geq b[v], v \in \{u \in V : b[u] < 0\}; \quad (15) \\ 0 \leq x[d] \leq h[d], d \in D\}.$$

2. В случае $\sum_{v \in V} b[v] > 0$ рассматривается задача

$$(\mathcal{D}, f) : f(x) = \sum_{d \in D} c[d] x[d] \rightarrow \min, \mathcal{D} = \{x = \{x[d], d \in D\} : \\ 0 \leq \sum_{d \in D_v^+} x[d] - \sum_{d \in D_v^-} x[d] \leq b[v], v \in \{u \in V : b[u] \geq 0\}; \\ \sum_{d \in D_v^+} x[d] - \sum_{d \in D_v^-} x[d] = b[v], v \in \{u \in V : b[u] < 0\}; \quad (16) \\ 0 \leq x[d] \leq h[d], d \in D\}.$$

Задачи (15) и (16) называются *транспортными задачами с нарушенным балансом*. Вместе с транспортными задачами они относятся к классу задач транспортного типа.

Метод п. 4.2 формально позволяет находить решение задач (15), (16). В случае задачи (15) вводят фиктивный пункт v_0 с потребностью $b[v_0] = -\sum_{v \in V} b[v]$ и соединяют его дугами со всеми пунктами транспортной сети так, чтобы дуги входили в пункт v_0 . Для новых дуг ограничения на пропускные способности полагают равными ∞ , а стоимости перевозок — нулю.

В задаче (16) поступают аналогичным образом, но дуги выходят из пункта v_0 , а стоимости перевозок по новым дугам должны быть достаточно большими числами.

§ 5. Транспортная задача в матричной постановке. Венгерский метод

5.1. Предварительные замечания. В этом параграфе рассматривается вычислительный метод решения транспортной задачи в матричной постановке (см. п. 2.3.2)

$$(\mathcal{D}, f): f(x) = \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j] \rightarrow \min,$$

$$\mathcal{D} = \{x \in R^{mn} : \sum_{j=1}^n x[i, j] = a[i], i \in 1:m;$$

$$\sum_{i=1}^m x[i, j] = b[j], j \in 1:n; x \geq 0\}.$$

Предполагается, что $\sum_{i=1}^m a[i] = \sum_{j=1}^n b[j]$. Кроме того, будем считать объемы производства $a[i], i \in 1:m$, и потребления $b[j], j \in 1:n$, целыми числами, что вполне оправданно в практических ситуациях.

Метод прост с вычислительной точки зрения и не боится вырожденности, что делает его особенно удобным для решения задачи о назначениях (см. пп. 2.3.8 и 8.1.5, 2°). В процессе решения осуществляется последовательный переход от некоторого недопустимого плана (не все потребности удовлетворены и, естественно, не весь продукт вывезен из пунктов производства) к допустимому, являющемуся решением задачи. Этот переход осуществляется за конечное число итераций q , не известное до окончания вычислений.

Идея метода предложена венгерским математиком Е. Эгервари еще в 1931 г., т. е. до построения общей теории линейного программирования. В настоящее время она широко используется при формировании специальных вычислительных процедур.

Мы не останавливаемся на связи венгерского метода с общими методами, описанными в гл. 4. Отметим только, что он наиболее близок к двойственному симплекс-методу.

5.2. Алгоритм. В процессе вычислений будут заполнены табл. $T_1, T_2, \dots, T_q$ вида

$$\begin{array}{|c|c|c|}
 \hline
 \dots & \dots & \dots \\
 \hline
 a^k[i] & \begin{array}{|c|c|}
 \hline
 c^k[i, j] \\
 x^k[i, j] \\
 \hline
 \end{array} & \dots \\
 \hline
 \dots & \dots & \dots \\
 \hline
 \end{array} \quad (1)$$

$b^k[j]$

Каждая из этих таблиц содержит m строк и n столбцов. Слева у каждой строки проставляются числа $a^k[i], i \in 1:m$, снизу у каждого столбца — числа $b^k[j], j \in 1:n$. В правом верхнем углу клетки таблицы, стоящей на пересечении i -й строки и j -го

столбца, проставляется элемент $c^k[i, j]$ матрицы $C^k = (c^k[i, j])$, в левом нижнем углу — элемент $x^k[i, j]$ плана $x^k = (x^k[i, j])$. Матрицы C^k , x^k и векторы b^k , a^k определяются в процессе вычислений. Для удобства в таблицах, а также слева от них и под ними будут записываться только положительные компоненты матриц x^k и векторов a^k , b^k .

Этап 0. Заполнение табл. T_1 . 1. В каждой строке матрицы C определяется минимальный элемент. Он вычитается из всех элементов этой строки. В каждом столбце преобразованной матрицы определяется минимальный элемент. Он вычитается из всех элементов столбца. Полученная после таких преобразований матрица обозначается через $C^1 = (c^1[i, j])$. В табл. T_1 проставляются элементы матрицы C^1 .

2. Положительные компоненты плана x^1 определяются следующим образом. Полагаем $a^1[i] = a[i]$, $i \in 1:m$, и $b^1[j] = b[j]$, $j \in 1:n$. Некоторым способом (например, сначала по строкам, потом по столбцам) нумеруются нулевые элементы матрицы C^1 и последовательно рассматриваются клетки табл. T_1 , в которых размещены эти элементы. Пусть $[s, t]$ — первая, еще не рассмотренная клетка. Если $\min\{a^1[s], b^1[t]\} = 0$, то считаем клетку рассмотренной и переходим к следующей; если $\min\{a^1[s], b^1[t]\} > 0$, то полагаем $x^1[s, t] = \min\{a^1[s], b^1[t]\}$, $a^1[s] = a^1[s] - x^1[s, t]$, $b^1[t] = b^1[t] - x^1[s, t]$, подчеркиваем нуль в этой клетке и, считая ее рассмотренной, переходим к следующей, еще не рассмотренной клетке. Построение заканчивается, если рассмотрены все клетки с $c^1[i, j] = 0$. Вычисленные положительные компоненты плана x^1 записываются в соответствующих клетках табл. T_1 . Для каждого $i \in 1:m$, $j \in 1:n$ ненулевые элементы $a^1[i]$, $b^1[j]$ — неиспользованные ресурсы и неудовлетворенные потребности — записываются соответственно слева и снизу от строки с номером i и столбца с номером j .

3. Полагаем номер итерации $k = 1$. Строку с номером i , для которой $a^k[i] > 0$, назовем *богатой*, столбец с номером j , для которого $b^k[j] > 0$, — *бедным*. Остальные строки и столбцы табл. T_k назовем соответственно *небогатыми* и *небедными*.

Этап 1. Разметка таблицы. 1. Если в табл. T_k нет богатых строк, то вычислительный процесс закончен. План x^k — оптимальный, его положительные компоненты содержатся в табл. T_k . В противном случае переходим к рекомендации 2.

2. Над всеми небедными столбцами ставим знак «+». Столбцы, над которыми стоит знак «+», будем считать отмеченными. Соответственно элемент матрицы C^k считается отмеченным, если он находится в отмеченном столбце или отмеченной строке (об отметках строк будет сказано ниже), и неотмеченным в противном случае.

3. В матрице C^k ищем первый по порядку (сначала среди элементов первой строки слева направо, затем среди элементов второй строки в том же порядке и т. д.) неотмеченный нуль.

Если его нет, переходим к этапу 3, в противном случае — к рекомендации 4.

4. У найденного нуля ставится штрих, а справа, у строки, где он расположен, — знак «+», после чего строка и все ее элементы считаются отмеченными. Возможны два случая.

4'. Нуль со штрихом расположен в бедной строке. В этой строке ищутся подчеркнутые нули, стоящие в отмеченных столбцах. Если их нет, то переходим к рекомендации 3. Если такие нули имеются, то к каждому из них приписываются два штриха и снимается отметка со столбца, в котором расположен такой нуль. Переходим к рекомендации 3.

4''. Нуль со штрихом расположен в богатой строке. Переходим к этапу 2.

Этап 2. Построение цепочки и преобразование плана. 1. Строится цепочка чередующихся $0'$ и $0''$, начинающаяся с $0'$, стоящего в богатой строке, и оканчивающаяся $0'$, стоящим в бедном столбце. Это делается так. Обозначим $0'$, стоящий в богатой строке, через $0'_1$. Если $0'_1$ расположен в бедном столбце, то цепочка построена: она состоит из одного элемента. В противном случае в столбце, содержащем $0'_1$, найдется (единственная) клетка, содержащая подчеркнутый $0''$. Обозначим его через $0''_1$ и в строке, в которой он расположен, найдем (единственную) клетку, содержащую $0'$. Обозначим его через $0'_2$. Процесс продолжается до тех пор, пока в конце цепочки не будет найден $0'$, стоящий в бедном столбце.

2. Пусть $[i_1, j_1], [i_2, j_1], [i_2, j_2], \dots, [i_r, j_r]$ — последовательность клеток таблицы, отвечающая цепочке $0'_1, 0''_1, 0'_2, \dots, 0'_r$. Вычисляем

$$e_k = \min \{a^k[i_1]; b^k[j_r]; x^k[i_{p+1}, j_p], p \in 1:r-1\}. \quad (2)$$

Заполняем табл. T_{k+1} . Компоненты плана x^{k+1} отличаются от компонент плана x^k только в клетках цепочки:

$$x^{k+1}[i_p, j_p] = x^k[i_p, j_p] + e_k, \quad p \in 1:r, \quad (3)$$

$$x^{k+1}[i_{p+1}, j_p] = x^k[i_{p+1}, j_p] - e_k, \quad p \in 1:r-1. \quad (4)$$

Матрица $C^{k+1} = C^k$; $a^{k+1}[i_1] = a^k[i_1] - e_k$; $b^{k+1}[j_r] = b^k[j_r] - e_k$; $a^{k+1}[i] = a^k[i]$, $i \in 1:m$, $i \neq i_1$, $b^{k+1}[j] = b^k[j]$, $j \in 1:n$, $j \neq j_r$. В табл. T_{k+1} подчеркиваем нулевые элементы матрицы C^{k+1} , находящиеся в клетках с положительными перевозками плана x^{k+1} . Полагаем $k = k + 1$. Переходим к этапу 1.

Этап 3. Преобразование матрицы C^k . 1. Среди неотмеченных элементов матрицы C^k определяется минимальный.

2. Заполняется табл. T_{k+1} . Элементы матрицы C^{k+1} вычисляются так: из неотмеченных строк матрицы C^k вычитается и к отмеченным столбцам прибавляется найденный минимальный элемент. Компоненты плана x^k и векторов a^k, b^k переписываются в табл. T_{k+1} без изменений: $x^{k+1} = x^k$, $a^{k+1} = a^k$,

$b^{k+1} = b^k$. Полагаем $k = k + 1$. Переходим к рекомендации 2 этапа 1. ■

5.3. Пример. В табл. (6)–(11) приведено решение венгерским методом транспортной задачи с матрицей стоимостей, объемами производства и потребления, указанными в следующей таблице:

12	7	1	2	10
13	5	3	7	3
12	4	8	1	6
12	11	2	2	9
	24	11	7	7

(5)

При переходе от табл. (6) к табл. (7) производится преобразование матрицы C^1 (процедура этапа 3), в результате которого появляется нуль в богатой строке и бедном столбце. Поэтому между третьим пунктом производства и первым пунктом потребления в табл. (7) поставлена перевозка, равная 5 единицам, и указаны новые значения $a[i]$ и $b[j]$. При снятии отметки знак «+» обводится кружком.

		+	+		
1	4	0	1	9	
	11				
	0'	0	4	0	+
	13				
5	1	7	0	5	
			7		
12	7	0	0	7	
	11		7		

(6)

		+	⊕		
1	3	0	1	8	
	11				
	0'	1	5	0	+
	13				
	0'	7	0''	4	+
5			7		+
12	6	0	0'	6	+
	6		7		

(7)

		⊕	+	⊕	
1	3	0	1	8	
	11				
	0''	1	5	0'	+
	13				
	0'	7	0''	4	+
11			1		+
6	6	0	0'	6	+
			6		
			7		

(8)

		⊕	+	+	
1	3	0	1	8	
	11				
	0''	1	5	0'	+
12			1		+
	0'	7	0	4	+
12					
5	6	0	0	6	
			7		
			6		

(9)

	$\oplus$	+	+		
1	0'	0	1	5	
	11				
	0''	4	8	0'	
12				1	
	0	10	3	4	
	12				
5	3	0	0	3	
		7			
					6

(10)

	$\oplus$	$\oplus$	+		
	0'	0''	1	5	
1	11				
	0''	4	8	0'	
11				2	
	0	10	3	4	
12					
5	3	0'	0	3	
		7			
					5

(11)

Цепочки для удобства обозначены ломаной пунктирной линией. Направление поиска нулей при конструировании цепочки везде одинаковое: от нуля со штрихом в последней, отмеченной знаком «+» строке. Оптимальный план записан в табл. (12)

	7	1	2	10	
6	6				
	5	3	7	3	
6				7	
	4	8	1	6	
12					
	11	2	2	9	
	5	7			

(12)

вместе с фактической матрицей стоимостей (а не матрицей C^7) рассматриваемой задачи. Минимальное значение целевой функции $f(x^7) = 171$.

5.4. Обоснование алгоритма. Пусть $d_1, d_2, \dots, d_m, h_1, h_2, \dots, h_n$ — произвольные числа, $c'[i, j] = c[i, j] - d_i - h_j$ для $i \in 1:m, j \in 1:n$. Рассмотрим транспортную задачу

$$(\mathcal{D}, f') : f'(x) = \sum_{i=1}^m \sum_{j=1}^n c'[i, j] x[i, j] \rightarrow \min, \quad x \in \mathcal{D}.$$

Обоснование алгоритма венгерского метода сводится к доказательству следующих утверждений.

1°. Если план $\bar{x}$ — решение задачи $(\mathcal{D}, f')$, то он является и решением задачи $(\mathcal{D}, f)$.

2°. При любом $k \in 1:q$ план $x^k \geq 0$.

3°. При любом $k \in 1:q$ положительным перевозкам плана x^k отвечают нулевые элементы матрицы C^k .

4°. Если в табл. T^k нет богатых строк, то x^k — решение задачи $(\mathcal{D}, f)$.

5°. Правило построения цепочки, описанное в рекомендации 1 этапа 2, позволяет (однозначно) строить цепочку чередующихся $0'$ и $0''$, начинающуюся $0'$, стоящим в богатой строке, и оканчивающуюся $0'$, стоящим в бедном столбце.

6°. Число итераций q конечно.

5.5. Доказательство утверждений 1°—6°.

1°. Действительно, для произвольного $x \in \mathcal{D}$ имеем

$$\begin{aligned} f'(x) &= \sum_{i=1}^m \sum_{j=1}^n (c[i, j] - d_i - h_j) x[i, j] = \\ &= \sum_{i=1}^m \sum_{j=1}^n c[i, j] x[i, j] - \sum_{i=1}^m d_i \sum_{j=1}^n x[i, j] - \\ &- \sum_{j=1}^n h_j \sum_{i=1}^m x[i, j] = f(x) - \sum_{i=1}^m d_i a[i] - \sum_{j=1}^n h_j b[j] = \\ &= f(x) + d, \end{aligned} \quad (13)$$

где $d = -\sum_{i=1}^m d_i a[i] - \sum_{j=1}^n h_j b[j]$, так что целевые функции в задачах $(\mathcal{D}, f)$ и $(\mathcal{D}, f')$ различаются только постоянным слагаемым. Поэтому каждый оптимальный план задачи $(\mathcal{D}, f)$ является и решением задачи $(\mathcal{D}, f')$. ■

2°. Утверждение следует из способа построения плана x^1 и формул (2)—(4), по которым осуществляется преобразование плана. ■

3°. При построении плана x^1 положительные перевозки размещаются в клетках с нулевыми элементами матрицы C^1 . Преобразование плана, осуществляемое на этапе 2, проводится вдоль цепочки нулевых элементов матрицы C^k . Остается доказать, что при преобразовании матрицы C^k в матрицу C^{k+1} элементы матрицы C^k , отвечающие положительным перевозкам, не изменяются. Для этого достаточно отметить, что при переходе к этапу 3 каждый подчеркнутый нуль матрицы C^k находится либо в отмеченной строке и неотмеченном столбце, либо в неотмеченном столбце и неотмеченной строке. Но преобразование этапа 3 такие элементы матрицы C^k не изменяет. ■

4°. Из условия утверждения 4° следует, что $x^k \in \mathcal{D}$, а из 3°—его оптимальность в транспортной задаче с множеством допустимых планов $\mathcal{D}$ и матрицей стоимостей C^k . Тогда утверждение следует из 1°. ■

5°. Из правила разметки матрицы следует, что:

1) либо столбец, содержащий $0'$, бедный, либо в этом столбце содержится единственный $0''$ (выделение $0''$ было связано со снятием отметки, после чего в этом столбце не мог появиться второй $0''$);

2) в каждой строке, содержащей $0''$, имеется единственный $0'$, так как $0''$ появляется только после выбора $0'$ в этой строке; строка становится отмеченной, и появление в ней второго $0'$ невозможно.

Следовательно, среди трех помеченных нулей цепочки $0'$, $0''$, $0'$ первый $0'$ был помечен после второго $0'$, так что все $0'$ в це-

почке различны, и построение цепочки через конечное число шагов закончится $0'$, стоящим в бедном столбце. ■

6°. Заметим, что каждая реализация этапа 2 уменьшает $\sum_{i=1}^m a^k[i]$ на целое положительное число, и после конечного числа таких реализаций эта сумма окажется равной нулю. Поэтому достаточно показать, что после некоторого числа переходов к этапу 3 осуществится переход к этапу 2. Действительно, очередной переход к этапу 3 приводит к увеличению числа отмеченных строк при разметке, и, следовательно, неизбежно появление $0'$ в богатой строке после конечного числа таких переходов. ■

УПРАЖНЕНИЯ

- 6.1. См. с. 214. 6.2. См. с. 217. 6.3. См. с. 217. 6.4. См. с. 224.
6.5. См. с. 224. 6.6. См. с. 232. 6.7. См. с. 232. 6.8. См. с. 233.
6.9. См. с. 238. 6.10. См. с. 238. 6.11. См. с. 240. 6.12. См. с. 245.

ПАРАМЕТРИЧЕСКИЕ ЗАДАЧИ

§ 1. Параметрические задачи
линейного программирования

1.1. Параметрические экстремальные задачи. Пусть задано некоторое множество Θ (предполагаем, что оно является подмножеством пространства R^n при некотором n) и для каждого $\theta \in \Theta$ определена экстремальная задача максимизации (минимизации) $(\mathcal{D}_\theta, f_\theta)$, где $\mathcal{D}_\theta \subset R^n$, а f_θ задана на $\mathcal{D}_\theta$ или R^n . Семейство задач $\{(\mathcal{D}_\theta, f_\theta), \theta \in \Theta\}$ назовем *параметрическим семейством экстремальных задач*, или *параметрической экстремальной задачей*.

Обозначим через $\bar{\Theta}$ множество тех θ из Θ , для которых экстремальная задача $(\mathcal{D}_\theta, f_\theta)$ имеет решение. Тогда *решением параметрической экстремальной задачи* назовем такое семейство векторов $\{\bar{x}_\theta, \theta \in \bar{\Theta}\}$, что для каждого $\theta \in \bar{\Theta}$ план $\bar{x}_\theta$ является решением задачи $(\mathcal{D}_\theta, f_\theta)$.

Решить параметрическую экстремальную задачу — значит найти множество $\bar{\Theta} \subset \Theta$ и семейство векторов $\{\bar{x}_\theta, \theta \in \bar{\Theta}\}$ с указанными выше свойствами. Параметрическая экстремальная задача не имеет решения, если множество $\bar{\Theta} = \emptyset$.

1.2. Параметрические задачи линейного программирования. *Параметрической задачей линейного программирования* называется такое семейство экстремальных задач $\{(\mathcal{D}_\theta, f_\theta), \theta \in \Theta\}$, что при каждом фиксированном $\theta \in \Theta$

$$(\mathcal{D}_\theta, f_\theta) : f_\theta(x) = c_\theta x \rightarrow \max, \quad \mathcal{D}_\theta = \{x \in R^n : A_\theta x = b_\theta, x \geq 0\} \quad (1)$$

является канонической задачей линейного программирования.

В общем случае предполагается функциональная зависимость от θ каждого элемента матрицы задачи, компонент вектора ограничений и коэффициентов целевой функции. Исследование таких задач вызывает значительные трудности, методы

их решения не разработаны. Удовлетворительные результаты получены только для некоторых частных случаев задачи (1).

Например, практическую ценность представляют следующие две параметрические задачи линейного программирования:

$$\{(\mathcal{D}, f_\theta), \theta \in \Theta \subset R\} : f_\theta(x) = c_\theta x = (c' + \theta c'')x \rightarrow \max, \\ \mathcal{D} = \{x \in R^n : Ax = b, x \geq 0\}; \quad (2)$$

$$\{(\mathcal{D}_\theta, f), \theta \in \Theta \subset R\} : f(x) = cx \rightarrow \max, \\ \mathcal{D}_\theta = \{x \in R^n : Ax = b_\theta = b' + \theta b'', x \geq 0\}, \quad (3)$$

где $A — m \times n$ матрица; векторы c', c'' и c из R^n ; b', b'' и $b —$ из R^m .

Задачу (2) будем называть *параметрической задачей линейного программирования с параметром в целевой функции*, а задачу (3) — *параметрической задачей линейного программирования с параметром в векторе ограничений*.

Можно привести следующую экономическую интерпретацию задач (2) и (3). Наличие параметра θ в целевой функции задачи (2) обычно связывают с сезонными колебаниями цен на сельскохозяйственную продукцию, а наличие θ в векторе ограничений — с колебаниями объема ресурсов (уровня запасов или поставок сырья) на предприятии или в некотором промышленном (аграрно-промышленном) комплексе. Предположение линейной (аффинной) зависимости цен и ресурсов от времени θ на некотором коротком временном интервале Θ выглядит естественным.

Возможность эффективного решения этих задач основана на следующих утверждениях.

1°. В задаче (2), если $\mathcal{D} \neq \emptyset$, можно указать такое разбиение прямой R на конечное множество интервалов

$$\Theta_1 = [\theta_1^-, \theta_1^+], \quad \Theta_2^+ = (\theta_1^+, \theta_2^+], \quad \Theta_3^+ = (\theta_2^+, \theta_3^+], \quad \dots, \quad \Theta_{\bar{q}}^+ = \\ = (\theta_{\bar{q}-1}^+ + \infty), \\ \Theta_2^- = [\theta_2^-, \theta_1^-), \quad \Theta_3^- = [\theta_3^-, \theta_2^-), \quad \dots, \quad \Theta_{\bar{q}}^- = (-\infty, \theta_{\bar{q}-1}^-) \quad (4)$$

(возможен случай совпадения одного из неограниченных интервалов с R) и такие последовательности базисов (см. п. 4.3.1)

$$\mathcal{B}_1, \mathcal{B}_2^+, \mathcal{B}_3^+, \dots, \mathcal{B}_{\bar{q}}^+; \mathcal{B}_2^-, \mathcal{B}_3^-, \dots, \mathcal{B}_{\bar{q}}^-, \quad (5)$$

что для любого из названных выше интервалов, кроме, быть может, $\Theta_{\bar{q}}^+, \Theta_{\bar{q}}^-$, при любом $\theta \in \Theta_1$ ($\theta \in \Theta_k^+, k \in 2:\bar{q}, \theta \in \Theta_k^-, k \in 2:\bar{q}$) план $x(\mathcal{B}_1)$ ($x(\mathcal{B}_k^+)$, $k \in 2:\bar{q}$, $x(\mathcal{B}_k^-)$, $k \in 2:\bar{q}$) является решением задачи $(\mathcal{D}, f_\theta)$. В промежутках $\Theta_{\bar{q}}^+, \Theta_{\bar{q}}^-$ целевая функция f_θ задачи $(\mathcal{D}, f_\theta)$ может быть не ограничена.

2°. В задаче (3) можно указать такое разбиение прямой R на конечное множество интервалов

$$\begin{aligned}\Theta_1 &= [\theta_1^-, \theta_1^+], \Theta_2^+ = (\theta_1^+, \theta_2^+], \Theta_2^- = (\theta_2^-, \theta_3^-], \dots, \Theta_q^+ = \\ &= (\theta_{q-1}^+, +\infty), \\ \Theta_2^- &= [\theta_2^-, \theta_1^-), \Theta_3^- = [\theta_3^-, \theta_2^-), \dots, \Theta_q^- = (-\infty, \theta_{q-1}^-) \quad (6)\end{aligned}$$

(возможен случай совпадения одного из неограниченных интервалов с R) и такие последовательности базисов

$$\mathcal{B}_1, \mathcal{B}_2^+, \mathcal{B}_3^+, \dots, \mathcal{B}_q^+; \mathcal{B}_2^-, \mathcal{B}_3^-, \dots, \mathcal{B}_q^-, \quad (7)$$

что для любого из названных выше интервалов, кроме, быть может, Θ_q^+ , Θ_q^- , при любом $\theta \in \Theta_1$ ($\theta \in \Theta_k^+$, $k \in 2:\bar{q}$, $\theta \in \Theta_k^-$, $k \in 2:\bar{q}$) план $x_\theta(\mathcal{B}_1) = B_1^{-1}b' + \theta B_1^{-1}b''$ ($x_\theta(\mathcal{B}_k^+)$, $k \in 2:\bar{q}$, $x_\theta(\mathcal{B}_k^-)$, $k \in 2:\bar{q}$) является решением задачи $(\mathcal{D}_\theta, f)$. В промежутках Θ_q^+ и Θ_q^- задача $(\mathcal{D}_\theta, f)$ может не иметь решения.

В силу утверждения 1° решение задачи (2) может быть задано конечным набором векторов из $\mathcal{D}$, каждый из которых является решением задачи $(\mathcal{D}, f_\theta)$ для каждого θ из пересечения множества Θ с соответствующим интервалом последовательности (4), кроме, быть может, θ , принадлежащих неограниченным интервалам Θ_q^+ и Θ_q^- .

Описание решения задачи (3) несколько иное. Здесь для каждого из интервалов (6) указывается аффинная функция, отображающая этот интервал в R^n и позволяющая для каждого θ из Θ найти решения задачи $(\mathcal{D}_\theta, f)$ (или устанавливается, что при каждом θ из Θ_q^+ или Θ_q^- задача не имеет решения).

Доказательство утверждения 1°. (Читатель должен быть знаком с содержанием пп. 4.3.1 и 4.3.3, понятия и обозначения которых здесь используются.) Прежде всего отметим, что для базиса $\mathcal{B}$ задачи (1), $N(\mathcal{B}) = \{j_1, \dots, j_m\}$ вектор $c(\mathcal{B}, \theta) = c'(\mathcal{B}) + \theta c''(\mathcal{B})$, где $c'(\mathcal{B}) = (c'[j_1], \dots, c'[j_m])$, $c''(\mathcal{B}) = (c''[j_1], \dots, c''[j_m])$, и $d_0(\mathcal{B}, \theta) = (-1, c(\mathcal{B}, \theta)B^{-1})$. Тогда

$$\begin{aligned}a_0(\mathcal{B}, \theta) &= d_0(\mathcal{B}, \theta) \bar{A} = -1 \cdot (c' + \theta c'') + [c'(\mathcal{B}) + \theta c''(\mathcal{B})] B^{-1} A = \\ &= [-c' + c'(\mathcal{B}) A(\mathcal{B})] + \theta [-c'' + c''(\mathcal{B}) A(\mathcal{B})] = a'_0(\mathcal{B}) + \theta a''_0(\mathcal{B}).\end{aligned}$$

Покажем, что если для некоторого $\theta_0 \in R$ целевая функция f_{θ_0} не ограничена сверху на $\mathcal{D}$, то функция f_θ также не ограничена сверху на $\mathcal{D}$ для каждого θ из некоторого бесконечного содержащего θ_0 интервала (неограниченного справа, слева или совпадающего со всей прямой R). Действительно, неограниченность f_{θ_0} можно обнаружить, например, на некоторой итерации

симплекс-метода (см. случай 3' этапа 1): для некоторого базиса $\mathcal{B}$, $N(\mathcal{B}) = \{j_1, \dots, j_m\}$ найдется такое $l \in I^-(\mathcal{B}, \theta_0) = \{j \in 1:n : a(\mathcal{B}, \theta_0)[0, j] = a'(\mathcal{B})[0, j] + \theta_0 a''(\mathcal{B})[0, j] < 0\}$, что $a'(\mathcal{B}) = B^{-1}a' \leq 0$. Следовательно (см. п. 4.3.6, 6°), функция f_θ не ограничена сверху для всех θ из множества $\Theta = \{\theta \in R : a'(\mathcal{B})[0, l] + \theta a''(\mathcal{B})[0, l] < 0\}$. Так как $\theta_0 \in \Theta$, то $\Theta \neq \emptyset$. Поэтому очевидно Θ — бесконечный интервал, содержащий θ_0 .

Предположим теперь, что при некотором $\theta_0 \in R$ задача $(\mathcal{D}, f_{\theta_0})$ имеет решение. Тогда (см. замечание в п. 4.3.7) найдется базисное решение $x(\mathcal{B})$, и по критерию оптимальности (см. конец п. 3.1.9 и доказательство утверждения 4° п. 4.3.7) вектор $u(\mathcal{B})$ является решением системы неравенств $uA \geq c' + \theta_0 c''$. Поэтому система числовых неравенств $u(\mathcal{B})A \geq c' + \theta c''$ очевидно выполняется для всех θ из некоторого интервала Θ_0 , содержащего θ_0 . Следовательно, для любого $\theta \in \Theta_0$ базис $\mathcal{B}$ будет оптимальным базисом задачи $(\mathcal{D}, f_\theta)$.

Поскольку из столбцов матрицы A можно образовать лишь конечное число базисов, прямая R может быть покрыта конечным числом интервалов. В каждом из них либо один и тот же вектор является решением всех задач $(\mathcal{D}, f_\theta)$ (θ из этого интервала), либо ни при каком θ из данного интервала задача $(\mathcal{D}, f_\theta)$ не имеет решения. Из рассматриваемой системы интервалов нетрудно выбрать конечную систему непересекающихся интервалов, обладающих свойствами, указанными в утверждении. ■

Замечание. При каждом $\theta \in \Theta$ задача, двойственная к задаче (3), содержит параметр θ в целевой функции: $f^*(u) = u(b' + \theta b'')$, поэтому задача, двойственная к задаче (3), очевидно эквивалентна параметрической задаче линейного программирования с параметром в целевой функции и в силу теории двойственности (см. пп. 3.5 и 3.7) утверждение 2° следует из утверждения 1°. Детальное доказательство этого предоставим читателю. — Упражнение 7.1.

§ 2. Алгоритм решения параметрической задачи линейного программирования с параметром в целевой функции

2.1. Предварительные замечания. Рассмотрим параметрическую задачу линейного программирования с параметром в целевой функции

$$\{(\mathcal{D}, f_\theta), \theta \in \Theta\} : f_\theta(x) = c_\theta x = (c' + \theta c'')x \rightarrow \max, \\ \mathcal{D} = \{x \in R^n : Ax = b, x \geq 0\}, \quad (1)$$

где $A = (a[i, j])$ — $m \times n$ матрица, ранг которой $r(A) = m$; $c' = (c'[1], \dots, c'[n])$, $c'' = (c''[1], \dots, c''[n])$, $b \in R^m$, $\Theta \subset R$.

В алгоритме п. 2.2 применяется вычислительная схема симплекс-метода п. 4.3.3 (возможно применение и вычислительной схемы п. 4.4.3). В дальнейшем используются обозначения и понятия п. 4.3.1. Новым является только следующее понятие: базис $\mathcal{B}$ называется θ -оптимальным, если план $x(\mathcal{B})$ есть решение задачи $(\mathcal{D}, f_\theta)$.

При описании алгоритма будем считать $\Theta = R$. Переход к общему случаю не вызывает затруднений и уже прокомментирован в п. 1.2 при доказательстве утверждения 1°.

Использование процедуры симплекс-метода предполагает отсутствие в множестве $\mathcal{D}$ вырожденных планов. Кроме того, для обоснования конечности алгоритма предположим, что при каждом $\theta \in R$ существует не более двух θ -оптимальных базисов. В этих предположениях, если $\mathcal{D} \neq \emptyset$, вычислительная процедура п. 2.2 позволяет за конечное число итераций $\bar{q} + \bar{q}$ построить последовательность базисов $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_{\bar{q}}$; $\mathcal{B}'_1 = \mathcal{B}_1, \mathcal{B}'_2, \dots, \mathcal{B}'_{\bar{q}}$ и соответствующие последовательности таких промежутков

$$\Theta^+(\mathcal{B}_1) = [0, \theta^+(\mathcal{B}_1)], \quad \Theta^+(\mathcal{B}_2) = (\theta^+(\mathcal{B}_1), \theta^+(\mathcal{B}_2)], \quad \dots, \quad \Theta^+(\mathcal{B}_{\bar{q}}) = (\theta^+(\mathcal{B}_{\bar{q}-1}), +\infty);$$

$$\Theta^-(\mathcal{B}_1) = [\theta^-(\mathcal{B}_1), 0], \quad \Theta^-(\mathcal{B}_2) = [\theta^-(\mathcal{B}_2), \theta^-(\mathcal{B}_1)], \quad \dots, \quad \Theta^-(\mathcal{B}'_{\bar{q}}) = (-\infty, \theta^-(\mathcal{B}'_{\bar{q}-1})),$$

что для любого $\theta \in \Theta^+(\mathcal{B}_k)$ ($\theta \in \Theta^-(\mathcal{B}'_k)$) при любом $k \in 1:\bar{q}$ ($k \in 1:\bar{q}$), кроме, быть может, $k = \bar{q}$ ($k = \bar{q}$), планы $x(\mathcal{B}_k)$ ($x(\mathcal{B}'_k)$) являются оптимальными. В промежутках $\Theta^+(\mathcal{B}_{\bar{q}})$ и $\Theta^-(\mathcal{B}'_{\bar{q}})$ целевая функция f_θ задачи $(\mathcal{D}, f_\theta)$ может быть неограниченной.

Случай пустоты множества $\mathcal{D}$ также определяется в процессе счета.

Сделанные выше предположения не являются принципиальными. Если в множестве $\mathcal{D}$ есть вырожденные планы, то для устранения заикливания может быть применен любой метод § 7 гл. 4. Невыполнение второго предположения также, вообще говоря, может привести к заикливанию. В п. 2.6 будет указано, какие изменения следует внести в этом случае в вычислительную процедуру.

2.2. Алгоритм. Пусть $\mathcal{B}$ — допустимый базис задачи, $N(\mathcal{B}) = \{j_1, \dots, j_m\}$. Положим $c'(\mathcal{B}) = (c'[j_1], \dots, c'[j_m])$, $c''(\mathcal{B}) = (c''[j_1], \dots, c''[j_m])$,

$$b'(\mathcal{B})[0] = c'(\mathcal{B})b(\mathcal{B}), \quad a'_0(\mathcal{B}) = -c' + c'(\mathcal{B})A(\mathcal{B}), \quad (2)$$

$$b''(\mathcal{B})[0] = c''(\mathcal{B})b(\mathcal{B}), \quad a''_0(\mathcal{B}) = -c'' + c''(\mathcal{B})A(\mathcal{B}). \quad (3)$$

В процессе вычислений будут построены последовательности таблиц

$$T(\mathcal{B}_1), T(\mathcal{B}_2), \dots, T(\mathcal{B}_q), \quad (4)$$

$$T(\mathcal{B}'_1), T(\mathcal{B}'_2), \dots, T(\mathcal{B}'_q), \quad (5)$$

$$\text{вида} \quad T(\mathcal{B}) = \begin{array}{|c|c|c|} \hline & b''(\mathcal{B})[0] & a''_0(\mathcal{B}) \\ \hline & b'(\mathcal{B})[0] & a'_0(\mathcal{B}) \\ \hline N(\mathcal{B}) & b(\mathcal{B}) & A(\mathcal{B}) \\ \hline \end{array}. \quad (6)$$

Последовательность (4) соответствует последовательности θ -оптимальных базисов $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_q$ задачи (1), которые, кроме первого, будут строиться на этапе 1 алгоритма вместе с соответствующими таблицами. Аналогичным образом на этапе 2 формируется последовательность (5). Она соответствует последовательности θ -оптимальных базисов $\mathcal{B}'_1 = \mathcal{B}_1, \mathcal{B}'_2, \dots, \mathcal{B}'_q$.

Этап 0. 1. Находим допустимый базис $\mathcal{B}$ задачи (1) (см., например, п. 4.3.8). Возможны два случая.

1'. Такой базис $\mathcal{B}$ найден. Переходим к рекомендации 2.

1''. Задача (1) не имеет допустимых базисов. Вычислительный процесс закончен, $\mathcal{D} = \emptyset$.

2. Полагаем $\theta = 0$ (вообще говоря, можно взять любое $\theta \in R$). Начиная с допустимого базиса $\mathcal{B}$ любым из методов § 3 или 4 гл. 4 решается каноническая задача линейного программирования $(\mathcal{D}, f_\theta)$. Если задача $(\mathcal{D}, f_\theta)$ имеет решение, то будут построены оптимальный базис $\mathcal{B}_1$ и оптимальный план $x(\mathcal{B}_1)$. Переходим к рекомендации 3. В противном случае (функция f_θ не ограничена сверху на $\mathcal{D}$) будут обнаружены такой базис $\mathcal{B}$ и такое $l \in 1:n \setminus N(\mathcal{B})$, что $a(\mathcal{B})[0, l] < 0$ и $a'(\mathcal{B}) \leq 0$. Переходим к этапу 3.

3. Полагаем $\theta^+(\mathcal{B}_0) = 0$.

4. По θ -оптимальному базису $\mathcal{B}_1$ заполняется табл. (6) при $\mathcal{B} = \mathcal{B}_1$. Если для нахождения базиса $\mathcal{B}_1$ использовалась вычислительная процедура п. 4.3.3, то табл. $T(\mathcal{B}_1)$ получается при соединении к табл. $T'(\mathcal{B}_1)$ (см. в п. 4.3.3 табл. (6) при $\mathcal{B} = \mathcal{B}_1$) сверху еще одной строки: $(b''(\mathcal{B})[0], a''_0(\mathcal{B}))$, элементы которой вычисляются по формулам (3).

5. Полагаем номер итерации $k = 1$. Переходим к этапу 1.

Этап 1. Разбиение полуоси $[0, +\infty)$.

1. В табл. $T(\mathcal{B}_k)$ рассматривается строка $a''_0(\mathcal{B}_k)$. Возможны два случая.

1'. $a''_0(\mathcal{B}_k) \geq 0$. Полагаем $\Theta^+(\mathcal{B}_k) = (\theta^+(\mathcal{B}_{k-1}), +\infty)$. План $x(\mathcal{B}_k)$ является оптимальным при любом $\theta \in \Theta^+(\mathcal{B}_k)$. В столбце $N(\mathcal{B}_k)$ табл. $T(\mathcal{B}_k)$ расположены номера его базисных компонент, рядом, в столбце $b(\mathcal{B}_k)$, — их значения. Остальные $n - m$

компонент плана $x(\mathcal{B}_k)$ равны нулю. Максимальное значение $f_\theta(x(\mathcal{B}_k))$ целевой функции f_θ равно $b'(\mathcal{B}_k)[0] + \theta b''(\mathcal{B}_k)[0]$. Полагаем $\theta^-(\mathcal{B}_0) = 0$, $\mathcal{B}'_1 = \mathcal{B}_1$, $k = 1$ и переходим к этапу 2.

1''. Существует такое j , что $a''(\mathcal{B}_k)[0, j] < 0$. Переходим к рекомендации 2.

2. Вычисляется

$$\theta^+(\mathcal{B}_k) = \min_{j: a''(\mathcal{B}_k)[0, j] < 0} -a'(\mathcal{B}_k)[0, j]/a''(\mathcal{B}_k)[0, j]. \quad (7)$$

Будет показано, что $\theta^+(\mathcal{B}_1) \geq 0$, $\theta^+(\mathcal{B}_k) \geq \theta^+(\mathcal{B}_{k-1})$. Полагаем l равным минимальному из индексов p , для которых

$$\theta^+(\mathcal{B}_k) = -a'(\mathcal{B}_k)[0, p]/a''(\mathcal{B}_k)[0, p],$$

и $\theta^+(\mathcal{B}_k) = (\theta^+(\mathcal{B}_{k-1}), \theta^+(\mathcal{B}_k)]$. При любом $\theta \in \theta^+(\mathcal{B}_k)$ план $x(\mathcal{B}_k)$ является оптимальным. Его и максимальное значение целевой функции f_θ можно выписать из табл. $T(\mathcal{B}_k)$ (см. рекомендацию 1').

3. Рассматривается вектор $a'(\mathcal{B}_k)$. Возможны два случая.

3'. $a'(\mathcal{B}_k) \leq 0$. Полагаем $\theta^+ = (\theta^+(\mathcal{B}_k), +\infty)$. При любом $\theta \in \theta^+$ функция f_θ не ограничена сверху на $\mathcal{D}$. Полагаем $\theta^-(\mathcal{B}_0) = 0$, $\mathcal{B}'_k = \mathcal{B}_k$, номер итерации $k = 1$. Переходим к этапу 2.

3''. Существует такое $i \neq 0$, что $a(\mathcal{B}_k)[i, l] > 0$. Переходим к рекомендации 4.

4. Вычисляется

$$\lambda(\mathcal{B}_k) = \min_{i: a(\mathcal{B}_k)[i, l] > 0} \frac{b(\mathcal{B}_k)[i]}{a(\mathcal{B}_k)[i, l]} = \frac{b(\mathcal{B}_k)[r]}{a(\mathcal{B}_k)[r, l]}. \quad (8)$$

Элемент $a(\mathcal{B}_k)[r, l]$ является ведущим. Если в множестве $\mathcal{D}$ отсутствуют вырожденные планы, то формула (8) определяет его однозначно (см. утверждение 11° п. 4.3.6). Столбец a^{lr} выводится из базиса $\mathcal{B}_k = \{j_1, \dots, j_m\}$, формируется новый базис

$$\mathcal{B}_{k+1} = \{a^{l_1}, \dots, a^{l_{r-1}}, a^l, a^{l_{r+1}}, \dots, a^{l_m}\},$$

$$N(\mathcal{B}_{k+1}) = \{j_1, \dots, j_{r-1}, l, i_{r+1}, \dots, i_m\}.$$

5. Заполняется табл. $T(\mathcal{B}_{k+1})$. В табл. $T(\mathcal{B}_{k+1})$ записывается столбец $N(\mathcal{B}_{k+1})$, остальная часть ее заполняется по формулам

$$b(\mathcal{B}_{k+1})[r] = \frac{b(\mathcal{B}_k)[r]}{a(\mathcal{B}_k)[r, l]}, \quad (9)$$

$$a(\mathcal{B}_{k+1})[r, j] = \frac{a(\mathcal{B}_k)[r, j]}{a(\mathcal{B}_k)[r, l]}, \quad j \in 1:n, \quad (10)$$

$$b(\mathcal{B}_{k+1})[i] = b(\mathcal{B}_k)[i] - a(\mathcal{B}_k)[i, l] b(\mathcal{B}_{k+1})[r], \quad i \in 1:m, i \neq r, \quad (11)$$

$$b'(\mathcal{B}_{k+1})[0] = b'(\mathcal{B}_k)[0] - a'(\mathcal{B}_k)[0, l] b(\mathcal{B}_{k+1})[r], \quad (11')$$

$$b''(\mathcal{B}_{k+1})[0] = b''(\mathcal{B}_k)[0] - a''(\mathcal{B}_k)[0, l] b(\mathcal{B}_{k+1})[r], \quad (11'')$$

$$a(\mathcal{B}_{k+1})[i, j] = a(\mathcal{B}_k)[i, j] - a(\mathcal{B}_k)[i, l]a(\mathcal{B}_{k+1})[r, j], \quad (12)$$

$$i \in 1:m, i \neq r, j \in 1:n,$$

$$a'(\mathcal{B}_{k+1})[0, j] = a'(\mathcal{B}_k)[0, j] - a'(\mathcal{B}_k)[0, l]a(\mathcal{B}_{k+1})[r, j], \quad (12')$$

$$j \in 1:n,$$

$$a''(\mathcal{B}_{k+1})[0, j] = a''(\mathcal{B}_k)[0, j] - a''(\mathcal{B}_k)[0, l]a(\mathcal{B}_{k+1})[r, j], \quad (12'')$$

$$j \in 1:n.$$

6. Полагаем номер итерации $k = k + 1$. Переходим к рекомендации 1.

Этап 2. Разбиение полуоси $(-\infty, 0]$.

1. В табл. $T(\mathcal{B}'_k)$ рассматривается строка $a'_0(\mathcal{B}'_k)$. Возможны два случая.

1'. $a'_0(\mathcal{B}'_k) \leq 0$. Полагаем $\Theta^-(\mathcal{B}'_k) = (-\infty, \Theta^-(\mathcal{B}'_{k-1}))$. Вычислительный процесс закончен. План $x(\mathcal{B}'_k)$ является оптимальным при любом $\theta \in \Theta^-(\mathcal{B}'_k)$. Этот план и максимальное значение целевой функции f_θ можно выписать из табл. $T(\mathcal{B}'_k)$ (см. рекомендацию 1' этапа 1).

1''. Существует такое j , что $a''(\mathcal{B}'_k)[0, j] > 0$. Переходим к рекомендации 2.

2. Вычисляется

$$\Theta^-(\mathcal{B}'_k) = \max_{l: a''(\mathcal{B}'_k)[0, l] > 0} \frac{-a'(\mathcal{B}'_k)[0, l]}{a''(\mathcal{B}'_k)[0, l]}. \quad (13)$$

Будет показано, что $\Theta^-(\mathcal{B}'_k) \leq 0$, $\Theta^-(\mathcal{B}'_k) \leq \Theta^-(\mathcal{B}'_{k-1})$. Полагаем l равным наименьшему из номеров p , для которых

$$\theta^-(\mathcal{B}'_k) = -a'(\mathcal{B}'_k)[0, p]/a''(\mathcal{B}'_k)[0, p],$$

$\Theta^-(\mathcal{B}'_k) = [\theta^-(\mathcal{B}'_k), \theta^-(\mathcal{B}'_{k-1}))$. При любом $\theta \in \Theta^-(\mathcal{B}'_k)$ план $x(\mathcal{B}'_k)$ является оптимальным. Этот план и максимальное значение целевой функции f_θ можно выписать из табл. $T(\mathcal{B}'_k)$ (см. рекомендацию 1' этапа 1).

3. Рассматривается вектор $a^l(\mathcal{B}'_k)$. Возможны два случая.

3'. $a^l(\mathcal{B}'_k) \leq 0$. Полагаем $\Theta^- = (-\infty, \theta(\mathcal{B}'_k))$. Вычислительный процесс закончен. При любом $\theta \in \Theta^-$ функция f_θ не ограничена сверху на $\mathcal{D}$.

3''. Существует такое $i \neq 0$, что $a(\mathcal{B}'_k)[i, l] > 0$. Выполняем рекомендации 4 и 5 этапа 1. Полагаем номер итерации $k = k + 1$. Переходим к рекомендации 1 этапа 2.

Этап 3. К этому этапу переход осуществляется, когда обнаружена неограниченность сверху целевой функции f_0 на $\mathcal{D}$. Таким образом (см. доказательство утверждения 1° п. 1.2), здесь будет определен неограниченный промежуток, содержащий точку $\theta = 0$. Для каждого θ из этого промежутка задача $(\mathcal{D}, f_\theta)$

не имеет решения (функция f_θ не ограничена сверху на $\mathcal{D}$). Если найденный промежуток не совпадает с R , то определяется θ , начиная с которого можно продолжить решение задачи (1) на этапе 1 или 2.

1. По базису $\mathcal{B}$ заполняется табл. (6).

2. Рассматриваются множество $J(\mathcal{B}) = \{j \in 1:n: a'(\mathcal{B})[0, j] < 0, a'(\mathcal{B}) \leq 0\}$ и строка $a''_0(\mathcal{B})$. Возможны два случая.

2'. Среди элементов $a''(\mathcal{B})[0, j]$ с индексами $j \in J(\mathcal{B})$ существует по крайней мере один, равный нулю, или одновременно два разного знака. Вычислительный процесс закончен. Для любого $\theta \in R$ целевая функция f_θ не ограничена сверху на $\mathcal{D}$.

2''. Если $a''(\mathcal{B})[0, j] > 0$ для любого $j \in J(\mathcal{B})$, то вычисляется

$$\theta^+ = \max_{j \in J(\mathcal{B})} -a'(\mathcal{B})[0, j]/a''(\mathcal{B})[0, j]. \quad (14)$$

При любом $\theta \in (-\infty, \theta^+)$ в задаче (1) функция f_θ не ограничена сверху на $\mathcal{D}$. Переходим к рекомендации 3.

Если же $a''(\mathcal{B})[0, j] < 0$ для любого $j \in J(\mathcal{B})$, то вычисляется

$$\theta^- = \min_{j \in J(\mathcal{B})} -a'(\mathcal{B})[0, j]/a''(\mathcal{B})[0, j]. \quad (15)$$

При любом $\theta \in (\theta^-, +\infty)$ в задаче (1) целевая функция f_θ не ограничена сверху на $\mathcal{D}$. Переходим к рекомендации 3.

3. Начиная с базиса $\mathcal{B}$ любым методом § 3 или 4 гл. 4 решается задача (1) при $\theta = \theta^+$ или $\theta = \theta^-$. Если найдены θ -оптимальный базис $\mathcal{B}_1$ и план $x(\mathcal{B}_1)$, то полагаем $\theta^+(\mathcal{B}_0) = \theta^+$ (или $\theta^-(\mathcal{B}_0) = \theta^-$) и переходим к выполнению рекомендации 4 этапа 0 (или, полагая $\mathcal{B}'_1 = \mathcal{B}_1$, — к этапу 2). Если $\theta = \theta^+$, то функция f_θ не ограничена сверху на $\mathcal{D}$ при любом $\theta \in (-\infty, \theta^+)$. Следовательно, после выполнения всех предписаний этапа 1 вычислительный процесс заканчивается.

Если функция f_{θ^+} (или f_{θ^-}) не ограничена сверху на $\mathcal{D}$, то будут обнаружены такой базис $\mathcal{B}$ и такое $l \in 1:n \setminus N(\mathcal{B})$, что $a(\mathcal{B})[0, l] < 0$ и $a'(\mathcal{B}) \leq 0$. Переходим к рекомендации 1 этапа 3. ■

2.3. Пример. В табл. (17), (18), состоящих соответственно из табл. $T(\mathcal{B}_1)$, $T(\mathcal{B}_2)$, $T(\mathcal{B}_3)$ ($\bar{q}=3$) и $T(\mathcal{B}'_1)$, $T(\mathcal{B}'_2)$, $T(\mathcal{B}'_3)$ ($\bar{q}=3$), приведено решение задачи $(\mathcal{D}, f_\theta)$:

$$\begin{aligned} f_\theta(x) &= (\theta - 1)x_1 + (3\theta - 2)x_2 + x_3 - 3x_4 + (4 + \theta)x_5 \rightarrow \max, \\ \mathcal{D} = \{x \in R^5: & \begin{array}{rrrrr} 3x_1 & & + x_3 + x_4 & - x_5 & = 3, \\ x_1 + x_2 & & + 4x_3 & & + 3x_5 = 2, \\ x_j & \geq 0, & j \in 1:5 \}. \end{array} \end{aligned} \quad (16)$$

$$\begin{aligned}
 T(\mathcal{B}_1) &= \begin{array}{c|cccccc} & 14/10 & 0 & -26/10 & 18/10 & 2/10 & 0 \\ & 1/10 & 0 & 31/10 & 27/10 & 23/10 & 0 \\ 1 & 11/10 & 1 & 1/10 & 7/10 & 3/10 & 0 \\ 5 & 3/10 & 0 & 3/10^* & 11/10 & -1/10 & 1 \end{array} \\
 \\
 T(\mathcal{B}_2) &= \begin{array}{c|cccccc} & 4 & 0 & 0 & 34/3 & -2/3 & 26/3 \\ & -3 & 0 & 0 & -26/3 & 10/3 & -31/3 \\ 1 & 1 & 1 & 1 & 1/3 & 1/3^* & -1/3 \\ 2 & 1 & 0 & 1 & 11/3 & -1/3 & 10/3 \end{array} \\
 \\
 T(\mathcal{B}_3) &= \begin{array}{c|cccccc} & 6 & 2 & 0 & 12 & 0 & 8 \\ & -13 & -10 & 0 & -12 & 0 & -7 \\ 4 & 3 & 3 & 0 & 1 & 1 & -1 \\ 2 & 2 & 1 & 1 & 4 & 0 & 3 \end{array}
 \end{aligned} \tag{17}$$

$$\begin{aligned}
 T(\mathcal{B}'_1) &= \begin{array}{c|cccccc} & 14/10 & 0 & -26/10 & 18/10 & 2/10 & 0 \\ & 1/10 & 0 & 31/10 & 27/10 & 23/10 & 0 \\ 1 & 11/10 & 1 & 1/10 & 7/10 & 3/10 & 0 \\ 5 & 3/10 & 0 & 3/10 & 11/10^* & -1/10 & 1 \end{array} \\
 \\
 T(\mathcal{B}'_2) &= \begin{array}{c|cccccc} & 10/11 & 0 & -34/11 & 0 & 4/11 & -18/11 \\ & -7/11 & 0 & 26/11 & 0 & 28/11 & -27/11 \\ 1 & 10/11 & 1 & -1/11 & 0 & 4/11^* & -7/11 \\ 3 & 3/11 & 0 & 3/11 & 1 & -1/11 & 10/11 \end{array} \\
 \\
 T(\mathcal{B}'_3) &= \begin{array}{c|cccccc} & 0 & -1 & -3 & 0 & 0 & -1 \\ & -7 & -7 & 3 & 0 & 0 & 2 \\ 4 & 10/4 & 11/4 & -1/4 & 0 & 1 & -7/4 \\ 3 & 2/4 & 1/4 & 1/4 & 1 & 0 & 3/4 \end{array}
 \end{aligned} \tag{18}$$

Отмечены ведущие элементы, выделены ведущие строки и столбцы. Решение начато с базиса $\mathcal{B}_1$ ($N(\mathcal{B}_1) = \{1, 5\}$). Этот базис является оптимальным при $\theta = 0$ (строка $a'_0(\mathcal{B}_1) = (0, {}^{31}/_{10}$,

$^{27/10}, ^{23/10}, 0) > 0$). Поскольку в строке $a_0''(\mathcal{B}_1) = (0, -^{26/10}, ^{18/10}, ^{2/10}, 0)$ есть отрицательный элемент, то по формуле (7) $\theta^+(\mathcal{B}_1) = -^{31/10}$; $-^{26/10} = ^{31/26}$, $l=2$ и $\theta^+(\mathcal{B}_1) = (0, ^{31/26})$. Следовательно, для любого $\theta \in [0, ^{31/26}]$ базис $\mathcal{B}_1$ является θ -оптимальным, поскольку (см. доказательство утверждения 1° п. 1.2) для таких θ

$$a(\mathcal{B}_1, \theta)[0, 1] = a'(\mathcal{B}_1)[0, 1] + \theta a''(\mathcal{B}_2)[0, 1] = 0 + \theta \cdot 0 = 0,$$

$$a(\mathcal{B}_1, \theta)[0, 2] = ^{31/10} + \theta \cdot (-^{26/10}) \geq 0,$$

$$a(\mathcal{B}_1, \theta)[0, 3] = ^{27/10} + \theta \cdot ^{18/10} > 0,$$

$$a(\mathcal{B}_1, \theta)[0, 4] = ^{23/10} + \theta \cdot ^{2/10} > 0,$$

$$a(\mathcal{B}_1, \theta)[0, 5] = 0 + \theta \cdot 0 = 0.$$

Отметим, что $a(\mathcal{B}_1, ^{31/26})[0, 2] = 0$ и $a(\mathcal{B}_1, \theta)[0, 2] < 0$ при $\theta > ^{31/26}$. Следовательно, $x(\mathcal{B}_1)$ не является решением задачи $(\mathcal{D}, f_0)$ при $\theta > ^{31/26}$. Полагаем $l=2$. В базис вводится столбец a^2 . Так как $a^2(\mathcal{B}_1) > 0$, то выполняются рекомендации 4 этапа 1.

Дальнейшие вычисления осуществляются в соответствии с процедурой симплекс-метода. Заполняется табл. $T(\mathcal{B}_2)$. В строке $a_0''(\mathcal{B}_2) = (0, 0, ^{34/3}, -^{2/3}, ^{26/3})$ один отрицательный элемент. Следовательно, $\theta^+(\mathcal{B}_2) = -^{10/3}$; $-^{2/3} = 5 > \theta^+(\mathcal{B}_1) = ^{31/26}$, и для каждого $\theta \in \theta^+(\mathcal{B}_2) = (^{31/26}, 5]$ базис $\mathcal{B}_2 = \{a^1, a^2\}$ является θ -оптимальным.

Вычисления на этапе 1 завершаются заполнением табл. $T(\mathcal{B}_3)$. Здесь $a_0''(\mathcal{B}_3) = (2, 0, 12, 0, 8)$, и, следовательно, для каждого $\theta \in (5, +\infty)$ базис $\mathcal{B}_3$ является θ -оптимальным. Луч R^+ разбит на три промежутка: $[0, ^{31/26}]$, $\theta^+(\mathcal{B}_2)$, $\theta^+(\mathcal{B}_3)$.

На этапе 2 (табл. $T(\mathcal{B}_1)$, $T(\mathcal{B}_2)$, $T(\mathcal{B}_3)$) разбивается луч $R^- = \{x \in R: x \leq 0\}$. Из табл. $T(\mathcal{B}_1) \equiv T(\mathcal{B}_1)$ имеем (см. (13)) $\theta^-(\mathcal{B}_1) = \max\{-^{27/10}; ^{18/10}, -^{23/10}; ^{2/10}\} = -^3/1$, $l=3$, и для каждого $\theta \in [-^3/2, 0]$ базис $\mathcal{B}_1$ является θ -оптимальным. Принимая во внимание результаты вычислений на этапе 1, получаем, что базис $\mathcal{B}_1$ является θ -оптимальным при любом $\theta \in [-^3/2, ^{31/26}]$, и т. д.

Итоги. Прямая R разбита на пять интервалов: $\theta^-(\mathcal{B}_3) = (-\infty, -7)$ (табл. $T(\mathcal{B}_3)$), $\theta^-(\mathcal{B}_2) = [-7, -^3/2]$ (табл. $T(\mathcal{B}_2)$), $[-^3/2, ^{31/26}]$, $\theta^+(\mathcal{B}_2)$ и $\theta^+(\mathcal{B}_3)$. Для каждого из них указан θ -оптимальный базис ($N(\mathcal{B}_3) = \{4, 3\}$, $N(\mathcal{B}_2) = \{1, 3\}$, $N(\mathcal{B}_1) = N(\mathcal{B}_1) = \{1, 5\}$, $N(\mathcal{B}_2) = \{1, 2\}$, $N(\mathcal{B}_3) = \{4, 2\}$). Из табл. (17), (18) можно выписать решения соответствующих задач (при θ из указанных промежутков): $x(\mathcal{B}_3) = (0, 0, ^{1/2}, ^{5/2}, 0)$, $x(\mathcal{B}_2) = (^{10/11}, 0, ^{3/11}, 0, 0)$, $x(\mathcal{B}_1) \equiv x(\mathcal{B}_1) = (^{11/10}, 0, 0, 0, ^{3/10})$, $x(\mathcal{B}_2) = (1, 1, 0, 0, 0)$, $x(\mathcal{B}_3) = (0, 2, 0, 3, 0)$; максимальные значения целевых функций: $f_0(x(\mathcal{B}_3)) = -7$, $f_0(x(\mathcal{B}_2)) = -^{7/11} + ^{10/11}\theta$, $f_0(x(\mathcal{B}_1)) = f_0(x(\mathcal{B}_1)) = ^{1/10} + ^{14/10}\theta$, $f_0(x(\mathcal{B}_2)) = -3 + 4\theta$, $f_0(x(\mathcal{B}_3)) = -13 + 6\theta$. Тем самым приведено описание решения параметрической задачи (16) (см. п. 1.1).

2.4. Обоснование алгоритма. Как уже отмечалось, основой алгоритма п. 2.2 является вычислительная процедура симплекс-метода п. 4.3.3. Поэтому утверждения 1° — 12° п. 4.3.6 и их доказательства в п. 4.3.7 (предполагается, что в множестве $\mathcal{D}$ нет вырожденных планов) служат обоснованием рекомендаций к нахождению оптимального базиса и оптимального плана для каждого $\theta \in (\theta^+(\mathcal{B}_k), \theta^+(\mathcal{B}_{k+1}))$ ($\theta \in [\theta^-(\mathcal{B}_{k+1}), \theta^-(\mathcal{B}_k)]$). Проверка формул (2), (3) предоставляется читателю. — **Упражнение 7.2.** Из утверждения 1° п. 1.2 следует принципиальная возможность разбиения прямой R на промежутки, удовлетворяющие свойствам, указанным в условии этого утверждения. Поэтому для завершения обоснования алгоритма остается установить, что алгоритм п. 2.3 осуществляет именно такое разбиение и за конечное число итераций.

1°. При любом $k = 1, 2, \dots$

$$\theta^+(\mathcal{B}_{k-1}) \leq \theta^+(\mathcal{B}_k), \quad \theta^-(\mathcal{B}'_{k-1}) \geq \theta^-(\mathcal{B}'_k). \quad (19)$$

2°. Если для каждого $\theta \in R$ существует не более двух θ -оптимальных базисов, то число итераций $\bar{q} + \bar{q}$ конечно.

2.5. Доказательство утверждений 1° — 2°.

1°. Так как множество $\mathcal{D}$ не содержит вырожденных планов, то в силу утверждения 4° п. 4.3.6 и комментария к нему базис $\mathcal{B}$ θ -оптимален в том и только том случае, если $a_0(\mathcal{B}, \theta) \geq 0$. При доказательстве утверждения 1° п. 1.2 установлено, что $a_0(\mathcal{B}, \theta) = a'_0(\mathcal{B}) + \theta a''_0(\mathcal{B})$. Следовательно, при $\theta = 0$ и $\mathcal{B} = \mathcal{B}_1$ получаем $a'(\mathcal{B})[0, j] \geq 0$, $j \in 1:n$, и неравенства (19) при $k = 1$ следуют из (7) и определения $\theta^+(\mathcal{B}_0)$ ($\theta^+(\mathcal{B}_0) = 0$).

Далее, в силу $\theta^+(\mathcal{B}_k)$ -оптимальности базиса $\mathcal{B}_k$ при каждом $j \in 1:n$ справедливо неравенство

$$a'(\mathcal{B}_k)[0, j] + \theta^+(\mathcal{B}_k)a''(\mathcal{B}_k)[0, j] \geq 0, \quad (20)$$

а для индекса l столбца, который вводится в базис по рекомендации 4 этапа 1:

$$a'(\mathcal{B}_k)[0, l] + \theta^+(\mathcal{B}_k)a''(\mathcal{B}_k)[0, l] = 0. \quad (21)$$

Тогда, используя формулы (12'), (12''), для любого $j \in 1:n$ получаем

$$\begin{aligned} & a'(\mathcal{B}_{k+1})[0, j] + \theta^+(\mathcal{B}_k)a''(\mathcal{B}_{k+1})[0, j] = \\ & = a'(\mathcal{B}_k)[0, j] - a'(\mathcal{B}_k)[0, l]a(\mathcal{B}_{k+1})[r, j] + \\ & + \theta^+(\mathcal{B}_k)(a''(\mathcal{B}_k)[0, j] - a''(\mathcal{B}_k)[0, l]a(\mathcal{B}_{k+1})[r, j]) = \\ & = a'(\mathcal{B}_k)[0, j] + \theta^+(\mathcal{B}_k)a''(\mathcal{B}_k)[0, j] \geq 0, \quad j \in 1:n, \end{aligned}$$

т. е. базис $\mathcal{B}_{k+1}$ является $\theta^+(\mathcal{B}_k)$ -оптимальным. Поэтому

$$\theta^+(\mathcal{B}_k) \leq -a'(\mathcal{B}_{k+1})[0, j]/a''(\mathcal{B}_{k+1})[0, j],$$

если $a''(\mathcal{B}_{k+1})[0, j] < 0$, и, следовательно, $\theta^+(\mathcal{B}_k) \leq \theta^+(\mathcal{B}_{k+1})$.

Доказательство неравенств (19) при $\mathcal{B} = \mathcal{B}'_k$ (оно осуществляется аналогично) полезно провести самостоятельно. — Упражнение 7.3. ■

2°. Так как для каждого $\theta \in R$ существует не более двух θ -оптимальных базисов, то при фиксированном k ($k = 1, 2, \dots$) может найтись не более одного $l \in 1:n \setminus N(\mathcal{B}_k)$, для которого выполнялось бы равенство (21) и $a(\mathcal{B}_k)[i, l] > 0$ при некотором $i \in 1:m$. В самом деле, как было отмечено при доказательстве утверждения 1°, вводя любой из таких столбцов в базис, получаем новые $\theta^+(\mathcal{B}_k)$ -оптимальные базисы. Поэтому либо найден (бесконечный) интервал неограниченности целевой функции, либо неравенства (20) выполняются как строгие неравенства для всех $j \in 1:n \setminus (N(\mathcal{B}_k) \cup \{l\})$. В последнем случае $\theta^+(\mathcal{B}_k) < \theta^+(\mathcal{B}_{k+1})$ ($\theta^-(\mathcal{B}_{k+1}) < \theta^-(\mathcal{B}_k)$), и интервал $(\theta^+(\mathcal{B}_k), \theta^+(\mathcal{B}_{k+1}))$ θ -оптимальности базиса $\mathcal{B}_{k+1}$ имеет положительную длину. Тогда конечность вычислительной процедуры следует из конечности числа допустимых базисов и невозможности возвращения к «старому» базису. ■

2.6. Измененная вычислительная процедура. Если при некотором $\theta \in R$ существует более двух θ -оптимальных базисов, то, как уже отмечалось, при решении задачи (1) методом, описанным в п. 2.2, возможно заикливание. Точнее, в процессе вычислений может встретиться такая последовательность базисов

$\mathcal{B}_k, \mathcal{B}_{k+1}, \dots, \mathcal{B}_{k+i}(\mathcal{B}'_k, \mathcal{B}'_{k+1}, \dots, \mathcal{B}'_{k+i})$, что

$$\theta^+(\mathcal{B}_k) = \theta^+(\mathcal{B}_{k+1}) = \dots = \theta^+(\mathcal{B}_{k+i}) \quad (\theta^-(\mathcal{B}'_k) = \theta^-(\mathcal{B}'_{k+1}) = \dots = \theta^-(\mathcal{B}'_{k+i})), \text{ где } \mathcal{B}_k = \mathcal{B}_{k+i} \quad (\mathcal{B}'_k = \mathcal{B}'_{k+i}), \quad i > 2.$$

Вероятно, условие $i > 2$ можно усилить и попытаться доказать, что заикливание невозможно, если в точке $\theta^+(\mathcal{B}_k)$ ($\theta^-(\mathcal{B}_k)$) существуют три, четыре и т. д. θ -оптимальных базиса. Нахождение точного числа, гарантирующего поступательное движение в алгоритме п. 2.2, является интересной задачей. — Упражнение 7.4.

Ниже приведена принципиальная схема метода, который «не боится» заикливания.

Этап 0. Полагаем $G = R$ и переходим к этапу 1.

Этап 1. 1. Рассматривается произвольное θ из G .

2. Любым методом гл. 4 решается задача $(\mathcal{D}, f_\theta)$. Возможны три случая.

2'. Множество $\mathcal{D} = \emptyset$. Вычислительный процесс закончен. Для каждого $\theta \in R$ задача $(\mathcal{D}, f_\theta)$ не имеет решения.

2''. В задаче $(\mathcal{D}, f_\theta)$ целевая функция f_θ не ограничена сверху на $\mathcal{D}$. Применяя процедуру этапа 3 алгоритма п. 2.2, определяем интервал неограниченности целевой функции f_θ . Если он совпадает с G , вычислительный процесс закончен. Для каждого $\theta \in G$ функция f_θ не ограничена сверху на $\mathcal{D}$. Если интервал неограниченности функции f_θ не совпадает с G , то

удаляем его из G и полагаем G равным оставшейся части. Переходим к рекомендации 1.

2'''. Задача $(\mathcal{D}, f_\theta)$ имеет решение $x(\theta)$. Определяется (например, так же, как это делается на этапах 1 и 2 алгоритма п. 2.2) интервал θ -оптимальности базиса $\mathcal{B}$. Если он совпадает с G , вычислительный процесс закончен (для любого $\theta \in G$ базис $\mathcal{B}$ является θ -оптимальным), если не совпадает, то удаляем его из G и полагаем G равным оставшейся части. Переходим к рекомендации 1. ■

Конечность приведенного алгоритма следует из утверждения 1° п. 1.2 и отсутствия вырожденных планов в множестве $\mathcal{D}$. ■

2.7. О методах решения параметрической задачи линейного программирования с параметром в векторе ограничений. В п. 1.2 (см. замечание) мы уже говорили о принципиальной возможности сведения решения параметрической задачи линейного программирования с параметром в векторе ограничений (см. задачу (3) п. 1.2) к решению некоторой параметрической задачи линейного программирования с параметром в целевой функции (см. задачу (2) п. 1.2). Однако такое сведение не всегда является рациональным, и поэтому интерес представляют методы непосредственного решения задачи (3) п. 1.2. В настоящей книге соответствующие алгоритмы не приводятся (подробное изложение одного из них содержится в [1]). Предлагаем читателю самостоятельно разработать алгоритм решения задачи (3) п. 1.2, используя алгоритм п. 4.5.3 (— Упражнение 7.5) или идеи вычислительной схемы п. 2.6 (— Упражнение 7.6).

УПРАЖНЕНИЯ

7.1. См. с. 257. 7.2. См. с. 265. 7.3. См. с. 266. 7.4. См. с. 266. 7.5. См. с. 267. 7.6. См. с. 267.

7.7. Разработать алгоритм решения задачи (1) п. 1.2, использующий вычислительную схему п. 4.4.3.

7.8. Пусть $(\mathcal{D}, f): f(x) = cx \rightarrow \max$, $\mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\}$.

а) Возможна ли неограниченность f сверху на $\mathcal{D}' = \{x \in R^n: Ax = b', x \geq 0\}$, если задача $(\mathcal{D}, f)$ имеет решение? Может ли множество $\mathcal{D}'$ быть пустым, если $\mathcal{D} \neq \emptyset$?

б) Возможна ли неограниченность сверху функции $f'(x) = c'x$ на множестве $\mathcal{D}$, если задача $(\mathcal{D}, f)$ имеет решение? Каким должно быть множество $\mathcal{D}$, чтобы задача максимизации $(\mathcal{D}, f')$ имела решение при любом $c' \in R^n$?

Глава 8

О МЕТОДАХ РЕШЕНИЯ ЗАДАЧ ДИСКРЕТНОГО ПРОГРАММИРОВАНИЯ

В этой главе будут рассмотрены только некоторые проблемы дискретного программирования — весьма обширной, важной, имеющей многочисленные приложения области математического программирования. Отбор материала явился, вероятно, единственно возможным в рамках небольшой главы учебного пособия. Для более глубокого изучения теории и методов дискретного программирования следует обратиться, например, к монографиям [2, 27, 28, 45].

§ 1. Унимодулярные матрицы и целочисленность решения задач транспортного типа

1.1. Некоторые определения и обозначения. План $x \in R^n$ назовем *целочисленным* (x — ц. ч.), если все его компоненты целые числа. Экстремальную задачу

$$(\mathcal{D}_c, f): f(x) = cx \rightarrow \max, \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0, x \text{ — ц. ч.}\}, \quad (1)$$

где $A — m \times n$ матрица ограничений задачи; $b \in R^m$; $c \in R^n$; b — вектор ограничений задачи; c — вектор целевой функции, назовем *канонической задачей целочисленного линейного программирования* (см. пп. 2.2.4 и 2.2.2). Частный случай такой задачи рассматривался в § 3 гл. 5. Задача

$$(\mathcal{D}, f): f(x) = cx \rightarrow \max, \mathcal{D} = \{x \in R^n: Ax = b, x \geq 0\},$$

которая получается из задачи (1), если отбросить условия целочисленности плана x , является обычной канонической задачей линейного программирования, методы решения которой подробно рассмотрены в гл. 4. Ясно, что множество $\mathcal{D}_c$ допустимых планов задачи (1) является подмножеством множества $\mathcal{D}$ допустимых планов задачи $(\mathcal{D}, f)$.

В дальнейшем без специальных ссылок будем пользоваться понятиями и обозначениями п. 4.3.1.

1.2. Постановка проблемы. Обсуждение. Решить задачу $(\mathcal{D}_n, f)$ — значит указать ее оптимальный план либо установить, что такого плана не существует: $\mathcal{D}_n = \emptyset$ или f не ограничена сверху на $\mathcal{D}_n$. При разработке методов решения задачи $(\mathcal{D}_n, f)$ естественно теперь выяснить возможность использования для этой цели уже известных вычислительных процедур гл. 4, и в первую очередь условий, при которых в процессе решения задачи $(\mathcal{D}, f)$ может быть найден оптимальный план нелинейной задачи $(\mathcal{D}_n, f)$.

Если $\mathcal{D} \neq \emptyset$, то $\mathcal{D}$ — многогранное выпуклое множество (см. п. 1.5.4), и так как $\mathcal{D} \subset R^{n+}$, то $\text{ext } \mathcal{D} \neq \emptyset$ (см. п. 1.6.3). Будем говорить, что *многогранное выпуклое множество $\mathcal{D}$ (и в частности, выпуклый многогранник) целочисленно*, если $\text{ext } \mathcal{D} \neq \emptyset$ и все его вершины являются целочисленными планами.

В случае целочисленности множества допустимых планов задачи $(\mathcal{D}, f)$ справедливы следующие утверждения.

1°. Каждому допустимому базису $\mathcal{B}$ задачи $(\mathcal{D}, f)$ отвечает целочисленный базисный план $x(\mathcal{B})$.

2°. Если $\bar{x}(\mathcal{B})$ — базисное решение задачи $(\mathcal{D}, f)$, то $\bar{x}(\mathcal{B})$ является также и оптимальным планом задачи $(\mathcal{D}_n, f)$.

3°. Оптимальный план задачи $(\mathcal{D}, f)$, полученный любым методом гл. 4, является решением задачи $(\mathcal{D}_n, f)$.

Доказательство. 1°. Из примера п. 1.5.5 следует, что базисный план $x(\mathcal{B})$ является вершиной целочисленного многогранного выпуклого множества $\mathcal{D}$ и, следовательно, $x(\mathcal{B})$ — целочисленный вектор. ■

2°. Утверждение следует из включения $\mathcal{D}_n \subset \mathcal{D}$. ■

3°. Так как оптимальный план задачи $(\mathcal{D}, f)$, полученный любым методом гл. 4, является базисным, то $\bar{x}(\mathcal{B})$ — решение задачи $(\mathcal{D}, f)$ — является также и решением задачи $(\mathcal{D}_n, f)$. ■

Особо отметим следующее обстоятельство. Очевидно, что если $\mathcal{D} = \emptyset$, то и $\mathcal{D}_n = \emptyset$. Однако задача $(\mathcal{D}_n, f)$ может иметь решение и в том случае, когда задача $(\mathcal{D}, f)$ решения не имеет: функция f может быть неограниченной сверху на множестве $\mathcal{D}$, в то время как множество $\mathcal{D}_n$ будет состоять из конечного числа точек. Построение такого примера предоставляем читателю. — Упражнение 8.1.

Подведем итоги. Если задача $(\mathcal{D}, f)$ имеет оптимальный план и множество ее допустимых планов целочисленно, то задача $(\mathcal{D}_n, f)$ также имеет оптимальный план, который может быть получен в результате решения канонической задачи линейного программирования $(\mathcal{D}, f)$ методами гл. 4 для невырожденных или двойственно невырожденных задач, а также для вырожденных задач с учетом замечаний, сделанных в п. 4.3.4

и § 7 гл. 4. Условия, позволяющие установить существование решения задачи $(\mathcal{D}, f)$ и целочисленность множества $\mathcal{D}$, существенно упрощают решение задачи $(\mathcal{D}_u, f)$. Условие существования решения задачи $(\mathcal{D}, f)$ было получено в п. 3.3.7. Для некоторых задач (например, для транспортной задачи в матричной постановке) существование оптимального плана очевидно. Ниже будут приведены некоторые условия целочисленности множества $\mathcal{D}$.

1.3. Унимодулярные матрицы. Условие целочисленности многогранного выпуклого множества. Матрицу A назовем *унимодулярной*, если каждый ее минор равен 1, -1 или 0. Ясно, что каждый элемент такой матрицы равен 1, -1 или 0. Примером унимодулярных матриц могут служить матрицы транспортных задач в матричной постановке (см. п. 3.1.11, 3°). Унимодулярность таких матриц будет установлена в п. 1.5.

Если $m \times n$ матрица A унимодулярна, вектор $b \in R^m$ целочислен и $\mathcal{D} = \{x \in R^n : Ax = b, x \geq 0\} \neq \emptyset$, то многогранное выпуклое множество $\mathcal{D}$ целочисленно.

Доказательство. В примере п. 1.5.5 установлено, что $x_0 \in \mathcal{D}$ в том и только том случае является экстремальной точкой множества $\mathcal{D}$, если система $\mathcal{B}_0 = \{a^j, j \in 1:n \setminus J(x_0)\}$ столбцов матрицы A , индексы которых отвечают положительным компонентам вектора x_0 , линейно независима. Обозначим через k число векторов в множестве $\mathcal{B}_0$, через B_0 — квадратную подматрицу матрицы A порядка k , составленную из элементов столбцов множества $\mathcal{B}_0$ и таких k строк матрицы A , что $\det B_0 \neq 0$, и через b_0 — вектор, составленный из k соответствующих компонент вектора b . Тогда вектор z_0 , составленный из k положительных компонент вектора x_0 , является решением системы линейных уравнений $B_0 z = b_0$, так что $z_0 = B_0^{-1} b_0$. Но $\det B_0 = \pm 1$ в силу унимодулярности матрицы A , и, следовательно, B_0^{-1} — целочисленная матрица: все элементы матрицы B_0^{-1} также равны 1, -1 или 0. — Упражнение 8.2. Поэтому z_0 — целочисленный вектор, а значит, целочислен и вектор x_0 . ■

Следствие. *Если в канонической задаче целочисленного линейного программирования (1) матрица A унимодулярна и вектор b целочислен, то оптимальный план задачи $(\mathcal{D}, f)$, полученный любым методом гл. 4, является решением задачи $(\mathcal{D}_u, f)$.*

Это следует из утверждения 3° предыдущего пункта. ■

1.4. Условие унимодулярности матрицы. Установим теперь, каким образом осуществлять проверку унимодулярности матрицы A . При больших m и n непосредственная проверка связана с вычислением огромного числа определителей и практически неосуществима. Приведем некоторые достаточные условия унимодулярности матрицы A , в ряде случаев более эффективные. Справедливо следующее утверждение.

Пусть $A = (a[i, j]) - m \times n$ матрица, элементы которой равны 1, -1 или 0. Предположим, что:

а) каждый столбец матрицы A содержит не более двух ненулевых элементов;

б) множество строк матрицы A можно разбить на два подмножества $\mathcal{A}_1 = \{a_i, i \in I_1\}$ и $\mathcal{A}_2 = \{a_i, i \in I_2\}$ ($I_1 \cup I_2 = 1:m$, $I_1 \cap I_2 = \emptyset$) так, что ненулевые элементы каждого столбца, имеющие одинаковые знаки, находятся в строках, принадлежащих разным множествам $\mathcal{A}_1$ и $\mathcal{A}_2$, а ненулевые элементы, имеющие разные знаки, — в строках одного из подмножеств $\mathcal{A}_1$ или $\mathcal{A}_2$.

Тогда матрица A унимодулярна.

Доказательство проводится индукцией по k — порядку минора матрицы A . Очевидно, что всякий минор матрицы A порядка 1 равен 1, -1 или 0. Предположим, что всякий минор матрицы A порядка k равен 1, -1 или 0, и пусть A' — квадратная подматрица матрицы A порядка $k+1$. Определитель матрицы A' очевидно равен нулю, если все элементы некоторого столбца этой матрицы равны нулю. Если в матрице A' есть столбец, содержащий ровно один ненулевой элемент, то, представив $\det A'$ в виде произведения этого элемента на его алгебраическое дополнение, получим, что $\det A'$ равен 1, -1 или 0 в силу предположения индукции. Если, наконец, каждый столбец матрицы A' содержит ровно два ненулевых элемента, то, как следует из определения множеств $\mathcal{A}_1$ и $\mathcal{A}_2$, возможны два случая: либо сумма строк матрицы A' , индексы которых входят в множество I_1 , равна сумме строк этой матрицы, индексы которых входят в множество I_2 , либо, если одно из этих множеств пустое, сумма строк другого — строка, все элементы которой равны нулю. И в том, и в другом случае $\det A' = 0$. ■

Отметим, что для матриц, удовлетворяющих условию а), условие б) является необходимым условием унимодулярности. Доказательство этого утверждения, принадлежащего Д. Гейлу, приведено в [30, с. 353, 354].

1.5. Целочисленность решений задач транспортного типа. Особенности структуры матрицы ограничений некоторых задач линейного программирования позволяют эффективно использовать критерий предыдущего пункта для проверки унимодулярности таких матриц. В первую очередь это относится к задачам транспортного типа.

1°. Матрица ограничений транспортной задачи в матричной постановке (см. п. 2.3.2) выписана в п. 3.1.11, 3°. Она очевидно удовлетворяет условию критерия предыдущего пункта и, следовательно, унимодулярна. Поэтому множество допустимых планов транспортной задачи в матричной постановке целочисленно, если только запасы $(a[i], i \in 1:m)$ и потребности $(b[j], j \in 1:n)$ — целые числа.

2°. Задачу о назначениях (см. п. 2.3.8) легко переформулировать как задачу нахождения целочисленного оптимального плана транспортной задачи в матричной постановке. — Упражнение 8.3. Применение алгоритма § 5 гл. 6 позволит найти такой оптимальный план транспортной задачи, который и будет оптимальным планом задачи о назначениях. Отметим, что транспортная задача в матричной постановке, отвечающая задаче о назначениях, обладает некоторой спецификой: $m = n$; $a[i] = b[j] = 1$, $i \in 1:n$, $j \in 1:n$. Учет этой специфики может облегчить решение задачи.

3°. Унимодулярной является и матрица ограничений транспортной задачи в сетевой постановке (см. пп. 2.3.3 или 6.3.6). Действительно, в п. 6.4.1 показано, что такую задачу можно записать как каноническую задачу линейного программирования с $m \times n$ матрицей ограничений A с элементами 1, -1 или 0, в каждом столбце которой два ненулевых элемента: 1 и -1 . Полагаем $I_1 = 1:m$, $I_2 = \emptyset$. Ясно, что для матрицы A выполняются условия утверждения п. 1.4, и, следовательно, A — унимодулярная матрица. ■

1.6. Обсуждение. Методы отсечения. В п. 1.2 установлено, что если множество $\mathcal{D} = \{x \in R^n : Ax = b, x \geq 0\}$ целочисленно, то для решения задачи $(\mathcal{D}_c, f)$ можно использовать алгоритмы решения задачи $(\mathcal{D}, f)$. Однако если множество $\mathcal{D}$ не является целочисленным, то оптимальный план $\bar{x}$ задачи $(\mathcal{D}, f)$ может не быть целочисленным. К сожалению, во многих случаях нельзя ограничиться и «приближенным решением» $[\bar{x}] = ([\bar{x}[1]], \dots, [\bar{x}[n]])$, где $[\bar{x}[j]]$ — целая часть числа $\bar{x}[j]$: мы не можем гарантировать допустимость вектора $[\bar{x}]$. Более того, даже если $[\bar{x}] \in \mathcal{D}_c$, $f([\bar{x}])$ может значительно отличаться от оптимального значения f на $\mathcal{D}_c$.

Один из подходов, возможных в такой ситуации, связан с аппроксимациями множества $v(\mathcal{D}_c)$ — выпуклой оболочки множества $\mathcal{D}_c$ — многогранными выпуклыми множествами, содержащими $\mathcal{D}_c$ и содержащимися в $\mathcal{D}$. Такие аппроксимации достигаются отсечением частей множества $\mathcal{D}$, не содержащих целых точек, с помощью дополнительных линейных ограничений. С реализациями такого подхода, осуществленными Р. Гомори, читатель может ознакомиться в [27, 44 или 45].

§ 2. Методы ветвей и границ

Задача выбора элемента конечного множества $\mathcal{D}$, на котором числовая функция f , определенная на этом множестве, принимает наибольшее или наименьшее значение, с принципиальной точки зрения достаточно тривиальна. В самом деле, для решения такой задачи достаточно перенумеровать элементы множества $\mathcal{D}$ и затем последовательно вычислять значения функции в этих точках, запоминая на каждом шаге тот из уже

проверенных элементов, для которого значение функции было наибольшим (наименьшим). Однако достаточно часто, особенно в экономических задачах, такая процедура практически нереализуема из-за большого числа элементов множества $\mathcal{D}$.

Другой подход к решению рассматриваемой задачи связан с такой организацией перебора, при которой удается отбрасывать не отдельные точки, а достаточно большие подмножества, состоящие из заведомо неоптимальных элементов множества $\mathcal{D}$. Именно такая идея и положена в основу алгоритмов, которые принято относить к методам ветвей и границ.

Сразу же необходимо отметить нерегулярность таких методов: их реализация для конкретных классов задач связана с анализом этих задач и, что существенно, с изобретательностью авторов алгоритмов. Кроме того, при решении задач любого достаточно широкого класса специально построенным для них алгоритмом метода ветвей и границ в этом классе может встретиться задача, решение которой сведется к полному перебору элементов множества $\mathcal{D}$.

2.1. Основные понятия и общие схемы методов ветвей и границ. Алгоритмом дробления конечного множества $\mathcal{D}$ назовем правило, определяющее на первом шаге разбиение множества $\mathcal{D}$ на несколько подмножеств, затем — разбиение каждого из полученных подмножеств и т. д. и действующее до тех пор, пока после конечного числа шагов элементами разбиения не станут отдельные точки множества $\mathcal{D}$.

Применение алгоритма дробления обычно осуществляется следующим образом. Назовем *концевыми множествами первого шага* подмножества, полученные при разбиении множества $\mathcal{D}$. Выберем какое-нибудь из этих множеств и применим к нему алгоритм дробления. Полученные при этом множества вместе с остальными концевыми множествами первого шага образуют *концевые множества второго шага*. Таким образом, концевыми множествами каждого шага являются все элементы очередного разбиения и все остальные, не подвергавшиеся дроблению концевые множества предыдущего шага.

Приведем пример. Пусть $\mathcal{D}$ — множество всех векторов $(x[1], x[2], \dots, x[n])$, у которых $x[j] = 0$ или $x[j] = 1$ при $j \in 1:n$. Рассмотрим следующий алгоритм дробления. На первом шаге множество $\mathcal{D}$ разбивается на два подмножества:

$$\mathcal{D}_0 = \{(0, x[2], x[3], \dots, x[n]), x[j] = 0 \text{ или } x[j] = 1, j \in 2:n\},$$

$$\mathcal{D}_1 = \{(1, x[2], x[3], \dots, x[n]), x[j] = 0 \text{ или } x[j] = 1, j \in 2:n\},$$

подмножество $\mathcal{D}_0$ алгоритм разбивает на два подмножества:

$$\mathcal{D}_{00} = \{(0, 0, x[3], \dots, x[n]), x[j] = 0 \text{ или } x[j] = 1, j \in 3:n\},$$

$$\mathcal{D}_{01} = \{(0, 1, x[3], \dots, x[n]), x[j] = 0 \text{ или } x[j] = 1, j \in 3:n\},$$

а множество $\mathcal{D}_1$ — на подмножества

$\mathcal{D}_{10} = \{(1, 0, x[3], \dots, x[n]), x[j] = 1 \text{ или } x[j] = 0, j \in 3:n\},$

$\mathcal{D}_{11} = \{(1, 1, x[3], \dots, x[n]), x[j] = 1 \text{ или } x[j] = 0, j \in 3:n\}$

и т. д.

Концевыми множествами первого шага дробления являются $\mathcal{D}_0$ и $\mathcal{D}_1$. Если на втором шаге было выбрано подмножество $\mathcal{D}_1$, то концевыми множествами второго шага являются подмножества $\mathcal{D}_0$, $\mathcal{D}_{10}$ и $\mathcal{D}_{11}$. Если после этого на третьем шаге выбирается подмножество $\mathcal{D}_{10}$, то концевыми множествами третьего шага оказываются $\mathcal{D}_0$, $\mathcal{D}_{11}$, $\mathcal{D}_{100}$, $\mathcal{D}_{101}$ и т. п.

Числовую функцию W , определенную на множестве подмножеств множества $\mathcal{D}$, которые могут быть получены в процессе дробления, назовем *системой оценок дробления*. Значение этой функции на подмножестве, полученном в результате дробления, назовем *оценкой этого множества*.

Рассмотрим теперь задачу дискретного программирования $(\mathcal{D}, f)$ с конечным множеством $\mathcal{D}$ допустимых планов. Методом ветвей и границ для решения задачи $(\mathcal{D}, f)$ назовем алгоритм дробления множества $\mathcal{D}$ и связанную с ним систему оценок дробления W , удовлетворяющую следующим условиям.

1) Оценка любого подмножества, полученного при дроблении, не превосходит минимального значения функции f на этом подмножестве — в задаче минимизации (или не меньше максимального значения функции f на этом подмножестве — в задаче максимизации).

2) Для любого множества, полученного при дроблении, его оценка не меньше оценки того множества, разбиением которого оно получено, — в задаче минимизации (или не больше оценки того множества, разбиением которого оно получено, — в задаче максимизации).

3) Оценка множества, состоящего из единственной точки, равна значению функции f в этой точке.

Реализация метода ветвей и границ может осуществляться по следующим двум основным схемам.

Схема полного ветвления.* На первом шаге вычисляется оценка $W(\mathcal{D})$. В соответствии с алгоритмом дробления осуществляется разбиение множества $\mathcal{D}$ и вычисляются оценки концевых множеств первого шага.

На втором шаге из концевых множеств первого шага выбирается множество, имеющее минимальную оценку (если таких несколько, то выбирается любое из них). Алгоритм дробления применяется к выбранному множеству. Вычисляются оценки вновь полученных концевых множеств.

* Приводится для задачи минимизации.

На каждом шаге алгоритма из всех конечных множеств предыдущего шага выбирается множество, имеющее минимальную оценку, осуществляется его дробление и вычисляются оценки вновь полученных конечных множеств.

Вычисления заканчиваются после получения одноэлементного множества, оценка которого не превосходит оценок всех конечных множеств данного шага. Найденный элемент является решением задачи $(\mathcal{D}, f)$. Это очевидным образом следует из свойств 1) — 3). ■

Изменения, которые необходимо внести в схему для случая задачи максимизации, читатель без труда осуществит самостоятельно. — Упражнение 8.4.

Проиллюстрируем схему полного ветвления следующим примером. Предположим, что $W(\mathcal{D}) = 10$ и множество $\mathcal{D}$ разбивается алгоритмом дробления на подмножества $\mathcal{D}_1$, $\mathcal{D}_2$ и $\mathcal{D}_3$; $W(\mathcal{D}_1) = 15$, $W(\mathcal{D}_2) = 10$, $W(\mathcal{D}_3) = 20$. Так как $W(\mathcal{D}_2) < W(\mathcal{D}_1)$ и $W(\mathcal{D}_2) < W(\mathcal{D}_3)$, то на втором шаге дроблению подлежит множество $\mathcal{D}_2$. Предположим, что оно разбивается на подмножества $\mathcal{D}_{21}$ и $\mathcal{D}_{22}$, и пусть $W(\mathcal{D}_{21}) = 17$, $W(\mathcal{D}_{22}) = 25$. Концевыми множествами второго шага являются $\mathcal{D}_1$, $\mathcal{D}_3$, $\mathcal{D}_{21}$ и $\mathcal{D}_{22}$. Среди них наименьшую оценку имеет множество $\mathcal{D}_1$. Следовательно, оно и подлежит дроблению на третьем шаге.

Схема одностороннего ветвления.* В отличие от схемы полного ветвления здесь на каждом шаге выбор конечного множества, подлежащего разбиению, производится не из всех конечных множеств предыдущего шага, а только из тех, которые впервые получены на предыдущем шаге дробления. Именно такое применение дробления называем *схемой одностороннего ветвления*. Процесс осуществляется до получения по крайней мере одного одноэлементного подмножества. Минимальная оценка таких одноэлементных подмножеств называется первым *рекордом*. Запоминаются элемент подмножества, имеющий эту оценку, и значение рекорда.

После вычисления рекорда из рассмотрения исключаются все концевые подмножества, имеющие оценки не меньше рекорда.

Среди оставшихся конечных подмножеств выбираем множество, имеющее наименьшую оценку, и подвергаем его дроблению в соответствии со схемой одностороннего ветвления. Однако здесь в отличие от процесса получения первого рекорда на каждом шаге исключаются из дальнейшего рассмотрения концевые множества, имеющие оценки не меньше рекорда. Эта процедура заканчивается появлением нового одноэлементного множества, имеющего оценку меньше рекорда, либо обрывается. В первом случае получаем новый рекорд, который запоминаем

* Приводится для задачи минимизации.

вместе с соответствующим элементом. В дальнейшем для исключения конечных множеств используется новый рекорд. Во втором случае из оставшихся конечных множеств выбираем множество с наименьшей оценкой и осуществляем его дробление в соответствии со схемой одностороннего ветвления.

Вычисления продолжают до тех пор, пока не будут исключены все конечные множества, имеющие оценки не меньше последнего рекорда. Этот рекорд равен минимальному значению целевой функции на $\mathcal{D}$, а соответствующий элемент является решением задачи $(\mathcal{D}, f)$. ■

Проиллюстрируем схему одностороннего ветвления. В схеме полного ветвления на третьем шаге дроблению подлежало множество $\mathcal{D}_1$. В схеме одностороннего ветвления на третьем шаге дроблению подлежит множество $\mathcal{D}_{21}$, так как здесь выбор множества, подлежащего дроблению, осуществляется только среди множеств $\mathcal{D}_{21}$ и $\mathcal{D}_{22}$, вновь полученных на втором шаге.

2.2. Применение метода ветвей и границ для решения задачи коммивояжера (см. п. 2.3.9).

1°. Обозначения и предварительные замечания. Напомним формулировку задачи коммивояжера. Множество допустимых планов $\mathcal{D}$ задачи состоит из всех маршрутов $\pi = (1 \rightarrow i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_{n-1} \rightarrow 1)$, где $i_1, i_2, \dots, i_{n-1}$ — перестановка чисел $2, 3, \dots, n$. Задана матрица $C = (c[i, j], i \in 1:n, j \in 1:n)$, где $c[i, j]$ — «стоимость путешествия» из i -го города в j -й при $i \neq j$, а $c[i, i] = +\infty, i \in 1:n$. Задача состоит в отыскании такого маршрута π , для которого $f(\pi) = c[1, i_1] + \sum_{k=1}^{n-2} c[i_k, i_{k+1}] + c[i_{n-1}, 1]$ — стоимость маршрута π — минимальна.

Пары $(1, i_1), (i_1, i_2), (i_2, i_3), \dots, (i_{n-1}, 1)$ назовем звеньями маршрута π . Пусть $i_1, i_2, \dots, i_k; j_1, j_2, \dots, j_k; i'_1, i'_2, \dots, i'_m; j'_1, j'_2, \dots, j'_m$ — числа из $1:n$. Через

$$\mathcal{D} \{ (i_1, j_1), \dots, (i_k, j_k) \| (i'_1, j'_1), \dots, (i'_m, j'_m) \} \quad (1)$$

обозначим множество всех маршрутов, имеющих звенья $(i_1, j_1), \dots, (i_k, j_k)$ и не содержащих звеньев $(i'_1, j'_1), \dots, (i'_m, j'_m)$. Пары $(i_1, j_1), \dots, (i_k, j_k)$ назовем фиксированными звеньями маршрутов из множества (1), а пары $(i'_1, j'_1), \dots, (i'_m, j'_m)$ — запрещенными звеньями для маршрутов из множества (1).

Примеры. Пусть $n=4$. Множество $\mathcal{D} = \{ (2, 3) \| (1, 4) \}$ состоит из единственного маршрута $(1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1)$. Множество $\mathcal{D} = \{ \emptyset \| (2, 4), (3, 1) \}$ — из маршрутов $(1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1)$, $(1 \rightarrow 3 \rightarrow 4 \rightarrow 2 \rightarrow 1)$ и $(1 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1)$. Множество $\mathcal{D} = \{ (1, 2), (4, 3) \| \emptyset \}$ включает единственный маршрут $(1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 1)$. Множество $\mathcal{D} = \{ (1, 2), (2, 3) \| (3, 4) \}$ — пустое.

Система $(i_1, j_1), (i_2, j_2), \dots, (i_k, j_k)$ звеньев маршрута образует цепочку, если эти звенья можно переупорядочить так

чтобы второй элемент предыдущего звена совпадал с первым элементом звена, следующего за ним. Если несколько звеньев упорядочено в цепочку, то первый элемент первого звена назовем *началом цепочки*, а второй элемент последнего звена — *концом цепочки*. Если начало цепочки совпадает с ее концом, то цепочка называется *циклом*. Ясно, что цепочка является циклом в том и только том случае, если она совпадает со всем маршрутом.

Очевидно, что $\mathcal{D}\{\emptyset\|\emptyset\} = \mathcal{D}$. Очевидно также, что для любой пары (i, j) чисел из $1:n$ множества

$$\mathcal{D}\{(i_1, j_1), \dots, (i_k, j_k), (i, j)\|(i'_1, j'_1), \dots, (i'_m, j'_m)\},$$

$$\mathcal{D}\{(i_1, j_1), \dots, (i_k, j_k)\|(i'_1, j'_1), \dots, (i'_m, j'_m), (i, j)\}$$

не пересекаются и их объединение есть множество (1).

2°. *Приведение матрицы*. При описании алгоритма нам понадобится следующая операция. Пусть $A = (a[i, j])$ — квадратная матрица порядка m . Положим

$$\begin{aligned} u[i] &= \min_{j \in 1:m} a[i, j], \quad i \in 1:m; \\ a'[i, j] &= a[i, j] - u[i], \quad i \in 1:m, \\ v[j] &= \min_{i \in 1:m} a'[i, j], \quad j \in 1:m; \\ \bar{a}[i, j] &= a'[i, j] - v[j], \quad i, j \in 1:m. \end{aligned} \quad (2)$$

Операцию построения матрицы $\bar{A}$ с элементами $\bar{a}[i, j]$, $i, j \in 1:m$, назовем *приведением* матрицы A .

Из формул (2) следует, что $\bar{a}[i, j] = a[i, j] - u[i] - v[j]$ и $\bar{a}[i, j] \geq 0$ для любых i и j из $1:m$. Кроме того, очевидно, что в каждой строке и каждом столбце матрицы $\bar{A}$ есть нулевой элемент. Число $d(A) = \sum_{i=1}^m u[i] + \sum_{j=1}^m v[j]$ назовем *числом приведения* матрицы A .

Пусть $p = (i_1, i_2, \dots, i_m)$ — перестановка чисел $1, 2, \dots, m$ и $g_A(p) = \sum_{k=1}^m a[k, i_k]$, $g_{\bar{A}}(p) = \sum_{k=1}^m \bar{a}[k, i_k]$. Тогда

$$g_A(p) = d(A) + g_{\bar{A}}(p). \quad (3)$$

Действительно, $g_A(p) = \sum_{k=1}^m (\bar{a}[k, i_k] + u[k] + v[i_k]) = g_{\bar{A}}(p) + \sum_{k=1}^m u[k] + \sum_{k=1}^m v[i_k] = g_{\bar{A}}(p) + d(A)$, поскольку $p = (i_1, i_2, \dots, i_m)$ — перестановка чисел $1, 2, \dots, m$. ■

3°. *Алгоритм дробления и вычисление оценок (на первом шаге)*. Применение метода ветвей и границ к решению задачи коммивояжера состоит в построении алгоритма дробления и вычисления оценок концевых подмножеств каждого шага, удовлетворяющих свойствам 1)–3) п. 2.1. Правила дробления концевого множества и вычисления оценок его подмножеств, получен-

ных при дроблении, будут определяться некоторой матрицей, связанной с этим конечным множеством. Поэтому в описании алгоритма мы не только укажем подмножества, на которые разбивается конечное множество, и их оценки, но также свяжем с каждым из этих подмножеств определенную матрицу.

Для удобства читателя отдельно опишем действие алгоритма на первом шаге и лишь затем — на произвольном.

С множеством $\mathcal{D} = \mathcal{D}(\emptyset \| \emptyset)$ свяжем матрицу $\bar{C} = C(\emptyset \| \emptyset)$, полученную приведением матрицы стоимостей C . Оценку $W(\mathcal{D})$ множества $\mathcal{D}$ положим равной числу приведения матрицы C : $W(\mathcal{D}) = d(C)$ (см. 2°). Тогда для любого маршрута $\pi = (1 \rightarrow i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_{n-1} \rightarrow 1) \in \mathcal{D}$

$$f(\pi) = W(\mathcal{D}) + \bar{c}[1, i_1] + \sum_{k=1}^{n-2} \bar{c}[i_k, i_{k+1}] + \bar{c}[i_{n-1}, 1] \quad (4)$$

(см. (3)) и $W(\mathcal{D}) \leq f(\pi)$, поскольку $\bar{C} \geq 0$.

На первом шаге разобьем множество $\mathcal{D}$ на два подмножества. Для этого в первой строке матрицы $\bar{C}$ найдем первый по порядку нулевой элемент. Пусть он расположен в столбце с номером j_1 . Множество $\mathcal{D}$ разобьем на подмножества $\mathcal{D}_1^1 = \mathcal{D}\{(1, j_1) \| \emptyset\}$ и $\mathcal{D}_1^0 = \mathcal{D}\{\emptyset \| (1, j_1)\}$. Они являются конечными множествами первого шага. Если $n > 2$, то $\mathcal{D}_1^1 = \mathcal{D}\{(1, j_1) \| (j_1, 1)\}$, так как в этом случае никакая цепочка, состоящая из двух звеньев, не может быть циклом.

Свяжем с множествами $\mathcal{D}_1^1$ и $\mathcal{D}_1^0$ матрицы $C_1^1 = C(1, j_1) \| (j_1, 1)$ и $C_1^0 = C\{\emptyset \| (1, j_1)\}$ и определим оценки множеств $\mathcal{D}_1^1$ и $\mathcal{D}_1^0$. Чтобы получить матрицу C_1^1 , заменим элемент $\bar{c}[j_1, 1]$ матрицы $\bar{C}$ на $+\infty$, вычеркнем первую строку и столбец с номером j_1 и приведем полученную матрицу. Оценку $W(\mathcal{D}_1^1)$ положим равной сумме $W(\mathcal{D})$ и числа приведения. Чтобы получить матрицу C_1^0 , заменим элемент $\bar{c}[1, j_1]$ матрицы $\bar{C}$ на $+\infty$ и приведем полученную матрицу. Оценку $W(\mathcal{D}_1^0)$ положим равной сумме $W(\mathcal{D})$ и числа приведения.

Отметим некоторые очевидные свойства матриц, связанных с конечными множествами первого шага и оценок этих множеств.

1) Строки матрицы C_1^1 занумерованы индексами $i \in 2:n$, а столбцы — индексами $j \in 1:n \setminus \{j_1\}$; строки матрицы C_1^0 занумерованы индексами $i \in 1:n$, а столбцы — индексами $j \in 1:n$. В матрице C_1^0 число элементов, равных $+\infty$, на единицу больше числа таких элементов в матрице $\bar{C}$.

2) Элементы матриц C_1^1 и C_1^0 неотрицательны, и по крайней мере один элемент каждой строки и каждого столбца этих матриц равен нулю.

3) Если i входит в множество индексов строк, а j — в мно-

жество индексов столбцов матрицы C_1^1 (матрицы C_1^0) и в множестве $\mathcal{D}_1^1$ (множестве $\mathcal{D}_1^0$) не существует ни одного маршрута, содержащего звено (i, j) , то элемент указанной матрицы с индексами i и j равен $+\infty$.

4) Стоимость любого маршрута из множества $\mathcal{D}_1^1$ (множества $\mathcal{D}_1^0$) равна оценке этого множества, сложенной с суммой элементов матрицы C_1^1 (матрицы C_1^0), отвечающих нефиксированным звеньям этого маршрута. Формальная проверка этого свойства осуществляется аналогично доказательству формулы (3) (или (4)) с учетом того, что $\bar{c}[1, j_1] = 0$.

$$5) \quad W(\mathcal{D}_1^1) \leq \min_{\pi \in \mathcal{D}_1^1} f(\pi); \quad W(\mathcal{D}_1^0) \leq \min_{\pi \in \mathcal{D}_1^0} f(\pi).$$

$$6) \quad W(\mathcal{D}_1^1) \geq W(\mathcal{D}); \quad W(\mathcal{D}_1^0) \geq W(\mathcal{D}).$$

4°. Алгоритм дробления произвольного конечного множества и вычисление оценок. Пусть

$$\mathcal{D}_m^k = \mathcal{D}\{(i_1, j_1), \dots, (i_k, j_k) \parallel (i'_1, j'_1), \dots, (i'_m, j'_m)\}$$

— некоторое конечное множество, $W(\mathcal{D}_m^k)$ — его оценка и

$$C_m^k = C\{(i_1, j_1), \dots, (i_k, j_k) \parallel (i'_1, j'_1), \dots, (i'_m, j'_m)\}$$

— матрица, связанная с множеством $\mathcal{D}_m^k$. Введем следующие предположения индукции.

1') C_m^k — квадратная матрица порядка $n - k$. Ее строки занумерованы индексами $i \in 1:n \setminus \{i_1, i_2, \dots, i_k\}$, столбцы — индексами $j \in 1:n \setminus \{j_1, j_2, \dots, j_k\}$.

2') Элементы матрицы C_m^k неотрицательны, и по крайней мере один элемент каждой строки и каждого столбца равен нулю.

3') Если i входит в множество индексов строк, а j — в множество индексов столбцов матрицы C_m^k и в множестве $\mathcal{D}_m^k$ нет ни одного маршрута, содержащего звено (i, j) , то элемент матрицы C_m^k с индексами i и j равен $+\infty$.

4') Стоимость любого маршрута из множества $\mathcal{D}_m^k$ равна оценке этого множества, сложенной с суммой элементов матрицы C_m^k , отвечающих нефиксированным звеньям этого маршрута.

$$5') \quad W(\mathcal{D}_m^k) \leq \min_{\pi \in \mathcal{D}_m^k} f(\pi).$$

Опишем правило разбиения конечного множества $\mathcal{D}_m^k$. В первой строке матрицы C_m^k найдем первый по порядку нулевой элемент. Пусть он расположен в столбце с номером j_{k+1} , а первая

строка имеет номер i_{k+1} . Множество $\mathcal{D}_m^k$ разобьем на два подмножества:

$$\mathcal{D}_{m+1}^{k+1} = \mathcal{D}\{(i_1, j_1), \dots, (i_k, j_k), (i_{k+1}, j_{k+1}) \parallel (i'_1, j'_1), \dots, (i'_m, j'_m)\},$$

$$\mathcal{D}_{m+1}^k = \mathcal{D}\{(i_1, j_1), \dots, (i_k, j_k) \parallel (i'_1, j'_1), \dots, (i'_m, j'_m), (i_{k+1}, j_{k+1})\}.$$

Рассмотрим цепочку максимальной длины, составленную из звеньев множества $\{(i_1, j_1), \dots, (i_k, j_k), (i_{k+1}, j_{k+1})\}$ и содержащую звено (i_{k+1}, j_{k+1}) . Такая цепочка может быть получена добавлением звена (i_{k+1}, j_{k+1}) справа или слева к некоторой цепочке, составленной из звеньев множества $\{(i_1, j_1), \dots, (i_k, j_k)\}$, либо стыковкой двух цепочек звеном (i_{k+1}, j_{k+1}) . Обозначим концы построенной цепочки через j'_{m+1} и i'_{m+1} .

Пример. Пусть $n = 5$, $\mathcal{D}_2^2 = \mathcal{D}\{(1, 2), (3, 4) \parallel (2, 1), (4, 3)\}$ и $(i_3, j_3) = (5, 1)$. Тогда из звеньев множества $\{(1, 2), (3, 4), (5, 1)\}$ составляется цепочка $5 \rightarrow 1 \rightarrow 2$, $j'_3 = 5$, $i'_3 = 2$ и

$$\begin{aligned} & \mathcal{D}\{(1, 2), (3, 4), (5, 1) \parallel (2, 1), (4, 3)\} = \\ & = \mathcal{D}\{(1, 2), (3, 4), (5, 1) \parallel (2, 1), (4, 3), (2, 5)\}. \end{aligned}$$

Если $(i_3, j_3) = (2, 3)$, то составляется цепочка $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$, $j'_3 = 1$, $i'_3 = 4$ и

$$\begin{aligned} & \mathcal{D}\{(1, 2), (3, 4), (2, 3) \parallel (2, 1), (4, 3)\} = \\ & = \mathcal{D}\{(1, 2), (3, 4), (2, 3) \parallel (2, 1), (4, 3), (4, 1)\}. \end{aligned}$$

Если $n > k + 2$, то никакая цепочка, длина которой не превосходит $k + 2$, не может быть циклом. Следовательно,

$$\begin{aligned} \mathcal{D}_{m+1}^{k+1} = \mathcal{D}\{(i_1, j_1), \dots, (i_k, j_k), (i_{k+1}, j_{k+1}) \parallel (i'_1, j'_1), \dots, \\ \dots, (i'_m, j'_m), (i'_{m+1}, j'_{m+1})\}. \end{aligned}$$

Множества $\mathcal{D}_{m+1}^{k+1}$, $\mathcal{D}_{m+1}^k$ — новые концевые множества. Свяжем с ними матрицы C_{m+1}^{k+1} , C_{m+1}^k и оценки $W(\mathcal{D}_{m+1}^{k+1})$, $W(\mathcal{D}_{m+1}^k)$. Чтобы получить матрицу C_{m+1}^{k+1} , заменим элемент матрицы C_m^k с индексами i'_{m+1} , j'_{m+1} на $+\infty$, вычеркнем строку с номером i_{k+1} и столбец с номером j_{k+1} и приведем полученную матрицу. Оценку $W(\mathcal{D}_{m+1}^{k+1})$ положим равной сумме $W(\mathcal{D}_m^k)$ и числа приведения. Чтобы получить матрицу C_{m+1}^k , заменим элемент матрицы C_m^k с индексами i_{k+1} , j_{k+1} на $+\infty$ и приведем полученную матрицу. Оценку $W(\mathcal{D}_{m+1}^k)$ положим равной сумме $W(\mathcal{D}_m^k)$ и числа приведения.

Из построения следует, что для матриц C_{m+1}^{k+1} , C_{m+1}^k и оценок $W(\mathcal{D}_{m+1}^{k+1})$, $W(\mathcal{D}_{m+1}^k)$ выполняются свойства 1') — 5'), если в формулировках этих свойств заменить k на $k + 1$ и m на

$m+1$ в случае множества $\mathcal{D}_{m+1}^{k+1}$ и m на $m+1$ — в случае множества $\mathcal{D}_{m+1}^k$; кроме того:

1'') в матрице C_{m+1}^k число элементов, равных $+\infty$, на единицу больше числа таких элементов в матрице C_m^k .

$$6') \quad W(\mathcal{D}_{m+1}^{k+1}) \geq W(\mathcal{D}_m^k); \quad W(\mathcal{D}_{m+1}^k) \geq W(\mathcal{D}_m^k).$$

5°. *Итоги.* Решение задачи коммивояжера методом ветвей и границ может осуществляться с использованием описанного в 3° и 4° алгоритма дробления и вычисления оценок в соответствии со схемами полного или одностороннего ветвления (п. 2.1). Получение одноэлементного множества за конечное число шагов обеспечивается свойствами 1'), 1''), 3'), а свойства оценок 1)–3) п. 2.1 выполняются в силу свойств 4')–6').

Для ускорения сходимости обычно используется следующий прием: пару (i_{k+1}, j_{k+1}) выбирают так, чтобы число приведения матрицы, полученной заменой в матрице C_m^k соответствующего нулевого элемента на $+\infty$, было максимальным. Такой прием может обеспечить более быстрое возрастание оценок $W(\mathcal{D}_{m+1}^k)$ и отсеечение неперспективных конечных множеств.

6°. *Пример.* Проиллюстрируем алгоритм дробления и вычисление оценок на примере задачи коммивояжера с матрицей стоимостей C порядка 4. В такой задаче множество $\mathcal{D}$ состоит из шести маршрутов:

$$\pi_1 = (1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1), \quad \pi_2 = (1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 1),$$

$$\pi_3 = (1 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1), \quad \pi_4 = (1 \rightarrow 3 \rightarrow 4 \rightarrow 2 \rightarrow 1),$$

$$\pi_5 = (1 \rightarrow 4 \rightarrow 2 \rightarrow 3 \rightarrow 1), \quad \pi_6 = (1 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1).$$

Пусть
$$C = \begin{bmatrix} \infty & 2 & 4 & 8 \\ 6 & \infty & 6 & 7 \\ 5 & 2 & \infty & 5 \\ 3 & 2 & 7 & \infty \end{bmatrix}.$$

Число приведения этой матрицы равно 13 ($u[1]=2$, $u[2]=6$, $u[3]=2$, $u[4]=2$, $v[1]=v[2]=v[3]=0$, $v[4]=1$), начальная оценка $W(\mathcal{D})=13$, приведенная матрица

$$\bar{C} = \begin{bmatrix} \infty & 0 & 2 & 5 \\ 0 & \infty & 0 & 0 \\ 3 & 0 & \infty & 2 \\ 1 & 0 & 5 & \infty \end{bmatrix}.$$

На первом шаге множество $\mathcal{D}$ дробится на два множества:

$$\mathcal{D}_1^1 = \mathcal{D} \{(1, 2) \parallel (2, 1)\} = \{\pi_1, \pi_2\} \text{ и } \mathcal{D}_1^0 = \mathcal{D} \{\emptyset \parallel (1, 2)\} = \{\pi_3, \pi_4, \pi_5, \pi_6\}$$

Это концевые множества первого шага. Их оценки соответственно $W(\mathcal{D}_1^1) = 16$, $W(\mathcal{D}_1^0) = 15$, матрицы

$$C_1^1 = 3 \begin{bmatrix} & 1 & 3 & 4 \\ 2 & \infty & 0 & 0 \\ 1 & 1 & \infty & 0 \\ 4 & 0 & 4 & \infty \end{bmatrix},$$

$$C_1^0 = 3 \begin{bmatrix} & 1 & 2 & 3 & 4 \\ 1 & \infty & \infty & 0 & 3 \\ 2 & 0 & \infty & 0 & 0 \\ 3 & 3 & 0 & \infty & 0 \\ 4 & 1 & 0 & 5 & \infty \end{bmatrix}.$$

На втором шаге множество $\mathcal{D}_1^0$ дробится на множества $\mathcal{D}_2^1 = \{(1, 3) \parallel (1, 2), (3, 1)\} = \{\pi_3, \pi_4\}$ и $\mathcal{D}_2^0 = \{\emptyset \parallel (1, 2), (1, 3)\} = \{\pi_5, \pi_6\}$. Их оценки $W(\mathcal{D}_2^1) = 15$, $W(\mathcal{D}_2^0) = 18$. Концевыми множествами второго шага являются $\mathcal{D}_1^1$, $\mathcal{D}_2^1$ и $\mathcal{D}_2^0$ (рис. 33).

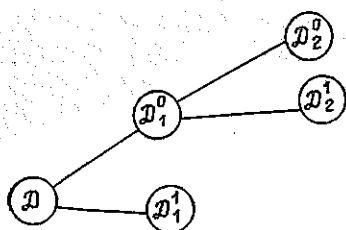


Рис. 33.

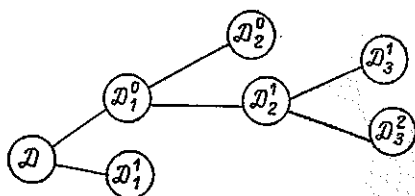


Рис. 34.

На третьем шаге множество $\mathcal{D}_2^1$, имеющее минимальную оценку среди концевых множеств второго шага, дробим на множества

$$\mathcal{D}_3^2 = \{(2, 1), (1, 3) \parallel (1, 2), (3, 1), (3, 2)\} = \{\pi_4\},$$

$$\mathcal{D}_3^1 = \{(1, 3) \parallel (1, 2), (3, 1), (2, 1)\} = \{\pi_3\}$$

(рис. 34). Это одноточечные множества, их оценки $W(\mathcal{D}_3^2) = 17$ и $W(\mathcal{D}_3^1) = 16$. Так как последняя оценка не превосходит оценок остальных концевых множеств третьего шага, то π_3 — оптимальный маршрут, $f(\pi_3) = 16$. Заметим, что развитие ветви $\mathcal{D} \rightarrow \mathcal{D}_1^1$ приведет к получению другого оптимального маршрута π_1 , $f(\pi_1) = 16$.

УПРАЖНЕНИЯ

8.1. См. с. 269. 8.2. См. с. 270. 8.3. См. с. 272. 8.4. См. с. 275.

Глава 9

МЕТОДЫ РЕШЕНИЯ НЕЛИНЕЙНЫХ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

В настоящей главе рассматриваются некоторые итеративные методы решения задач выпуклого программирования. Общие идеи градиентного спуска, на которых они основаны, восходят к Ньютону и Коши и находят достаточно широкую сферу приложений в вычислительной математике. Здесь мы сознательно ограничиваем область приложения таких методов классом задач выпуклого программирования, так как изучение именно этого класса экстремальных задач составляет основное содержание настоящего курса математического программирования. Кроме того, эффективность итеративных методов для произвольных экстремальных задач значительно ниже их эффективности для задач выпуклого программирования, поскольку в первом случае они позволяют находить, вообще говоря, лишь отдельные точки локального экстремума.

При оценке итеративных методов и алгоритмов решения экстремальных задач большое значение приобретают такие факторы, как возможность практической реализации соответствующих вычислительных процессов на ЭВМ и эффективность вычислений. Первый из этих факторов связан с внутренними характеристиками задачи $(\mathcal{D}, f)$: конкретными свойствами и формой задания функции f и функций, определяющих множество $\mathcal{D}$. Второй связан с внешними обстоятельствами: требуемой точностью расчетов и мощностью ЭВМ, имеющейся в распоряжении исследователя. Эти вопросы в дальнейшем не обсуждаются: нас будет интересовать лишь принципиальная реализуемость методов и алгоритмов. Мы не претендуем и на полноту обзора итеративных методов и алгоритмов решения задач выпуклого программирования. Читателю, специально интересующемуся такими вопросами, следует обратиться к соответствующей литературе (см., например, [13, 16, 35, 43]).

§ 1. Общие схемы итеративных методов

1.1. Предварительные замечания. В этой главе рассматриваются методы, приводящие к построению, вообще говоря, бесконечной последовательности точек $x_1, x_2, \dots$. Предполагается, что если уже построены точки $x_1, x_2, \dots, x_k$, то построение следующей за ними точки x_{k+1} осуществляется по определенному правилу, которое будем называть *коррекцией*, или *корректирующим отображением*, и обозначать через сог . Предполагается также, что задана некоторая *оценочная* функция $P: R^n \rightarrow R$, позволяющая для любых двух точек из R^n определить, в какой мере одна из них «лучше» другой. Чаще всего в качестве оценочной функции выбирают целевую функцию f экстремальной задачи $(\mathcal{D}, f)$. В дальнейшем будем рассматривать только такие итеративные методы решения экстремальных задач $(\mathcal{D}, f): f(x) \rightarrow \max, x \in \mathcal{D}$, в которых $P(\text{сог } x) \geq P(x)$ для любого $x \in \mathcal{D}$. В ряде случаев оказывается нецелесообразным ограничить поиск только множеством оптимальных планов. Поэтому в $\mathcal{D}$ выделяют подмножество π приемлемых точек, обычно содержащее множество решений задачи $(\mathcal{D}, f)$. Точка, не принадлежащая множеству π , называется неприемлемой. Построение последовательности не может быть прервано в неприемлемой точке. Поэтому предполагается возможность эффективной проверки принадлежности точки $x \in \mathcal{D}$ множеству π .

Перейдем теперь к формальному описанию некоторых схем итеративных методов решения экстремальных задач. Еще раз отметим, что в этом параграфе (и во всей главе) экстремальные задачи формулируются как задачи максимизации.

1.2. Принципиальная схема метода. Пусть $(\mathcal{D}, f)$ — экстремальная задача, π — множество приемлемых точек в $\mathcal{D}$, $\text{сог}: \mathcal{D} \rightarrow \mathcal{D}$ — корректирующее отображение и $P: R^n \rightarrow R$ — оценочная функция. Построение последовательности $x_1, x_2, \dots$ начинается с некоторого допустимого плана $x_1 \in \mathcal{D}$. После построения точек $x_1, x_2, \dots, x_k$ вычисляется $\text{сог } x_k \in \mathcal{D}$ и $P(\text{сог } x_k)$. Если $P(\text{сог } x_k) > P(x_k)$, полагаем $x_{k+1} = \text{сог } x_k$, вычисляем $\text{сог } x_{k+1}$, $P(\text{сог } x_{k+1})$ и т. д. Если $P(\text{сог } x_k) = P(x_k)$, то построение последовательности закончено. В этом случае полагаем $\bar{x} = x_k$.

Теорема. Если множество $\mathcal{D}$ замкнуто, функция P непрерывна в каждой точке множества $\mathcal{D} \setminus \pi$ или ограничена сверху на $\mathcal{D}$ и для каждой точки $x \in \mathcal{D} \setminus \pi$ существуют такие числа $\varepsilon > 0$ и $\delta > 0$, что $P(\text{сог } z) - P(z) \geq \delta$ для любого $z \in \mathcal{D} \cap \overline{\mathcal{D}}(x, \varepsilon)$ ($\overline{\mathcal{D}}(x, \varepsilon)$ — замкнутый шар с радиусом ε и центром в точке x), то последовательность $x_1, x_2, \dots$ либо конечна и ее последний элемент $\bar{x}$ принадлежит π , либо бесконечна и каждая ее сходящаяся подпоследовательность сходится к точке множества π .

Доказательство. Из условия теоремы следует, что если $x_k \in \mathcal{D} \setminus \pi$, то $P(\text{сог } x_k) > P(x_k)$. Поэтому если $P(\text{сог } x_k) = P(x_k)$ при некотором k , то $x_k \in \pi$.

Пусть теперь последовательность $x_1, x_2, \dots$ бесконечна и $\{x_k\}_{k \in K}$ — ее подпоследовательность, сходящаяся к точке $x_0 \in \mathcal{D}$ ($\mathcal{D}$ — замкнутое множество). Предположим, что $x_0 \in \mathcal{D} \setminus \pi$. Покажем, что это предположение приводит к противоречию. По условию теоремы найдутся такие числа $\varepsilon > 0, \delta > 0$, что $P(\text{co}g z) - P(z) \geq \delta$ для любого $z \in \mathcal{D} \cap \overline{\mathcal{D}}(x_0, \varepsilon)$. Так как $x_k \rightarrow x_0$ при $k \rightarrow \infty, k \in K$, то найдется такое натуральное число k' , что $x_k \in \mathcal{D} \cap \overline{\mathcal{D}}(x_0, \varepsilon)$ при любом $k \geq k', k \in K$. Поэтому $P(\text{co}g x_k) - P(x_k) \geq \delta$ при любом $k \geq k', k \in K$.

Последовательность $\{P(x_k)\}_{k \in K}$ монотонно возрастает и, следовательно, имеет предел, равный $P(x_0)$, если функция P непрерывна в точке x_0 , или некоторому числу α , если P ограничена на $\mathcal{D}$. И в том, и в другом случаях существует такое натуральное число k'' , что $|P(x_k) - P(x_m)| < \delta/2$ для любых $k \geq k'', k \in K, m \geq k'', m \in K$. Пусть теперь $k_1 \geq \max\{k', k''\}, k_1 \in K$, и k_2 — следующий за k_1 номер члена последовательности $\{x_k\}_{k \in K}$. Тогда $P(x_{k_2}) - P(x_{k_1}) < \delta/2$. Вместе с тем

$$P(x_{k_2}) - P(x_{k_1}) = \sum_{i=k_1}^{k_2-1} [P(x_{i+1}) - P(x_i)] \geq P(x_{k_2}) - P(x_{k_1}) \geq \delta. \blacksquare$$

1.3. Приближенное вычисление приемлемой точки. В приложениях оказывается полезным следующее утверждение.

Пусть $\sigma > 0, \pi(\sigma) = \pi + \mathcal{D}(0, \sigma)$ ($\pi(\sigma)$ — σ -окрестность множества π) и последовательность $\{x_k\}_{k=1}^{\infty}$ содержит сходящуюся подпоследовательность. При выполнении условий теоремы п. 1.2 в последовательности $x_1, x_2, \dots$ через конечное число шагов будет построена точка, принадлежащая множеству $\pi(\sigma)$.

Действительно, если $\{x_k\}_{k \in K}$ — подпоследовательность последовательности $\{x_k\}_{k=1}^{\infty}$, сходящаяся в силу теоремы п. 1.2 к точке $x_0 \in \pi$, то найдется такой номер $k_0 \in K$, что $\|x_k - x_0\| < \sigma$ для любого $k \geq k_0, k \in K$, и, следовательно, $x_{k_0} \in \pi(\sigma)$. $\blacksquare$

Нетрудно понять смысл доказанного здесь утверждения: если вычисление приемлемой точки достаточно осуществить с заданной степенью точности σ , то применение схемы п. 1.2 позволит закончить вычисления после конечного числа итераций.

1.4. Неоднозначные коррекции. Некоторая модификация описанной в п. 1.2 схемы получается в случае, когда корректирующее отображение каждой точке $x \in \mathcal{D}$ ставит в соответствие подмножество $\text{Co}g x \subset \mathcal{D}$, а выбор элемента $\text{co}g x \in \text{Co}g x$ осуществляется произвольно, например при помощи некоторого случайного механизма.

Если множество $\mathcal{D}$ замкнуто, функция P непрерывна в каждой точке множества $\mathcal{D} \setminus \pi$ или ограничена сверху на $\mathcal{D}$ и для каждой точки $x \in \mathcal{D} \setminus \pi$ существуют такие числа $\varepsilon > 0$ и $\delta > 0$, что $P(z') - P(z) \geq \delta$ для любого $z \in \mathcal{D} \cap \overline{\mathcal{D}}(x, \varepsilon)$ и любого

$z' \in \text{Сог } z$, то последовательность $x_1, x_2, \dots$ либо конечна и ее последний элемент принадлежит π , либо бесконечна и каждая ее сходящаяся подпоследовательность сходится к точке множества π . Очевидно, что утверждение непосредственно следует из теоремы п. 1.2.

1.5. Схема с приближенной коррекцией. В некоторых случаях точное вычисление $\text{сог } x$ может оказаться практически неосуществимым. Поэтому в настоящем пункте будет предложена схема методов, позволяющих строить итеративную последовательность $x_1, x_2, \dots$ по приближенным значениям $\text{сог } x$. Перейдем к описанию схемы.

Пусть $(\mathcal{D}, f)$ — экстремальная задача, π — множество приемлемых точек в $\mathcal{D}$, $\text{сог}: \mathcal{D} \rightarrow \mathcal{D}$ — корректирующее отображение и $P: R^n \rightarrow R$ — оценочная функция. Пусть σ — некоторое положительное число. Построение начинается с точки $x_1 \in \mathcal{D}$. Предположим, что уже построены такие точки $x_1, x_2, \dots, x_k$ и такие неотрицательные целые числа $m_1, m_2, \dots, m_{k-1}$, что $\|x_{i+1} - \text{сог } x_i\| \leq \sigma/2^{m_i} = \sigma_i$, $P(x_{i+1}) - P(x_i) \geq \sigma_i$ для каждого $i \in 1:k-1$. Для построения точки x_{k+1} применим следующую процедуру. Выберем $y_0 \in \mathcal{D} \cap \bar{\mathcal{D}}(\text{сог } x_k, \sigma)$. Если $P(y_0) - P(x_k) \geq \sigma$, то положим $x_{k+1} = y_0$, $m_k = 0$, $\sigma_k = \sigma$. В противном случае выберем $y_1 \in \mathcal{D} \cap \bar{\mathcal{D}}(\text{сог } x_k, \sigma/2)$ и проверим справедливость неравенства $P(y_1) - P(x_k) \geq \sigma/2$. Если это неравенство справедливо, то положим $x_{k+1} = y_1$, $m_k = 1$, $\sigma_k = \sigma/2$. В противном случае продолжим этот процесс. Тогда либо после конечного числа шагов будет найдена такая точка $y_m \in \mathcal{D} \cap \bar{\mathcal{D}}(\text{сог } x_k, \sigma/2^m)$, что $P(y_m) - P(x_k) \geq \sigma/2^m$, и положим $x_{k+1} = y_m$, $m_k = m$, $\sigma_{m_k} = \sigma/2^m$, либо будет построена такая бесконечная последовательность точек $\{y_m\}_{m=1}^\infty$, что $y_m \in \mathcal{D} \cap \bar{\mathcal{D}}(\text{сог } x_k, \sigma/2^m)$ и $P(y_m) - P(x_k) \leq \sigma/2^m$ при каждом $m = 0, 1, \dots$. В последнем случае будем говорить, что k -я итерация *циклическая*.

Теорема. Если $\mathcal{D}$ — замкнутое множество, P — равномерно непрерывная на $\mathcal{D}$ функция и для каждой точки $x \in \mathcal{D} \setminus \pi$ найдутся такие $\varepsilon > 0$ и $\delta > 0$, что $P(\text{сог } z) - P(z) \geq \delta$ для любого $z \in \mathcal{D} \cap \bar{\mathcal{D}}(x, \varepsilon)$, то либо при некотором k встретится циклическая итерация и тогда $x_k \in \pi$, либо будет построена такая бесконечная последовательность $x_1, x_2, \dots$, что любая ее сходящаяся подпоследовательность сходится к приемлемой точке.

Доказательство. Предположим сначала, что k -я итерация циклическая. Тогда существует такая последовательность $y_0, y_1, \dots$, что $P(y_m) - P(x_k) \leq \sigma/2^m$, $\|\text{сог } x_k - y_m\| \leq \sigma/2^m$, $m = 1, 2, \dots$. Следовательно, $y_m \rightarrow \text{сог } x_k$ при $m \rightarrow \infty$, и в силу непрерывности функции P имеем $P(y_m) \rightarrow P(\text{сог } x_k)$ при $m \rightarrow \infty$, а значит, $P(\text{сог } x_k) \leq P(x_k)$. Поскольку из условия теоремы следует, что $P(\text{сог } x) > P(x)$ для любой точки $x \in \mathcal{D} \setminus \pi$, то $x_k \in \pi$.

Предположим теперь, что ни одна итерация алгоритма не является циклической. В этом случае будет построена такая бесконечная последовательность $x_1, x_2, \dots$, что $\|x_{k+1} - \text{сog } x_k\| \leq \sigma_k$, $P(x_{k+1}) - P(x_k) \geq \sigma_k$. Пусть $\{x_k\}_{k \in K}$ — подпоследовательность последовательности $\{x_k\}_{k=1}^\infty$, сходящаяся к точке $x_0 \in \mathcal{D}$ ($\mathcal{D}$ — замкнутое множество). Предположим, что $x_0 \in \mathcal{D} \setminus \pi$, и покажем, что это предположение приводит к противоречию.

Если $x_0 \in \mathcal{D} \setminus \pi$, то существуют такие $\varepsilon > 0$, $\delta > 0$, что $P(\text{сog } z) - P(z) \geq \delta$ для любого $z \in \mathcal{D} \cap \overline{\mathcal{D}}(x_0, \varepsilon)$. Следовательно, найдется такое k' , что

$$P(\text{сog } x_k) - P(x_k) \geq \delta \quad (1)$$

для любого $k \geq k'$, $k \in K$. Для двух последовательных индексов k_1 и k_2 из K имеем

$$P(x_{k_2}) - P(x_{k_1}) = \sum_{i=k_1}^{k_2-1} [P(x_{i+1}) - P(x_i)] \geq P(x_{k_1+1}) - P(x_{k_1}) \geq \sigma_{k_1} \quad (2)$$

и так как $P(x_k) \rightarrow P(x_0)$ при $k \rightarrow \infty$, $k \in K$, в силу непрерывности P , то $\sigma_k \rightarrow 0$ при $k \rightarrow \infty$, $k \in K$. Кроме того, функция P равномерно непрерывна, и, следовательно, можно найти такое $\varepsilon' > 0$, что $|P(u) - P(v)| < \delta/2$, если $\|u - v\| \leq \varepsilon'$. Выберем такое k'' , что $\sigma_k < \min\{\varepsilon', \delta/2\}$ при $k \geq k''$, $k \in K$. Но тогда $\|x_{k+1} - \text{сog } x_k\| \leq \sigma_k < \varepsilon'$, и, следовательно,

$$P(\text{сog } x_k) - P(x_{k+1}) < \delta/2 \quad (3)$$

при $k \geq k''$, $k \in K$. Положим $k_0 = \max\{k', k''\}$. Тогда при $k \geq k_0$, $k \in K$, выполняются оба неравенства (1) и (3), а значит, $P(x_{k+1}) - P(x_k) > \delta/2$ при $k \geq k_0$, $k \in K$, что, как мы уже видели (см. (2)), противоречит сходимости последовательности $\{P(x_k)\}_{k \in K}$ к $P(x_0)$. ■

З а м е ч а н и е. Если множество $\mathcal{D}$ ограничено или известно, что $\|x - \text{сog } x\| \leq r$ для любого $x \in \mathcal{D}$ (или даже $\|x_k - \text{сog } x_k\| \leq r$ при $k = 1, 2, \dots$), то, как следует из приведенного выше доказательства теоремы, ее утверждение остается верным в предположении непрерывности (а не равномерной непрерывности) функции P на $\mathcal{D}$. Чтобы убедиться в этом, достаточно заметить, что при сделанных здесь предположениях все интересующие нас точки находятся в компактном множестве, содержащем точку x_0 , на котором непрерывная функция P равномерно непрерывна.

§ 2. Задачи безусловной оптимизации

В этом параграфе обсуждается экстремальная задача (R^n, f) максимизации непрерывно дифференцируемой вогнутой функции $f: R^n \rightarrow R$. Методы, которые здесь рассматриваются, являются по существу методами допустимых направлений, о ко-

торых уже говорилось в § 2 гл. 4. Их конкретное описание будет связано с содержательным определением таких понятий, входящих в общие схемы § 1, как множество приемлемых точек, корректирующее отображение и оценочная функция, а доказательство сходимости — с проверкой условий соответствующих теорем § 1.

2.1. Приемлемые точки. В задачах безусловной оптимизации чаще всего в качестве множества π приемлемых точек выбирают множество стационарных точек функции f , т. е. множество $\{x \in R^n : \nabla f(x) = 0\}$. Известно, что условие $\nabla f(\bar{x}) = 0$ необходимо для того, чтобы $\bar{x} \in R^n$ была точкой локального максимума (минимума) непрерывно дифференцируемой функции f . В случае вогнутой (выпуклой) функции это условие является достаточным условием оптимальности вектора $\bar{x}$.

Для оптимальности вектора $\bar{x} \in R^n$ в экстремальной задаче (R^n, f) максимизации вогнутой (минимизации выпуклой) непрерывно дифференцируемой на R^n функции $f: R^n \rightarrow R$ необходимо и достаточно, чтобы $\nabla f(\bar{x}) = 0$.

Это утверждение непосредственно следует из «рабочего» критерия оптимальности (см. п. 3.1.7). Однако по содержанию оно значительно проще этого критерия, и мы приведем здесь его прямое доказательство.

Достаточность. Пусть $\nabla f(\bar{x}) = 0$. Тогда в силу вогнутости функции f (см. п. 1.8.7) для любого $x \in R^n$

$$f(x) - f(\bar{x}) \leq \nabla f(\bar{x})(x - \bar{x}) = 0,$$

т. е. $f(x) \leq f(\bar{x})$ для любого $x \in R^n$. ■

Необходимость. Пусть $f(x) \leq f(\bar{x})$ для любого $x \in R^n$. Тогда $(1/\lambda)[f(\bar{x} + \lambda s) - f(\bar{x})] \leq 0$ для любого $s \in R^n$ и любого $\lambda > 0$. Следовательно (см. п. 1.2.10),

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} [f(\bar{x} + \lambda s) - f(\bar{x})] = \nabla f(\bar{x})s \leq 0. \quad (1)$$

Полагая в (1) $s = \nabla f(\bar{x})$, имеем $\|\nabla f(\bar{x})\|^2 \leq 0$, т. е. $\nabla f(\bar{x}) = 0$. ■

Отметим также, что $\nabla f(\bar{x}) = 0$ является необходимым и достаточным условием оптимальности вектора $\bar{x}$ в экстремальной задаче выпуклого программирования $(\mathcal{D}, f)$, если $\bar{x}$ — внутренняя точка выпуклого множества $\mathcal{D}$.

В дальнейшем при исследовании задачи безусловной оптимизации (R^n, f) выпуклого программирования будем полагать $\pi = \{x \in R^n : \nabla f(x) = 0\}$.

2.2. Оценочная функция. Всюду далее в качестве оценочной функции принимается целевая функция задачи.

В п. 1.1 мы условились рассматривать только такие методы решения экстремальных задач, в которых $P(\text{сog } x) \geq P(x)$ для любого $x \in \mathcal{D}$. Поэтому будем строить только такие коррекции, что $f(\text{сog } x) \geq f(x)$ для любого $x \in R^n$.

Особо отметим, что для рассматриваемого нами класса экстремальных задач функция f непрерывна на R^n , поскольку предполагается ее непрерывная дифференцируемость. Впрочем, в силу теоремы п. 1.8.5 непрерывность функции f на R^n следует и из вогнутости.

В дальнейшем нам понадобится следующая лемма.

Лемма. Пусть $f: R^n \rightarrow R$ — замкнутая выпуклая функция (см. п. 1.8.4) и для некоторого $\alpha_0 \in R$ множество $L_{\alpha_0} = \{x \in R^n: f(x) \leq \alpha_0\}$ непусто и ограничено. Тогда при любом $\alpha \in R$ множество $L_\alpha = \{x \in R^n: f(x) \leq \alpha\}$ компактно.

Доказательство. В п. 1.8.2, 1° было установлено, что L_α — замкнутое (см. упражнение 1.12) выпуклое (см. п. 1.8.2, 1°) множество в R^n при любом $\alpha \in R$. Докажем его ограниченность. Предположим противное: существует такое $\alpha \in R$, что множество L_α не ограничено.

Тогда (см. теорему 1 п. 1.6.2) существует вектор $s \in R^n$, $s \neq 0$, принадлежащий конусу неограниченности $K_{L_\alpha}^\infty$ множества L_α .

Если x — какая-нибудь точка множества L_α , то луч $\{x + \lambda s, \lambda \geq 0\} \subset L_\alpha$. Но тогда луч $\{(x + \lambda s, \alpha) = (x, \alpha) + \lambda(s, 0), \lambda \geq 0\}$ принадлежит замкнутому выпуклому множеству $\text{epi } f$, и, следовательно, вектор $(s, 0)$ принадлежит конусу неограниченности $K_{\text{epi } f}^\infty$ множества $\text{epi } f$ (рис. 35).

Так как $L_{\alpha_0} \neq \emptyset$, то существует $x_0 \in L_{\alpha_0}$, $(x_0, \alpha_0) \in \text{epi } f$, и, следовательно (см. теорему 2 п. 1.6.2), луч $\{(x_0 + \lambda s, \alpha_0), \lambda \geq 0\} \subset \text{epi } f$. Но тогда $\{x_0 + \lambda s, \lambda \geq 0\} \subset L_{\alpha_0}$, что противоречит ограниченности множества L_{α_0} . ■

Из доказанной леммы непосредственно следует, что если $f: R^n \rightarrow \bar{R}$ — замкнутая вогнутая функция и при некотором $\alpha_0 \in R$ множество $\hat{L}_{\alpha_0} = \{x \in R^n: f(x) \geq \alpha_0\}$ непусто и ограничено, то для любого $\alpha \in R$ множество $\hat{L}_\alpha = \{x \in R^n: f(x) \geq \alpha\}$ компактно.

Докажем теперь утверждение, на которое будем ссылаться как в этом, так и в следующем параграфах.

Теорема. Если $f: R^n \rightarrow \bar{R}$ — непрерывная на $\text{dom } f$ вогнутая функция и существует такая точка $x_0 \in R^n$, что множество $\hat{L}_{f(x_0)} = \{x \in R^n: f(x) \geq f(x_0)\}$ ограничено, то задача (R^n, f) имеет решение и множество ее оптимальных планов содержится в $\hat{L}_{f(x_0)}$; если к тому же $\mathcal{D}$ — непустое замкнутое множество в R^n , то экстремальная задача максимизации $(\mathcal{D}, f)$ имеет решение и при любом $z_0 \in \mathcal{D} \cap \text{dom } f$ множество решений задачи $(\mathcal{D}, f)$ содержится в $\hat{L}_{f(z_0)}$.

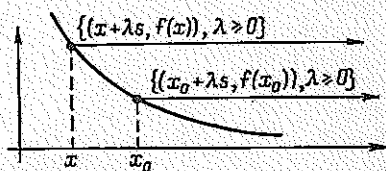


Рис. 35.

Доказательство. Достаточно доказать вторую часть утверждения, первая следует из нее при $\mathcal{D} = R^n$. Очевидно, что $x_0 \in \text{dom } f$: в противном случае $\hat{L}_{f(x_0)} = R^n$. Из непрерывности функции f на $\text{dom } f$ следует ее замкнутость. Поэтому в силу леммы множество $\hat{L}_{f(z_0)}$ компактно. Итак, $\hat{L}_{f(z_0)} \cap \mathcal{D}$ — непустое компактное множество. По теореме Вейерштрасса непрерывная функция f принимает наибольшее на этом множестве значение в некоторой его точке $\bar{x}$. Но тогда $f(\bar{x}) \geq f(z_0)$ и $f(\bar{x})$ — наибольшее значение функции f на $\mathcal{D}$, так как $f(x) \leq f(z_0)$ вне $\hat{L}_{f(z_0)}$. По этой же причине любое решение задачи $(\mathcal{D}, f)$ содержится в $\hat{L}_{f(z_0)}$. ■

2.3. Модифицированный алгоритм наискорейшего спуска. Перейдем теперь к непосредственному изложению некоторых алгоритмов решения экстремальных задач максимизации вогнутой функции f на R^n . Формальная процедура алгоритма построения итеративной последовательности определяется конкретным заданием корректирующего отображения.

Пусть $G(x) = (g_{ij}(x))$ — квадратная матрица порядка n , элементами которой являются непрерывные функции $g_{ij}: R^n \rightarrow R$. Предположим, что при любом $x \in R^n$ матрица $G(x)$ симметрична и положительно определена.

Для любого $x \in R^n$ положим

$$\text{сог}(x) = x + \lambda(x) G(x) \nabla f(x), \quad (2)$$

где $G(x) \nabla f(x)$ — произведение матрицы $G(x)$ на вектор $\nabla f(x)$, а $\lambda(x)$ — наименьшее из неотрицательных λ , удовлетворяющих условию

$$f(x + \lambda G(x) \nabla f(x)) = \max_{\mu \geq 0} f(x + \mu G(x) \nabla f(x)). \quad (3)$$

Ниже будет показано, что такое число $\lambda(x)$ существует для любого $x \in R^n$, если множество $L_{f(x)}$ компактно. Более того, будут указаны способы вычисления $\lambda(x)$ с любой заданной степенью точности.

Обозначим $G(x) \nabla f(x)$ через $s(x)$.

Лемма 1. Если $\nabla f(x) \neq 0$, то $s(x) = G(x) \nabla f(x)$ — прогрессивное направление в точке x .

Доказательство. В п. 3.1.4 установлено, что необходимым и достаточным условием прогрессивности направления s в точке x является неравенство $\nabla f(x)s > 0$. Но $\nabla f(x)s(x) = \nabla f(x)G(x)\nabla f(x) > 0$, так как по условию матрица $G(x)$ положительно определена, а $\nabla f(x) \neq 0$. ■

Лемма 2. Если множество $\hat{L}_{f(x)} = \{z \in R^n : f(z) \geq f(x)\}$ компактно, то существует число $\lambda(x)$ — наименьшее из неотрицательных чисел λ , удовлетворяющих условию (3).

Доказательство. Обозначим $M = \{\mu \geq 0 : x + \mu s(x) \in \hat{L}_{f(x)}\}$. Так как $L_{f(x)}$ — компактное выпуклое множество, то

$M = [0, \bar{\mu}]$. Функция $\varphi(\mu) = f(x + \mu s(x))$ непрерывна на M , и по теореме Вейерштрасса множество

$$\Lambda = \{\lambda \in M : f(x + \lambda s(x)) = \max_{\mu \geq 0} f(x + \mu s(x))\} \neq \emptyset.$$

Кроме того, φ — вогнутая функция на M , так что Λ — непустое замкнутое выпуклое подмножество отрезка M (см. п. 1.8.2), т. е. отрезок $[\lambda_1, \lambda_2]$ (на луче $\{z \in R^n : z = x + \mu s(x), \mu \geq 0\}$ — отрезок $[a, b]$, рис. 36). Так как $\varphi(\mu) \geq f(x)$ при любом $\mu \in M$, а $\varphi(\mu) < f(x)$ при $\mu > \bar{\mu}$, то Λ совпадает с множеством неотрицательных чисел λ , удовлетворяющих условию (3), и $\lambda(x) = \lambda_1$. ■

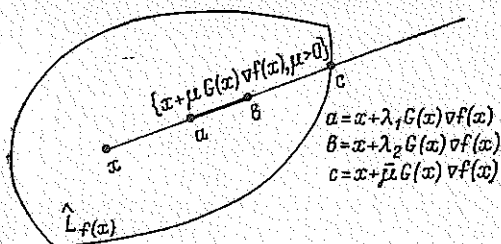


Рис. 36.

Следствие. Если при некотором $x_0 \in R^n$ множество $\hat{L}_{f(x_0)}$ ограничено, то формула (2) определяет корректирующее преобразование на R^n , причем $f(\text{сог}(x)) > f(x)$, если $x \in \pi = \{z \in R^n : \nabla f(z) = 0\}$.

Доказательство. Из леммы п. 2.2 следует, что множество $L_{f(x)}$ компактно при любом $x \in R^n$. Поэтому в силу леммы 2 определено отображение $\lambda : R^n \rightarrow R$, а значит (см. (2)), и отображение $\text{сог} : R^n \rightarrow R^n$ (см. (2)). Очевидно, если $\nabla f(x) = 0$, то $f(\text{сог}(x)) = f(x)$. Если $\nabla f(x) \neq 0$, то в силу леммы 1 существует такое $\mu > 0$, что $f(x + \mu s(x)) > f(x)$. Но тогда тем более $\lambda(x) > 0$ и $f(\text{сог}(x)) > f(x)$. ■

Опишем теперь процесс построения итеративной последовательности в соответствии со схемой метода п. 1.2.

Построение последовательности $x_1, x_2, \dots$ начинается с произвольной точки $x_1 \in R^n$. Предположим, что уже построены векторы $x_1, x_2, \dots, x_k$. Вычислим $\nabla f(x_k)$. Если $\nabla f(x_k) = 0$, то $\bar{x} = x_k$ — решение задачи (R^n, f) . В противном случае, т. е. если $\nabla f(x_k) \neq 0$, находим $G(x_k)$, $s(x_k) = G(x_k) \nabla f(x_k)$ и для вычисления $\lambda(x_k)$ решаем задачу одномерной оптимизации. Полагаем $x_{k+1} = x_k + \lambda(x_k) s(x_k)$.

Теорема. Если при некотором $a \in R$ множество L_a непусто и ограничено, то алгоритм осуществляет построение итеративной последовательности $x_1, x_2, \dots$. Последовательность $x_1, x_2, \dots$

либо конечна и ее последний элемент $\bar{x}$ есть решение задачи (R^n, f) , либо бесконечна. В последнем случае она содержит подпоследовательность, сходящуюся к решению задачи (R^n, f) , и всякая сходящаяся подпоследовательность последовательности $\{x_k\}_{k=1}^{\infty}$ сходится к некоторому решению задачи (R^n, f) .

Доказательство. Мы уже установили, что алгоритм осуществляет построение итеративной последовательности (см. следствие из лемм 1, 2). Элементы этой последовательности принадлежат компактному множеству $\bar{L}_f(x_1)$ (см. лемму п. 2.2), и если последовательность $x_1, x_2, \dots$ бесконечна, то она содержит сходящуюся подпоследовательность $\{x_k\}_{k \in K}$, $x_k \rightarrow \bar{x}$ при $k \rightarrow \infty, k \in K$. Покажем, что $\bar{x} \in \pi = \{z \in R^n: \nabla f(z) = 0\}$. В силу теоремы п. 1.2 достаточно для любого $x \in R^n$, для которого $\nabla f(x) \neq 0$, установить существование таких $\varepsilon > 0, \delta > 0$, что $f(\operatorname{cog} z) - f(z) \geq \delta$ для любого $z \in \bar{\mathcal{D}}(x, \varepsilon)$.

Рассмотрим функцию $g: R^n \rightarrow R$, $g(z) = f(z + \lambda(x)s(z)) - f(z)$. Функции f и $s: R^n \rightarrow R$, $s(z) = G(z)\nabla f(z)$ непрерывны (f — непрерывно дифференцируемая функция, g_{ii} — непрерывные функции) и, следовательно, непрерывна функция g . Поскольку $\nabla f(x) \neq 0$, то, как указано в следствии из лемм 1 и 2, $g(x) = f(\operatorname{cog} x) - f(x) > 0$. Положим $g(x) = 2\delta$. В силу непрерывности функции g в точке x найдется такое $\varepsilon > 0$, что $g(z) \geq \delta$ для любого $z \in \bar{\mathcal{D}}(x, \varepsilon)$. Но тогда в соответствии с определением $\lambda(z)$

$$\begin{aligned} & f(\operatorname{cog} z) - f(z) = \\ & = f(z + \lambda(z)s(z)) - f(z) \geq f(z + \lambda(x)s(z)) - f(z) = g(z) \geq \delta \end{aligned}$$

для любого $z \in \bar{\mathcal{D}}(x, \varepsilon)$. ■

Замечание 1. Если существует единственное решение задачи (R^n, f) (например, в случае, когда f — строго вогнутая функция, принимающая наибольшее значение в некоторой точке), то последовательность $x_1, x_2, \dots$ либо конечна и ее последний элемент $\bar{x}$ — решение задачи (R^n, f) , либо сходится к решению этой задачи.

Замечание 2. Если $\{x_k\}_{k \in K}$ — подпоследовательность последовательности $\{x_k\}_{k=1}^{\infty}$, сходящаяся к решению $\bar{x}$ задачи (R^n, f) при $k \rightarrow \infty, k \in K$, то, как следует из непрерывности f , последовательность $\{f(x_k)\}_{k \in K}$ сходится к $f(\bar{x})$ при $k \rightarrow \infty, k \in K$. Но тогда из монотонности последовательности $\{f(x_k)\}_{k=1}^{\infty}$ следует (без предположения ограниченности множества L_α при некотором $\alpha \in R$), что $f(x_k) \rightarrow f(\bar{x})$ при $k \rightarrow \infty$. Таким образом, последовательность значений целевой функции в точках $x_1, x_2, \dots$ сходится и притом монотонно к оптимальному значению целевой функции.

2.4. Две конкретизации корректирующего отображения.

1. Пусть $G(x) \equiv I$, где I — единичная матрица порядка n . Тогда $s(x) = \nabla f(x)$. В этом случае описанный выше алгоритм есть алгоритм наискорейшего спуска Коши.

2. Предположим, что функция f дважды непрерывно дифференцируема и ее матрица Гессе $G(x) = \left(\frac{\partial^2 f(x)}{\partial x[i] \partial x[j]} \right), i \in 1:n, j \in 1:n$ отрицательно определена при любом $x \in R^n$ (см. п. 1.8.7). Положим при фиксированном $x_0 \in R^n$

$$\bar{f}_0(x) = f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^T G(x_0)(x - x_0).$$

Функция $\bar{f}_0$ есть вогнутая квадратичная аппроксимация функции f в точке x_0 . Она принимает наибольшее значение в точке x^* , для которой $\nabla \bar{f}_0(x^*) = 0$. Но $\nabla \bar{f}(x) = \nabla f(x_0) + G(x_0)(x - x_0)$. Полагая $\nabla \bar{f}_0(x^*) = \nabla f(x_0) + G(x_0)(x^* - x_0) = 0$, получаем

$$x^* = x_0 - G^{-1}(x_0) \nabla f(x_0).$$

Таким образом, вектор $s(x_0) = -G^{-1}(x_0) \nabla f(x_0)$ есть направление из точки x_0 в точку максимума квадратичной аппроксимации функции f в точке x_0 . Имея в виду, что матрица $-G^{-1}(x)$ положительно определена при любом $x \in R^n$, зададим корректирующее отображение формулой

$$\text{cor } x = x - \lambda(x) G^{-1}(x) \nabla f(x). \quad (4)$$

При таком определении коррекции алгоритм п. 2.3 есть алгоритм Ньютона.

2.5. Построение итеративной последовательности по приближенным значениям $\lambda(x)$ и $s(x)$. Алгоритмы вычисления шага, которые рассматриваются в следующем пункте, в общем случае приводят к построению бесконечной последовательности положительных чисел $\lambda_1, \lambda_2, \dots$, сходящейся к $\lambda(x)$. Поэтому внесем некоторые коррективы в формулировку алгоритма п. 2.3 построения итеративной последовательности $x_1, x_2, \dots$. А именно установим, что возможность вычисления числа $\lambda(x)$, а также матрицы $G(x)$ и вектора $\nabla f(x)$ с любой заданной степенью точности позволяет строить итеративную последовательность в соответствии со схемой алгоритма п. 1.5. Сходимость итеративной последовательности в смысле теоремы п. 1.5 устанавливается в предположении ограниченности множества $\hat{L}_{f(x)}$ при некотором $x_0 \in R^n$.

Зададимся положительным числом σ . Выберем точку $x_1 \in R^n$. Предположим, что уже найдены такие точки $x_1, x_2, \dots, x_k$ и такие неотрицательные целые числа $m_1, m_2, \dots, m_{k-1}$, что

$$\|x_{i+1} - \text{cor } x_i\| \leq \sigma/2^{m_i}, \quad f(x_{i+1}) - f(x_i) \geq \sigma/2^{m_i}$$

для каждого $i \in 1:k-1$. Вычислим приближенные значения $\tilde{G}(x_k)_0$, $\tilde{\nabla} f(x_k)_0$ и $\tilde{\lambda}(x_k)_0$ так, чтобы

$$\|\lambda(x_k) G(x_k) \nabla f(x_k) - \tilde{\lambda}(x_k)_0 \tilde{G}(x_k)_0 \tilde{\nabla} f(x_k)_0\| \leq \sigma.$$

Если $f(x_k + \tilde{\lambda}(x_k)_0 \tilde{G}(x_k)_0 \tilde{\nabla} f(x_k)_0) - f(x_k) \geq \sigma$, то положим $x_{k+1} = x_k + \tilde{\lambda}(x_k)_0 \tilde{G}(x_k)_0 \tilde{\nabla} f(x_k)_0$. В противном случае продолжим процесс построения приближенного значения $\text{сог } x_k$, заменив σ на $\sigma/2$, и т. д. Если за конечное число шагов найдем такое число m_k и такие приближенные значения $\tilde{G}(x_k)_{m_k}$, $\tilde{\nabla} f(x_k)_{m_k}$ и $\tilde{\lambda}(x_k)_{m_k}$, что

$$\|\lambda(x_k) G(x_k) \nabla f(x_k) - \tilde{\lambda}(x_k)_{m_k} \tilde{G}(x_k)_{m_k} \tilde{\nabla} f(x_k)_{m_k}\| \leq \sigma/2^{m_k},$$

$$f(x_k + \tilde{\lambda}(x_k)_{m_k} \tilde{G}(x_k)_{m_k} \tilde{\nabla} f(x_k)_{m_k}) - f(x_k) \geq \sigma/2^{m_k},$$

то $x_{k+1} = x_k + \tilde{\lambda}(x_k)_{m_k} \tilde{G}(x_k)_{m_k} \tilde{\nabla} f(x_k)_{m_k}$. В противном случае k -я итерация циклическая и $\bar{x} = x_k$.

Очевидно, что описанная выше процедура укладывается в схему алгоритма п. 1.5. Кроме того, все точки последовательности $x_1, x_2, \dots$ содержатся в компактном множестве $\tilde{L}_{f(x_0)}$, на котором функция f равномерно непрерывна. Из доказательства теоремы п. 2.3 следует, что выполнены условия теоремы п. 1.5 (см. также замечание к этой теореме). Поэтому верна такая теорема.

Теорема. Если при некотором $x_0 \in R^n$ множество $L_{f(x_0)}$ ограничено, то при любом $\sigma > 0$ алгоритм осуществляет построение итеративной последовательности $x_1, x_2, \dots$. При этом либо при некотором k встретится циклическая итерация, и тогда $\nabla f(x_k) = 0$, $\bar{x} = x_k$ — решение задачи (R^n, f) , либо будет построена бесконечная последовательность $\{x_k\}_{k=1}^\infty$, содержащая подпоследовательность $\{x_k\}_{k \in K}$, сходящуюся к решению задачи (R^n, f) , и всякая сходящаяся подпоследовательность последовательности $\{x_k\}_{k=1}^\infty$ сходится к некоторому решению задачи (R^n, f) .

2.6. Задача вычисления шага $\lambda(x)$. Вычисление шага $\lambda(x)$, т. е. наименьшего положительного числа λ , удовлетворяющего условию (4), проводим в предположении компактности множества $\tilde{L}_{f(x_0)} = \{z \in R^n : f(z) \geq f(x_0)\}$ при некотором (а следовательно (см. теорему п. 2.2), и при любом) $x_0 \in R^n$.

Вычисление $\lambda(x)$ проводится в два этапа. Сначала определяется отрезок $[\lambda', \lambda'']$ длиной Δ , содержащий точку $\lambda(x)$, а затем осуществляется построение такой вложенной последовательности отрезков

$$[\lambda', \lambda''] = [\lambda'_1, \lambda''_1] \supset [\lambda'_2, \lambda''_2] \supset \dots, \quad (5)$$

что $\lambda(x) \in [\lambda'_k, \lambda''_k]$ при любом k и $\lambda''_k - \lambda'_k \rightarrow 0$ при $k \rightarrow \infty$.

Отрезок $[\lambda', \lambda'']$ определяем следующим образом. Сначала находим такое число $\Delta > 0$, что $f\left(x + \frac{\Delta}{2} s(x)\right) > f(x)$. Это можно осуществить, последовательно вычисляя $f\left(x + \frac{\mu'}{2m} s(x)\right)$ при $m = 1, 2, \dots$ для некоторого $\mu' > 0$. Число Δ полагаем равным μ'/m' , если $f\left(x + \frac{\mu'}{2m'} s(x)\right) > f(x)$ (существование такого m' следует из леммы 1 п. 2.3). Затем последовательно вычисляются числа $f(x)$, $f\left(x + \frac{\Delta}{2} s(x)\right)$, $f\left(x + 2 \frac{\Delta}{2} s(x)\right)$, $\dots$. Процесс вычисления оканчивается, когда будет найдено такое натуральное m'' , что

$$f\left(x + (m'' - 1) \frac{\Delta}{2} s(x)\right) < f\left(x + m'' \frac{\Delta}{2} s(x)\right)$$

$$\text{и} \quad f\left(x + m'' \frac{\Delta}{2} s(x)\right) \geq f\left(x + (m'' + 1) \frac{\Delta}{2} s(x)\right)$$

(существование такого m'' следует из леммы 2 п. 2.3). Полагаем $\lambda' = (m'' - 1) \frac{\Delta}{2}$, $\lambda'' = (m'' + 1) \frac{\Delta}{2}$. Нетрудно проверить, что из вогнутости функции $\varphi(\mu) = f(x + \mu s(x))$ следует $\lambda(x) \in [\lambda', \lambda'']$. — Упражнение 9.1.

Построение последовательности (5) может быть осуществлено разными способами.

1°. *Метод золотого сечения, или поиск Фибоначчи.* Положим $F_1 = (3 - \sqrt{5})/2 \approx 0,38$, $F_2 = (\sqrt{5} - 1)/2 \approx 0,62$. Отметим, что $F_1 + F_2 = 1$ и $F_2^2 = F_1$. На первом шаге вычислим $\mu'_1 = \lambda'_1 + F_1(\lambda''_1 - \lambda'_1) = F_2\lambda'_1 + F_1\lambda''_1$, $\mu''_1 = \lambda'_1 + F_2(\lambda''_1 - \lambda'_1) = F_1\lambda'_1 + F_2\lambda''_1$, $\varphi(\mu'_1)$ и $\varphi(\mu''_1)$. Из неравенств

$$\varphi(\lambda'_1) < \varphi\left(\frac{\lambda'_1 + \lambda''_1}{2}\right), \quad \varphi(\lambda''_1) \leq \varphi\left(\frac{\lambda'_1 + \lambda''_1}{2}\right)$$

и вогнутости функции φ следует, что

$$\varphi(\lambda'_1) < \varphi(\mu'_1), \quad \varphi(\mu''_1) \geq \varphi(\lambda''_1).$$

а) Если $\varphi(\mu'_1) < \varphi(\mu''_1)$, то положим $\lambda'_2 = \mu'_1$, $\lambda''_2 = \lambda''_1$. Тогда $\lambda'_2 < \mu''_1 < \lambda''_2$, $\varphi(\lambda'_2) < \varphi(\mu''_1)$, $\varphi(\mu''_1) \geq \varphi(\lambda''_2)$. Следовательно, $\lambda(x) \in [\lambda'_2, \lambda''_2]$ и $\lambda''_2 - \lambda'_2 = \lambda''_2 - (F_2\lambda'_1 + F_1\lambda''_2) = (1 - F_1)\lambda'_1 - F_2\lambda''_1 = F_2\Delta$ (рис. 37,а).

б) Если $\varphi(\mu'_1) \geq \varphi(\mu''_1)$, то положим $\lambda'_2 = \lambda'_1$, $\lambda''_2 = \mu''_1$. Тогда $\lambda'_2 < \mu'_1 < \lambda''_2$, $\varphi(\lambda'_2) < \varphi(\mu'_1)$, $\varphi(\mu'_1) \geq \varphi(\lambda''_2)$. Следовательно, $\lambda(x) \in [\lambda'_2, \lambda''_2]$ и $\lambda''_2 - \lambda'_2 = F_1\lambda'_1 + F_2\lambda''_1 - \lambda'_1 = F_2\lambda''_1 - (1 - F_1)\lambda'_1 = F_2\Delta$ (рис. 37,б).

Итак, как в случае а), так и в случае б) $\lambda(x) \in [\lambda'_2, \lambda''_2]$ и $\lambda''_2 - \lambda'_2 = F_2 \Delta$.

На следующем шаге вычислим $\mu'_2 = F_2 \lambda'_2 + F_1 \lambda''_2$, $\mu''_2 = F_1 \lambda'_2 + F_2 \lambda''_2$, $\varphi(\mu'_2)$ и $\varphi(\mu''_2)$. Отметим, что в случае а)

$$\begin{aligned}\mu'_2 &= F_2(F_2 \lambda'_1 + F_1 \lambda''_1) + F_1 \lambda''_1 = F_1 \lambda'_1 + (1 - F_2)(1 + F_2) \lambda''_1 = \\ &= F_1 \lambda'_1 + (1 - F_1) \lambda''_1 = \mu''_1,\end{aligned}$$

а в случае б)

$$\mu''_2 = F_1 \lambda'_2 + F_2 \lambda''_2 = F_1 \lambda'_1 + F_2(F_1 \lambda'_1 + F_2 \lambda''_1) = F_2 \lambda'_1 + F_1 \lambda''_1 = \mu'_1.$$

Это обстоятельство естественно сокращает требуемый объем новых вычислений на каждой итерации. Кроме того, как в случае а), так и в случае б) $\varphi(\lambda'_2) < \varphi(\mu'_2)$, $\varphi(\mu''_2) \geq \varphi(\lambda''_2)$, и построение отрезка $[\lambda'_3, \lambda''_3]$, содержащего $\lambda(x)$, осуществляется

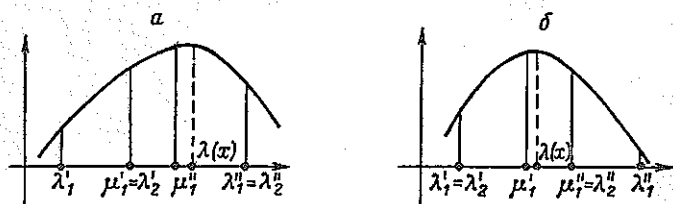


Рис. 37.

в соответствии с описанной выше процедурой. Остается отметить, что $\lambda''_k - \lambda'_k = (F_2)^k \Delta \rightarrow 0$ при $k \rightarrow \infty$.

2°. *Метод деления пополам (метод Больцано)*. Из вогнутости функции f и неравенств

$$\varphi(\lambda'_1) < \varphi\left(\frac{\lambda'_1 + \lambda''_1}{2}\right), \quad \varphi\left(\frac{\lambda'_1 + \lambda''_1}{2}\right) \geq \varphi(\lambda''_1)$$

следует, что производные $d\varphi(\lambda'_1)/d\mu > 0$, $d\varphi(\lambda''_1)/d\mu \leq 0$.

Вычислим $d\varphi\left(\frac{\lambda'_1 + \lambda''_1}{2}\right)/d\mu$. Если $d\varphi\left(\frac{\lambda'_1 + \lambda''_1}{2}\right)/d\mu > 0$, то положим $\lambda'_2 = \frac{\lambda'_1 + \lambda''_1}{2}$, $\lambda''_2 = \lambda''_1$. В противном случае если $d\varphi\left(\frac{\lambda'_1 + \lambda''_1}{2}\right)/d\mu \leq 0$, положим $\lambda'_2 = \lambda'_1$, $\lambda''_2 = \frac{\lambda'_1 + \lambda''_1}{2}$. Из вогнутости функции φ следует, что $\lambda(x) \in [\lambda'_2, \lambda''_2]$.

Затем вычислим $d\varphi\left(\frac{\lambda'_2 + \lambda''_2}{2}\right)/d\mu$, определим λ'_3, λ''_3 и т. д. Остается отметить, что $\lambda''_k - \lambda'_k = \frac{1}{2^{k-1}} \Delta \rightarrow 0$ при $k \rightarrow \infty$.

2.7. Метод локальных вариаций. В этом пункте рассмотрим еще один метод решения задачи выпуклого программирования (R^n, f) . По сравнению с методами, изложенными в пп. 2.3—2.6, он обладает тем достоинством, что его применение не связано с вычислением производных функции f . Он оказывается особенно эффективным, если функция представлена в виде $f(x) = \sum_{i=1}^n f_i(x[i])$. Как и ранее, будем предполагать, что множество $\hat{L}_{f(x_0)}$ ограничено при некотором $x_0 \in R^n$.

Пусть $e_1, e_2, \dots, e_n$ — канонический базис пространства R^n (см. п. 1.1.1). Возьмем какую-нибудь точку $x_1 \in R^n$ и число $\sigma = \sigma_1 > 0$. Предположим, что уже построены точки $x_1, x_2, \dots, x_k$ и числа $\sigma_1, \sigma_2, \dots, \sigma_k$. Опшем k -ю итерацию — процесс построения точки x_{k+1} и числа σ_{k+1} .

Последовательно вычисляем числа $f(x_k + \sigma_k e_1), f(x_k - \sigma_k e_1), f(x_k + \sigma_k e_2), f(x_k - \sigma_k e_2), \dots$. Если все члены этой последовательности не превосходят $f(x_k)$, то полагаем $x_{k+1} = x_k, \sigma_{k+1} = \frac{1}{2} \sigma_k$. В противном случае, т. е. если на каком-то шаге этих вычислений встретится число больше $f(x_k)$, обозначаем через $z_k^{(1)}$ то значение аргумента, при котором встретится это число, и последовательно вычисляем

$$f(z_k^{(1)} + \sigma_k e_1), f(z_k^{(1)} - \sigma_k e_1), f(z_k^{(1)} + \sigma_k e_2), f(z_k^{(1)} - \sigma_k e_2), \dots$$

Если все члены этой последовательности не превосходят $f(z_k^{(1)})$, то полагаем $x_{k+1} = z_k^{(1)}, \sigma_{k+1} = \frac{1}{2} \sigma_k$. В противном случае обозначаем через $z_k^{(2)}$ то значение аргумента, при котором встретилось число больше $f(z_k^{(1)})$, и продолжаем этот процесс. За конечное число шагов (множество $\hat{L}_{f(x_k)}$ компактно) будет построена такая точка $z_k^{(m)}$, что все элементы последовательности

$$f(z_k^{(m)} + \sigma_k e_1), f(z_k^{(m)} - \sigma_k e_1), \dots, f(z_k^{(m)} + \sigma_k e_n), f(z_k^{(m)} - \sigma_k e_n)$$

не будут превосходить $f(z_k^{(m)})$, и положим $x_{k+1} = z_k^{(m)}, \sigma_{k+1} = \sigma_k/2$.

Теорема. Последовательность $\{x_k\}_{k=1}^{\infty}$ содержит подпоследовательность, сходящуюся к решению задачи (R^n, f) , и всякая сходящаяся подпоследовательность последовательности $\{x_k\}_{k=1}^{\infty}$ сходится к решению задачи (R^n, f) . В частности, если найдется такой номер k_0 , что $x_i = x_{k_0}$ при всех $i \geq k_0$, то x_{k_0} — решение задачи (R^n, f) .

Доказательство. Существование сходящейся подпоследовательности следует из компактности множества $\hat{L}_{f(x_0)}$.

Пусть $\{x_k\}_{k \in K}$ — подпоследовательность последовательности $\{x_k\}_{k=1}^{\infty}$, сходящаяся к точке $\bar{x}$. По построению при любом на-

туральном $k \geq 2$ и любом $i \in 1:n$

$$f(x_k + \sigma_{k-1}e_i) \leq f(x_k) \quad \text{и} \quad f(x_k - \sigma_{k-1}e_i) \leq f(x_k).$$

Так как f — непрерывно дифференцируемая вогнутая функция, то (см. п. 1.8.7)

$$0 \leq f(x_k) - f(x_k + \sigma_{k-1}e_i) \leq -\sigma_{k-1} \nabla f(x_k + \sigma_{k-1}e_i) e_i,$$

$$0 \leq f(x_k) - f(x_k - \sigma_{k-1}e_i) \leq \sigma_{k-1} \nabla f(x_k - \sigma_{k-1}e_i) e_i,$$

и, следовательно,

$$\nabla f(x_k + \sigma_{k-1}e_i) e_i \leq 0, \quad \nabla f(x_k - \sigma_{k-1}e_i) e_i \geq 0 \quad (6)$$

при любом $k \geq 2$ и любом $i \in 1:n$. Перейдем теперь к пределу в неравенствах (6) при $k \rightarrow \infty$, $k \in K$. Так как $x_k \rightarrow \bar{x}$ при $k \rightarrow \infty$, $k \in K$, $\sigma_{k-1} = (1/2^{k-2})\sigma \rightarrow 0$ при $k \rightarrow \infty$ и ∇f — непрерывная на R^n функция, то для любого $i \in 1:n$ получаем $\nabla f(\bar{x})e_i = 0$, и, следовательно, $\nabla f(\bar{x}) = 0$. ■

З а м е ч а н и е. При выполнении условий теоремы последовательность $\{f(x_k)\}_{k=1}^{\infty}$ значений целевой функции в точках $x_1, x_2, \dots$ монотонно сходится к оптимальному значению целевой функции (ср. с замечанием 2 п. 2.3). Если к тому же задача (R^n, f) имеет единственное решение $\bar{x}$, то и последовательность $\{x_k\}_{k=1}^{\infty}$ сходится к $\bar{x}$ при $k \rightarrow \infty$.

§ 3. Методы штрафных функций

В этом параграфе будут рассмотрены методы штрафных функций решения задачи выпуклого программирования

$$(\mathcal{D}, f): f(x) \rightarrow \max, \quad \mathcal{D} = \{x \in R^n: f_i(x) \leq 0, i \in 1:m\},$$

где f — вогнутая, а f_i , $i \in 1:m$, — выпуклые функции из R^n в R , т. е. функции, принимающие только конечные значения. Легко видеть, что такая форма записи задачи по существу не ограничивает общности: достаточно, как уже было отмечено в п. 2.2.1, каждое уравнение в приведенных там или в гл. 3 формулировках задачи выпуклого программирования заменить равносильной ему системой из двух неравенств, а условия неотрицательности переменных включить в общую систему неравенств.

Методы штрафных функций позволяют свести решение задачи условной оптимизации к решению серии задач безусловной оптимизации: из последовательности оптимальных планов этих задач или из их аппроксимаций будет формироваться итеративная последовательность для исходной задачи.

Отметим, что из предположений вогнутости и выпуклости функций f и f_i , $i \in 1:m$, следует их непрерывность на R^n (см. п. 1.8.5).

3.1. Внешние штрафные функции. Последовательность неотрицательных непрерывных на R^n функций $p_1, p_2, \dots$ называется

последовательностью внешних штрафных функций для множества $\mathcal{D}$, если $p_k(x) = 0$ для любого $x \in \mathcal{D}$, $p_{k+1}(x) > p_k(x)$, $k = 1, 2, \dots$, и $p_k(x) \rightarrow +\infty$ при $k \rightarrow \infty$ для любого $x \notin \mathcal{D}$ (рис. 38). Поскольку нас интересует теория выпуклого программирования, будем также предполагать выпуклость функций p_k на R^n при $k = 1, 2, \dots$ (необходимость такого предположения выяснится при описании метода в п. 3.2).

Последовательность выпуклых внешних штрафных функций для множества $\mathcal{D} = \{x \in R^n : f_i(x) \leq 0, i \in 1:m\}$ может быть построена следующим образом. Пусть $\{\alpha_k\}_{k=1}^\infty$ — строго возрастающая последовательность положительных чисел и $\alpha_k \rightarrow \infty$ при $k \rightarrow \infty$. Положим

$$p_k(x) = \alpha_k \sum_{i=1}^m [\max \{f_i(x), 0\}]^2, \quad k = 1, 2, \dots \quad (1)$$

Легко видеть, что формулы (1) определяют последовательность выпуклых внешних штрафных функций для $\mathcal{D}$, непрерывно дифференцируемых на R^n , если каждая из функций f_i , $i \in 1:m$, непрерывно дифференцируема на R^n .

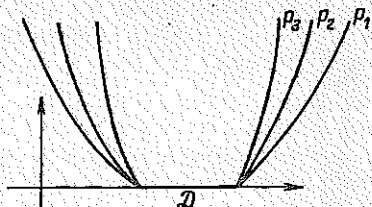


Рис. 38.

Действительно, так как $f_i(x) \leq 0$, $i \in 1:m$, для любого $x \in \mathcal{D}$, то $p_k(x) = 0$ при $k = 1, 2, \dots$ для любого $x \in \mathcal{D}$. Если $x \notin \mathcal{D}$, то найдется такое $i \in 1:m$, что $f_i(x) > 0$. Но тогда и $p_k(x) > 0$, поскольку $\alpha_k > 0$. Свойства $p_{k+1}(x) > p_k(x)$, $k = 1, 2, \dots$, и $p_k(x) \rightarrow +\infty$ при $k \rightarrow \infty$ и $x \notin \mathcal{D}$ следуют из свойств последовательности $\{\alpha_k\}_{k=1}^\infty$. Доказательство выпуклости функций p_k при условии выпуклости функций f_i , $i \in 1:m$, и непрерывной дифференцируемости функций p_k при условии непрерывной дифференцируемости функций f_i , $i \in 1:m$, мы оставляем читателю. — Упражнение 9.2.

Нетрудно проверить, что при выполнении указанных выше условий формулы

$$p_k(x) = \alpha_k \sum_{i=1}^m [\max \{f_i(x), 0\}]^{\beta}, \quad k = 1, 2, \dots,$$

определяют последовательность выпуклых непрерывно дифференцируемых внешних штрафных функций для $\mathcal{D}$ при любом $\beta \geq 2$. — Упражнение 9.3.

Далее понадобится следующая лемма.

Лемма. Пусть $\{p_k\}_{k=1}^\infty$ — последовательность внешних штрафных функций множества $\mathcal{D}$, $\{x_k\}_{k=1}^\infty$ — последовательность точек множества $R^n \setminus \mathcal{D}$, сходящаяся к точке x_0 , и пусть существует такое число $\alpha > 0$, что $p_k(x_k) \leq \alpha$ при $k = 1, 2, \dots$. Тогда $x_0 \in \mathcal{D}$.

Доказательство. Предположим, что $x_0 \in \mathcal{D}$. Тогда $p_k(x_0) \rightarrow +\infty$ при $k \rightarrow \infty$ и найдется такой номер k_0 , что $p_k(x_0) > \alpha$ при $k \geq k_0$. Из непрерывности функции p_{k_0} в точке x_0 следует, что найдется такой шар $\mathcal{D}(x_0, r)$, что $p_{k_0}(x) > \alpha$ для любого $x \in \mathcal{D}(x_0, r)$. Так как $x_k \rightarrow x_0$ при $k \rightarrow \infty$, то найдется такой номер k_1 , что $x_k \in \mathcal{D}(x_0, r)$ при $k \geq k_1$. Пусть $k > \max\{k_0, k_1\}$. Тогда $p_k(x_k) > p_{k_0}(x_k) > \alpha$, что противоречит условию $p_k(x_k) \leq \alpha$ при $k = 1, 2, \dots$ ■

3.2. О методе внешних штрафных функций. Идея метода внешних штрафных функций состоит в следующем. Используя некоторую последовательность штрафных функций, строят последовательность экстремальных задач выпуклого программирования $\{R^n, g_k\}_{k=1}^\infty$, в каждой из которых взимается «штраф», если ее решение не содержится в множестве $\mathcal{D}$, причем этот штраф неограниченно возрастает при $k \rightarrow \infty$. Эта система штрафов и принуждает стремиться к выполнению ограничений задачи $(\mathcal{D}, f)$. Так как для решения каждой из задач (R^n, g_k) можно воспользоваться одним из алгоритмов безусловной оптимизации, возникает метод решения задачи $(\mathcal{D}, f)$, который и называется *методом внешних штрафных функций*.

Перейдем к точным формулировкам. Пусть $(\mathcal{D}, f)$ — задача выпуклого программирования и $\{p_k\}_{k=1}^\infty$ — последовательность внешних штрафных функций (см. п. 3.1) для множества $\mathcal{D}$. Положим $g_k(x) = f(x) - p_k(x)$, $k = 1, 2, \dots$, и рассмотрим последовательность $\{(R^n, g_k)\}_{k=1}^\infty$ задач безусловной оптимизации — экстремальных задач максимизации выпуклого программирования (так как f — вогнутая, а p_k — выпуклая функции, то $g_k = f - p_k$ — вогнутая функция).

Предположим, что существует такая точка $x_0 \in R^n$, что множество $\hat{L}_{f(x_0)} = \{z \in R^n : f(z) \geq f(x_0)\}$ компактно. При этом предположении, если только $\mathcal{D} \neq \emptyset$, задача $(\mathcal{D}, f)$ имеет решение, и при любом $x \in \mathcal{D}$ компактное множество $\hat{L}_{f(x)}$ содержит множество оптимальных планов задачи $(\mathcal{D}, f)$ (см. теорему п. 2.2). Кроме того, при любом $x \in \mathcal{D}$ замкнутые множества $\{z \in R^n : f(z) - p_k(z) \geq f(x)\}$, $k = 1, 2, \dots$, содержатся в $\hat{L}_{f(x)}$ и, следовательно, компактны. Поэтому каждая из задач (R^n, g_k) , $k = 1, 2, \dots$, имеет решение, и при любом $x \in \mathcal{D}$ множество $\hat{L}_{f(x)}$ содержит множества оптимальных планов задач (R^n, g_k) , $k = 1, 2, \dots$

Содержание метода внешних штрафных функций определяет теорема, которая будет доказана в следующем пункте.

3.3. Теорема сходимости метода внешних штрафных функций. Установим сначала справедливость следующих двух утверждений.

Лемма 1. Пусть $\bar{x}_k$ — оптимальный план задачи (R^n, g_k) , $\bar{x}_{k+1}$ — оптимальный план задачи (R^n, g_{k+1}) (k — натуральное

число) и $\bar{x}$ — оптимальный план задачи $(\mathcal{D}, f)$. Тогда

$$g_k(\bar{x}_k) \geq g_{k+1}(\bar{x}_{k+1}) \geq f(\bar{x}).$$

Доказательство. Так как $\bar{x}_k$ — оптимальный план задачи (R^n, g_k) и $p_k(\bar{x}_{k+1}) \leq p_{k+1}(\bar{x}_{k+1})$, то

$$\begin{aligned} g_k(\bar{x}_k) &\geq g_k(\bar{x}_{k+1}) = f(\bar{x}_{k+1}) - p_k(\bar{x}_{k+1}) \geq \\ &\geq f(\bar{x}_{k+1}) - p_{k+1}(\bar{x}_{k+1}) = g_{k+1}(\bar{x}_{k+1}). \end{aligned}$$

Поскольку $\bar{x}_{k+1}$ — оптимальный план задачи (R^n, g_{k+1}) и $\bar{x} \in \mathcal{D}$,
 $g_{k+1}(\bar{x}_{k+1}) \geq g_{k+1}(\bar{x}) = f(\bar{x}) - p_{k+1}(\bar{x}) = f(\bar{x})$. ■

Лемма 2. Если задача $(\mathcal{D}, f)$ имеет решение и при некотором k_0 оптимальный план $\bar{x}_{k_0}$ задачи (R^n, g_{k_0}) принадлежит $\mathcal{D}$, то $\bar{x}_{k_0}$ — решение задачи $(\mathcal{D}, f)$, и при любом $k > k_0$ любой оптимальный план задачи (R^n, g_k) принадлежит $\mathcal{D}$ и является решением задачи $(\mathcal{D}, f)$.

Доказательство. Пусть $\bar{x}_{k_0} \in \mathcal{D}$ — оптимальный план задачи (R^n, g_{k_0}) и $\bar{x}$ — решение задачи $(\mathcal{D}, f)$. Тогда

$$g_{k_0}(\bar{x}_{k_0}) = f(\bar{x}_{k_0}) - p_{k_0}(\bar{x}_{k_0}) = f(\bar{x}_{k_0}) \leq f(\bar{x}).$$

Вместе с тем в силу леммы 1 $g_{k_0}(\bar{x}_{k_0}) \geq f(\bar{x})$. Значит, $f(\bar{x}_{k_0}) = f(\bar{x})$ и $\bar{x}_{k_0}$ — оптимальный план задачи $(\mathcal{D}, f)$. Пусть теперь $k > k_0$ и $\bar{x}_k$ — оптимальный план задачи (R^n, g_k) . Тогда если $\bar{x}_k \in \mathcal{D}$, то $p_{k_0}(\bar{x}_k) < p_k(\bar{x}_k)$ и

$$\begin{aligned} f(\bar{x}) &\leq g_k(\bar{x}_k) = f(\bar{x}_k) - p_k(\bar{x}_k) < f(\bar{x}_k) - p_{k_0}(\bar{x}_k) = \\ &= g_{k_0}(\bar{x}_k) \leq g_{k_0}(\bar{x}_{k_0}) = f(\bar{x}), \end{aligned}$$

что невозможно. Следовательно, $\bar{x}_k \in \mathcal{D}$ и $\bar{x}_k$ — решение задачи $(\mathcal{D}, f)$. ■

Предположим, что множество $\mathcal{D} \neq \emptyset$ и при некотором $x_0 \in R^n$ множество $\hat{L}_{f(x_0)}$ компактно. Рассмотрим последовательность $\{(R^n, g_k)\}_{k=1}^{\infty}$ экстремальных задач безусловной оптимизации, порожденных экстремальной задачей выпуклого программирования $(\mathcal{D}, f)$ и последовательностью $\{p_k\}_{k=1}^{\infty}$ выпуклых внешних штрафных функций для множества $\mathcal{D}$. Как мы уже отмечали (п. 3.2) в предположении компактности множества $\hat{L}_{f(x_0)}$, задача $(\mathcal{D}, f)$ и каждая из задач (R^n, g_k) , $k = 1, 2, \dots$, имеют решение и оптимальные планы этих задач содержатся в множестве $\hat{L}_{f(x_0)}$, если $x_0 \in \mathcal{D}$. Пусть $\{\bar{x}_k\}$ — последовательность оптимальных планов задач (R^n, g_k) , $k = 1, 2, \dots$. Тогда верна следующая теорема.

Теорема. Либо найдется такое k_0 , что $\bar{x}_{k_0} \in \mathcal{D}$, и в этом случае $\bar{x}_{k_0}$ — решение задачи $(\mathcal{D}, f)$, либо $\bar{x}_k \in R^n \setminus \mathcal{D}$ при

$k = 1, 2, \dots$; тогда последовательность $\{\bar{x}_k\}_{k=1}^{\infty}$ содержит подпоследовательность, сходящуюся к оптимальному плану задачи $(\mathcal{D}, f)$, и всякая сходящаяся подпоследовательность последовательности $\{\bar{x}_k\}_{k=1}^{\infty}$ сходится к оптимальному плану задачи $(\mathcal{D}, f)$.

Если задача $(\mathcal{D}, f)$ имеет единственное решение, то последовательность $\{\bar{x}_k\}_{k=1}^{\infty}$ сходится к этому решению.

Доказательство. Мы уже видели, что если $\bar{x}_{k_0} \in \mathcal{D}$ при некотором k_0 , то $\bar{x}_{k_0}$ — решение задачи $(\mathcal{D}, f)$ (см. лемму 2). Поэтому рассмотрим случай, когда $\bar{x}_k \in R^n \setminus \mathcal{D}$ при $k = 1, 2, \dots$. Так как все точки $\bar{x}_k$ принадлежат компактному множеству $\hat{L}_{f(x_0)}$ (при некотором $x_0 \in \mathcal{D}$), то существует подпоследовательность $\{\bar{x}_{k_i}\}_{i \in K}$, сходящаяся к $\bar{x}$. Докажем, что $\bar{x}$ — решение задачи $(\mathcal{D}, f)$.

Положим $\gamma = \max_{x \in \hat{L}_{f(x_0)}} f(x)$ (так как функция f непрерывна,

то по теореме Вейерштрасса такое число γ существует). Обозначим через $\bar{z}$ какое-нибудь решение задачи $(\mathcal{D}, f)$. В силу леммы 1 $g_k(\bar{x}_k) = f(\bar{x}_k) - p_k(\bar{x}_k) \geq f(\bar{z})$, так что $p_k(\bar{x}_k) \leq f(\bar{x}_k) - f(\bar{z}) \leq \gamma - f(\bar{z})$ при $k = 1, 2, \dots$. Следовательно, существует такое $\alpha > 0$, что $p_k(\bar{x}_k) \leq \alpha$ при любом $k \in K$, и в силу леммы п. 3.1 $\bar{x} \in \mathcal{D}$. Так как $f(\bar{x}_k) - p_k(\bar{x}_k) \geq f(\bar{z})$ при любом $k \in K$ и $p_k(\bar{x}_k) > 0$, то $f(\bar{x}_k) > f(\bar{z})$, $k \in K$, и из непрерывности f следует $f(\bar{x}) \geq f(\bar{z})$. Но $\bar{x} \in \mathcal{D}$, а $\bar{z}$ — решение задачи $(\mathcal{D}, f)$. Значит, $f(\bar{x}) = f(\bar{z})$ и $\bar{x}$ — решение задачи $(\mathcal{D}, f)$.

Следствие. Положим $\bar{y} = \max_{x \in \mathcal{D}} f(x)$. Числовая последовательность $\{g_k(\bar{x}_k)\}_{k=1}^{\infty}$ монотонно сходится к $\bar{y}$ при $k \rightarrow \infty$.

Доказательство. В силу леммы 1 $\{g_k(\bar{x}_k)\}_{k=1}^{\infty}$ — невозрастающая последовательность вещественных чисел и достаточно доказать, что некоторая ее подпоследовательность сходится к $\bar{y}$. Утверждение теоремы следует из леммы 2, если $\bar{x}_{k_0} \in \mathcal{D}$ при некотором k_0 . Пусть $\bar{x}_k \in R^n \setminus \mathcal{D}$ при $k = 1, 2, \dots$. Тогда в силу теоремы 2 существует подпоследовательность $\{\bar{x}_{k_i}\}_{i \in K}$, сходящаяся к решению $\bar{x}$ задачи $(\mathcal{D}, f)$ при $k \rightarrow \infty$, $k \in K$. Действительно, в силу леммы 1 $g_k(\bar{x}_k) \geq f(\bar{x})$ при $k = 1, 2, \dots$, и так как функции p_k неотрицательны, то

$$0 \leq g_k(\bar{x}_k) - f(\bar{x}) = f(\bar{x}_k) - p_k(\bar{x}_k) - f(\bar{x}) \leq f(\bar{x}_k) - f(\bar{x}).$$

Но в силу непрерывности функции f

$$f(\bar{x}_k) \rightarrow f(\bar{x}) \text{ при } k \rightarrow \infty, \quad k \in K. \blacksquare$$

Замечание. Итеративные методы решения задач безусловной оптимизации, вообще говоря, не обеспечивают построения точного решения таких задач за конечное число шагов. Поэтому построение последовательности $\bar{x}_1, \bar{x}_2, \dots$ практически не-

осуществимо. Однако для решения задачи $(\mathcal{D}, f)$ методом внешних штрафных функций нет необходимости строить эту последовательность: достаточно ограничиться построением приближенных решений задач (R^n, g_k) , $k = 1, 2, \dots$

Действительно, зададимся какой-нибудь последовательностью $\{\varepsilon_k\}_{k=1}^\infty$ положительных чисел, стремящейся к нулю при $k \rightarrow \infty$, и будем строить такие векторы $\bar{x}_1, \bar{x}_2, \dots$, что $\|\bar{x}_k - \bar{x}\| \leq \varepsilon_k$, $k = 1, 2, \dots$. Тогда очевидно, что если $\{\bar{x}_k\}_{k \in K}$ — подпоследовательность последовательности $\{\bar{x}_k\}_{k=1}^\infty$, сходящаяся к решению $\bar{x}$ задачи $(\mathcal{D}, f)$, то и подпоследовательность $\{\bar{x}_k\}_{k \in K}$ сходится к тому же вектору $\bar{x}$. Кроме того (см. следствие), последовательность $\{g_k(\bar{x}_k)\}_{k=1}^\infty$ очевидно будет сходиться к $\bar{y} = \max_{x \in \mathcal{D}} f(x)$.

Другая возможность состоит в построении такой последовательности векторов $\{\hat{x}_k\}_{k=1}^\infty$, что $|g_k(\hat{x}_k) - g_k(\bar{x}_k)| < \varepsilon_k$ для $k = 1, 2, \dots$. Тогда очевидно, что $\lim_{k \rightarrow \infty} g_k(\hat{x}_k) = \lim_{k \rightarrow \infty} g_k(\bar{x}_k) = \bar{y}$. Кроме того, если $\{\hat{x}_k\}_{k \in K}$ — подпоследовательность последовательности $\{\hat{x}_k\}_{k=1}^\infty$, сходящаяся к вектору $\hat{x}$, то $\hat{x} \in \mathcal{D}$ и $f(\hat{x}) = \bar{y}$. Доказательство последнего утверждения мы оставляем читателю. — Упражнение 9.4.

3.4. Внутренние штрафные функции. Метод внутренних штрафных функций, или, как его иначе называют, метод барьеров, о котором пойдет речь, тоже сводит решение задачи условной оптимизации $(\mathcal{D}, f)$ к решению серии задач безусловной оптимизации. Однако сфера его применения уже сферы применения метода внешних штрафных функций: в частности, это проявляется в том, что возникают трудности в применении этого метода к задачам, для которых $\dim \mathcal{D} < n$. Для таких задач метод внутренних штрафных функций часто применяют в комбинации с другими вычислительными алгоритмами.

Ниже будут рассматриваться только такие задачи выпуклого программирования $(\mathcal{D}, f)$, для которых $\dim \mathcal{D} = n$. В силу теоремы п. 1.3.6 это условие эквивалентно тому, что $\mathcal{D}^0 \neq \emptyset$, или, что то же, выполнению условия С (см. п. 3.1.5) для всех ограничений, определяющих множество $\mathcal{D}$. Кроме того, как и в § 2, будем предполагать, что множество $\hat{L}_{f(x_0)} = \{x \in R^n : f(x) \geq f(x_0)\}$ ограничено при $x_0 \in R^n$ (см. лемму и теорему п. 2.2) и, следовательно, в силу непрерывности функции f компактно в R^n . При этом предположении задача $(\mathcal{D}, f)$ имеет решение, и множество ее оптимальных планов содержится в множестве $L_{f(x)}$ при любом $x \in \mathcal{D}$.

Последовательность $\{q_k\}_{k=1}^\infty$ положительных замкнутых (см. п. 1.8.4) выпуклых функций $q_k: R^n \rightarrow \bar{R}$ назовем последовательностью *внутренних штрафных функций* для выпуклого множества $\mathcal{D}$, $\dim \mathcal{D} = n$, если при каждом $k = 1, 2, \dots$ $\text{dom } q_k = \mathcal{D}^0$,

$q_k(x) > q_{k+1}(x)$ для любого $x \in \mathcal{D}^0$ и $q_k(x) \rightarrow 0$ при $k \rightarrow \infty$ для любого $x \in \mathcal{D}^0$ (рис. 39). Из условия $\text{dom } q_k = \mathcal{D}^0$ и замкнутости функции q_k следует, что $q_k(x) = +\infty$ для любого $x \in \mathcal{D}^0$ и $q_k(x_m) \rightarrow +\infty$ при $m \rightarrow \infty$, если $\{x_m\}_{m=1}^\infty$ — последовательность точек из $\mathcal{D}^0$, сходящаяся к граничной точке множества $\mathcal{D}$. Кроме того, каждая из функций q_k непрерывна на $\mathcal{D}^0$ (п. 1.8.5).

Последовательность $\{q_k\}_{k=1}^\infty$ внутренних штрафных функций для множества $\mathcal{D} = \{x \in R^n: f_i(x) \leq 0, i \in 1:m\}$, где при каждом $i \in 1:m$ функция $f_i: R^n \rightarrow R$ выпукла, может быть построена следующим образом. Для $i \in 1:m$ положим $h_i(x) = -1/f_i(x)$, если $f_i(x) < 0$, и $h_i(x) = +\infty$, если $f_i(x) \geq 0$. Пусть $\{\alpha_k\}_{k=1}^\infty$ — убывающая последовательность положительных чисел, стремящаяся к нулю при $k \rightarrow \infty$. Тогда

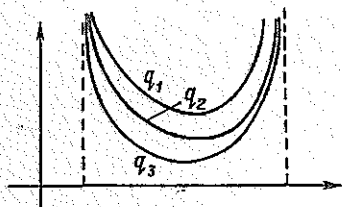


Рис. 39.

последовательность функций $\{q_k\}_{k=1}^\infty$, где $q_k(x) = \alpha_k \sum_{i=1}^m h_i(x)$, есть последовательность внутренних штрафных функций для выпуклого множества $\mathcal{D}$, если $\dim \mathcal{D} = n$.

В самом деле, нетрудно показать, что h_i — замкнутая выпуклая функция из R^n в $\bar{R}$. —

Упражнение 9.5. Следовательно,

q_k — положительная замкнутая выпуклая функция, отличная от $+\infty$ только в таких точках $x \in R^n$, для которых $f_i(x) < 0$ при всех $i \in 1:m$, т. е. в точках $\mathcal{D}^0$. Строгая монотонность последовательности функций $\{q_k\}_{k=1}^\infty$ на $\mathcal{D}^0$ и свойство $q_k(x) \rightarrow 0$ при $k \rightarrow \infty$ для $x \in \mathcal{D}^0$ следуют из свойств последовательности $\{\alpha_k\}_{k=1}^\infty$.

3.5. Метод внутренних штрафных функций. Пусть $(\mathcal{D}, f) \rightarrow$ задача выпуклого программирования, $\dim \mathcal{D} = n$ и множество $\hat{L}_{f(x_0)}$ ограничено при некотором $x_0 \in R^n$. Пусть $\{q_k\}_{k=1}^\infty$ — последовательность внутренних штрафных функций для множества $\mathcal{D}$. Для каждого $k = 1, 2, \dots$ построим функцию $g_k: R^n \rightarrow \bar{R}$, $g_k(x) = f(x) - q_k(x)$. Так как f — вогнутая, а q_k — выпуклая функции, то g_k — вогнутая функция на R^n и $\text{dom } g_k = \mathcal{D}^0$. Рассмотрим последовательность $\{(R^n, g_k)\}_{k=1}^\infty$ экстремальных задач максимизации вогнутых функций на R^n .

Пусть $x_0 \in \mathcal{D}^0$. Так как $\hat{L}_{f(x_0)}$ непустое компактное множество в R^n , то при любом $k = 1, 2, \dots$ множество $\hat{L}_{g_k(x_0)} = \{x \in R^n: f(x) \geq g_k(x_0)\}$ компактно (см. лемму п. 2.2). Из положительности функций q_k на $\mathcal{D}^0$ следует, что при любом $k = 1, 2, \dots$ множество $\hat{L}_{g_k(x_0)}$ содержит непустое замкнутое

выпуклое множество $\hat{M}_{g_k(x_0)} = \{x \in R^n : g_k(x) \geq g_k(x_0)\} = \{x \in R^n : f(x) - q_k(x) \geq f(x_0) - q_k(x_0)\}$. Более того, так как $q_k(x_0) < q_1(x_0)$ при $k = 1, 2, \dots$, то $\hat{M}_{g_k(x_0)} \subset \hat{L}_{g_1(x_0)}$ при $k = 1, 2, \dots$. Следовательно, каждое из множеств $\hat{M}_{g_k(x_0)}$ компактно. Очевидно также, что при любом k множество $M_{g_k(x_0)} \subset \mathcal{D}^0$: вне $\mathcal{D}^0$ каждая из функций g_k принимает значение $-\infty$. Поэтому каждая из задач (R^n, g_k) , $k = 1, 2, \dots$, имеет решение (оптимальный план), принадлежащее множеству $\mathcal{D}^0$ (теорема п. 2.2). Если функции f и q_k непрерывно дифференцируемы на $\mathcal{D}^0$, то для решения этих задач могут быть применены итеративные методы решения задач безусловной оптимизации. Важно только, чтобы построение итеративных последовательностей начиналось с некоторой точки множества $\mathcal{D}^0$. Задачу нахождения такой точки обсудим в следующем параграфе (п. 4.7).

Обозначим через $\{\bar{x}_k\}_{k=1}^\infty$ последовательность оптимальных планов задач (R^n, g_k) , $k = 1, 2, \dots$, и пусть $\bar{y} = \max_{x \in \mathcal{D}} f(x)$. Так как $q_1(x) > q_2(x) > \dots > 0$ при любом $x \in \mathcal{D}^0$, то $g_1(\bar{x}_1) < g_2(\bar{x}_2) < \dots < \bar{y}$.

Теорема. *Последовательность $\{\bar{x}_k\}_{k=1}^\infty$ содержит подпоследовательность, сходящуюся к оптимальному плану задачи $(\mathcal{D}, f)$, и всякая сходящаяся подпоследовательность последовательности $\{\bar{x}_k\}_{k=1}^\infty$ сходится к оптимальному плану задачи $(\mathcal{D}, f)$. Кроме того, последовательность $\{g_k(\bar{x}_k)\}_{k=1}^\infty$ монотонно сходится к $\bar{y}$.*

Если задача $(\mathcal{D}, f)$ имеет единственное решение, то последовательность $\{\bar{x}_k\}_{k=1}^\infty$ сходится к этому решению.

Доказательство. Последовательность $\{\bar{x}_k\}_{k=1}^\infty$ содержится в компактном множестве $\hat{L}_{g_1(x_0)}$ (при некотором $x_0 \in \mathcal{D}^0$) и, следовательно, содержит подпоследовательность $\{\bar{x}_k\}_{k \in K}$, сходящуюся к $\bar{x}$ при $k \rightarrow \infty$, $k \in K$. Так как $\bar{x}_k \in \mathcal{D}^0$ при $k = 1, 2, \dots$ и множество $\mathcal{D}$ замкнуто, то $\bar{x} \in \mathcal{D}$. Покажем, что $\bar{x}$ — решение задачи $(\mathcal{D}, f)$ и $g_k(\bar{x}_k) \rightarrow f(\bar{x}) = \bar{y}$ при $k \rightarrow \infty$, $k \in K$. Тогда в силу монотонности последовательности $\{g_k(\bar{x}_k)\}_{k=1}^\infty$ можно заключить, что $g_k(\bar{x}_k) \rightarrow f(\bar{x})$ при $k \rightarrow \infty$.

Покажем сначала, что $g_k(\bar{x}_k) \rightarrow f(\bar{x})$ при $k \rightarrow \infty$, $k \in K$. Пусть $\varepsilon > 0$. Так как $\bar{x}_k \rightarrow \bar{x}$ при $k \rightarrow \infty$, $k \in K$, то в силу непрерывности функции f найдется такой номер $k_0 \in K$, что $|f(\bar{x}) - f(\bar{x}_k)| < \varepsilon/2$ при $k \geq k_0$, $k \in K$. Точка $\bar{x}_{k_0} \in \mathcal{D}^0$, и в соответствии с определением последовательности внутренних штрафных функций найдется такое $k_1 \geq k_0$, что $q_k(\bar{x}_{k_0}) < \varepsilon/2$ для любых $k \geq k_1$. Пусть $k \geq k_1$, $k \in K$. Так как $\bar{x}_k$ — решение за-

дачи (R^n, g_k) , то $g_k(\bar{x}_k) \geq g_k(\bar{x}_{k_0})$,

$$f(\bar{x}) - g_k(\bar{x}_k) \leq f(\bar{x}) - g_k(\bar{x}_{k_0}) = f(\bar{x}) - f(\bar{x}_{k_0}) + q_k(\bar{x}_{k_0}) < \varepsilon,$$

$$f(\bar{x}) - g_k(\bar{x}_k) = f(\bar{x}) - f(\bar{x}_k) + q_k(\bar{x}_k) > f(\bar{x}) - f(\bar{x}_k) > -\varepsilon/2.$$

Следовательно, $|f(\bar{x}) - g_k(\bar{x}_k)| < \varepsilon$ при $k \geq k_1$, $k \in K$.

Покажем теперь, что $\bar{x}$ — решение задачи $(\mathcal{D}, f)$. Предположим противное. Тогда существует такая точка $\bar{z} \in \mathcal{D}$, что $f(\bar{z}) - f(\bar{x}) = \Delta > 0$. Так как $\mathcal{D} = \bar{\mathcal{D}}^0$ (см. теорему п. 1.3.8), то найдется последовательность $\{x_k\}_{k \in K}$ точек множества $\mathcal{D}^0$, сходящаяся к $\bar{z}$ при $k \rightarrow \infty$, $k \in K$. В силу непрерывности f найдется такой номер $k_0 \in K$, что $|f(\bar{z}) - f(x_k)| < \Delta/4$ при $k \geq k_0$, $k \in K$. Выберем также такой номер $k_1 > k_0$, что $q_k(x_{k_0}) < \Delta/4$ при $k \geq k_1$. Поскольку $\bar{x}_k$ — решение задачи (R^n, g_k) , то при $k > k_1$, $k \in K$,

$$\begin{aligned} g_k(\bar{x}_k) &\geq g_k(x_{k_0}) = f(x_{k_0}) - q_k(x_{k_0}) > \\ &> f(\bar{z}) - \Delta/4 - \Delta/4 = f(\bar{x}) + \Delta/2, \end{aligned}$$

что невозможно, если последовательность $\{g_k(\bar{x}_k)\}_{k \in K}$ сходится к $f(\bar{x})$ при $k \rightarrow \infty$, $k \in K$. ■

В заключение заметим, что для решения задачи методом внутренних штрафных функций нет необходимости в точном вычислении векторов $\bar{x}_k$. Здесь можно дословно повторить все сказанное по этому поводу в замечании к п. 3.3.

§ 4. Методы допустимых направлений

Методы допустимых направлений мы уже рассматривали в § 2 гл. 4. Там же обсуждались идеи конечных алгоритмов решения задач линейного программирования, основанные на этих методах.

Идея построения итеративной последовательности, связанная с выбором допустимого прогрессивного направления, достаточно ясна. Но конкретная реализация этой идеи, приводящая к построению итеративной последовательности, сходящейся к решению задачи, совсем нетривиальна. Эта реализация впервые была осуществлена Г. Зойтендейком [15]. Впоследствии методы допустимых направлений для решения экстремальных задач в той или иной форме стали предметом многочисленных исследований. В настоящем параграфе мы ставим своей задачей ознакомление читателя с главной идеей методов. Поэтому здесь будет рассмотрен только один такой алгоритм решения задачи выпуклого программирования, укладывающийся в схему п. 1.4, и доказана его сходимости.

4.1. Общие замечания. Рассмотрим задачу выпуклого программирования (см. п. 2.2.1 или начало гл. 3)

$$(\mathcal{D}, f) : f(x) \rightarrow \max, \quad \mathcal{D} = \{x \in R^n : f_i(x) \leq 0, i \in I_n;\} \\ f_i(x) = a_i x - b[i] = 0, i \in I_y; x[j] \geq 0, j \in J\}.$$

Предположим, что f и $f_i, i \in I_n$, — непрерывно дифференцируемые на R^n функции и для множества ограничений $\{f_i(x) \leq 0, i \in I_n\}$ выполнено условие регулярности С (см. п. 3.1.5). Кроме того, предположим, что для некоторой точки $x_0 \in R^n$ множество $\tilde{L}_{f(x_0)} = \{x \in R^n : f(x) \geq f(x_0)\}$ компактно. В силу теоремы п. 2.2 множество $L_{f(x)}$ компактно для любого $x \in R^n$ ($x \in \mathcal{D}$).

Как мы уже отмечали, построение итеративной последовательности $x_1, x_2, \dots$ будет осуществлено в соответствии со схемой п. 1.4. Поэтому для описания алгоритма нам следует определить множество приемлемых точек π , корректирующее отображение $\text{cor} : \mathcal{D} \rightarrow \mathcal{D}$ и оценочную функцию p . Затем для доказательства сходимости итеративной последовательности в смысле утверждения п. 1.4 установим выполнение условий утверждения этого пункта. Особо будет рассмотрена задача нахождения начального допустимого плана x_1 . В качестве оценочной функции возьмем целевую функцию f задачи $(\mathcal{D}, f)$.

4.2. Множество приемлемых точек. Пусть $x \in \mathcal{D}$, $I(x) = \{i \in I_n : f_i(x) = 0\}$, $J(x) = \{j \in J : x[j] = 0\}$ (см. п. 3.1.3). Рассмотрим следующую задачу линейного программирования:

$$(S(x), g) : g(s, z) = z \rightarrow \max,$$

$$S(x) = \{(s, z) \in R^{n+1}, s \in R^n, z \in R:$$

$$\forall f_i(x) s + z \leq 0, i \in I(x) \cap I_n, \quad (1)$$

$$\forall f_i(x) s = a_i s \leq 0, i \in I(x) \setminus I_n, \quad (2)$$

$$\forall f_i(x) s = a_i s = 0, i \in I_y, \quad (3)$$

$$s[j] \geq 0, j \in J(x), \quad (4)$$

$$-1 \leq s[j] \leq 1, j \in 1:n, \quad (5)$$

$$z \leq \nabla f(x) s\}. \quad (6)$$

При любом $x \in \mathcal{D}$ задача $(S(x), g)$ имеет решение. Действительно, множество $S(x) \neq \emptyset$, так как вектор $0 \in R^{n+1}$ очевидно принадлежит $S(x)$. Кроме того, из (5) и (6) следует, что $\|s\| \leq \sqrt{n}$ и $g(s, z) = z \leq \nabla f(x) s \leq \|\nabla f(x)\| \cdot \|s\| \leq \sqrt{n} \|\nabla f(x)\|$ для любого вектора $(s, z) \in S(x)$, т. е. целевая функция g задачи $(S(x), g)$ ограничена сверху на $S(x)$. Следовательно, задача $(S(x), g)$ имеет решение (см. п. 3.3.6, 5°, а также доказательство теоремы п. 3.3.7). ■

Условия (5), которые вместе с (6) обеспечивают существование решения задачи $(S(x), g)$, называются условиями нор-

мализации. Можно предложить условия нормализации, отличные от (5).

Обозначим через $(\bar{s}(x), \bar{z}(x))$ какое-нибудь решение задачи $(S(x), g)$. Наибольшее значение $\bar{z}(x)$ функции g на $S(x)$ определяется однозначно, и очевидно, что $\bar{z}(x) \geq 0$. Точку $x \in \mathcal{D}$ будем считать приемлемой в том и только том случае, если $\bar{z}(x) = 0$, т. е. положим $\pi = \{x \in \mathcal{D} : \bar{z}(x) = 0\}$. Мы сейчас установим, что определенное так множество π приемлемых точек задачи $(\mathcal{D}, f)$ фактически совпадает с множеством ее решений.

Теорема. Вектор $\bar{x} \in \mathcal{D}$ в том и только том случае является оптимальным планом задачи $(\mathcal{D}, f)$, если $\bar{z}(\bar{x}) = 0$.

Доказательство. Необходимость. Пусть $\bar{x}$ — оптимальный план задачи $(\mathcal{D}, f)$. Покажем, что $\bar{z}(\bar{x}) = 0$. Предположим противное: пусть $\bar{z}(\bar{x}) > 0$. Тогда $0 < \bar{z}(\bar{x}) \leq \forall f(\bar{x}) \bar{s}(\bar{x})$, а следовательно, $\bar{s}(\bar{x}) \neq 0$ и $\bar{s}(\bar{x})$ — прогрессивное направление в точке $\bar{x}$ (см. п. 3.1.4). Кроме того, для вектора $\bar{s}(\bar{x})$ выполняются условия (5) — (8) п. 3.1.3 при $x = \bar{x}$, и, следовательно, $\bar{s}(\bar{x})$ — допустимое направление в точке $\bar{x}$. Но тогда (см. п. 3.1.2) $\bar{x}$ не является оптимальным планом задачи $(\mathcal{D}, f)$. ■

Достаточность. Предположим, что x не является оптимальным планом задачи $(\mathcal{D}, f)$. Тогда (см. п. 3.1.2) существует допустимое прогрессивное направление s в точке x . Так как для задачи $(\mathcal{D}, f)$ выполняется условие регулярности C , то существует такой вектор $\hat{x} \in \mathcal{D}$, что $f_i(\hat{x}) < 0$, $i \in I_x$, и в соответствии с замечанием 2 п. 3.1.6 вектор $s(\lambda) = s + \lambda(\hat{x} - x)$ удовлетворяет условиям (5'') — (8'') п. 3.1.6 при любом $\lambda > 0$. Кроме того, $\nabla f(x)s > 0$, и, следовательно, найдется такое $\bar{\lambda} > 0$, что $\nabla f(x)s(\bar{\lambda}) = \nabla f(x)s + \bar{\lambda}\nabla f(x)(\hat{x} - x) > 0$. Очевидно существует такое положительное число α , что для вектора $s' = (1/\alpha)s(\bar{\lambda})$ выполняются условия (5). Положим $z' = \min\{-\nabla f_i(x)s', i \in I(x) \cap I_x, \nabla f_i(x)s'\}$. Очевидно $(s', z') \in S(x)$ и $z' > 0$. Но тогда и $\bar{z}(x) > 0$. Следовательно, если $\bar{z}(\bar{x}) = 0$, то $\bar{x}$ — оптимальный план задачи $(\mathcal{D}, f)$. ■

4.3. σ -активные множества индексов. Основная идея Г. Зойтендейка состоит в том, что при построении допустимого прогрессивного направления в точке $x \in \mathcal{D}$ используют не только те граничные гиперповерхности, на которых находится точка x , но и «близкие» к этой точке гиперповерхности, ограничивающие множество $\mathcal{D}$. Этот технический прием и обеспечивает сходимость алгоритма.

Пусть $x \in \mathcal{D}$ и $\sigma \geq 0$. Рассмотрим множества

$$I(x, \sigma) = \{i \in I_x : -\sigma \leq f_i(x) \leq 0\},$$

$$J(x, \sigma) = \{j \in J : 0 \leq x[j] \leq \sigma\}.$$

Множества $I(x, \sigma)$ и $J(x, \sigma)$ называются σ -активными множествами индексов задачи $(\mathcal{D}, f)$ в точке x . Установим несколько утверждений, связанных с этими понятиями.

1°. $I(x, 0) = I(x)$, $J(x, 0) = J(x)$ при любом $x \in \mathcal{D}$. Утверждение очевидно. ■

2°. Если $\sigma_1 \leq \sigma_2$, то $I(x, \sigma_1) \subset I(x, \sigma_2)$, $J(x, \sigma_1) \subset J(x, \sigma_2)$ для любого $x \in \mathcal{D}$. Доказательство очевидно. ■

3°. Для любого $x \in \mathcal{D}$ существует такое $\sigma' > 0$, что $I(x, \sigma) = I(x)$, $J(x, \sigma) = J(x)$ при любом $\sigma \in [0, \sigma']$. Действительно, так как $\sigma' = \min\{-f_i(x), i \in I_n \setminus I(x); x[j], j \in J \setminus J(x)\} > 0$, то $I(x, \sigma) = I(x)$, $J(x, \sigma) = J(x)$, если $0 \leq \sigma < \sigma'$. ■

4°. Пусть $x \in \mathcal{D}$ и $\sigma > 0$. Существует такое число $r > 0$, что $I(x', \sigma) \subset I(x, \sigma)$, $J(x', \sigma) \subset J(x, \sigma)$ для любого $x' \in \bar{\mathcal{D}}(x, r) \cap \mathcal{D}$ ($\bar{\mathcal{D}}(x, r)$ — замкнутый шар с центром в точке x и радиусом r).

Доказательство. Так как $f_i(x) < -\sigma$ при каждом $i \in I_n \setminus I(x, \sigma)$, то в силу непрерывности функций f_i найдется такое число $r' > 0$, что $f_i(x') < -\sigma$ при любом $i \in I_n \setminus I(x, \sigma)$ для всех $x' \in \bar{\mathcal{D}}(x, r') \cap \mathcal{D}$. Аналогично если $j \in J \setminus J(x, \sigma)$, то $x[j] > \sigma$ и найдется такое $r'' > 0$, что $x'[j] > \sigma$ при любом $j \in J \setminus J(x, \sigma)$ для всех $x' \in \bar{\mathcal{D}}(x, r'') \cap \mathcal{D}$. Положим $r = \min\{r', r''\}$. Тогда если $x' \in \bar{\mathcal{D}}(x, r) \cap \mathcal{D}$, $i \in I_n$, и $f_i(x') \geq -\sigma$, то $i \in I(x, \sigma)$, т. е. $I(x', \sigma) \subset I(x, \sigma)$. Аналогично $J(x', \sigma) \subset J(x, \sigma)$, если $x' \in \bar{\mathcal{D}}(x, r) \cap \mathcal{D}$. ■

Пусть $x \in \mathcal{D}$ и $\sigma \geq 0$. Положим

$$S(x, \sigma) = \{(s, z) \in R^{n+1} : \nabla f_i(x)s + z \leq 0, i \in I(x, \sigma) \cap I_n,$$

$$a_i s \leq 0, i \in I(x, \sigma) \setminus I_n, a_i s = 0, i \in I_y, s[j] \geq 0,$$

$$j \in J(x, \sigma), -1 \leq s[j] \leq 1, j \in 1:n, z \leq \nabla f(x)s\}.$$

Очевидно, что при любом $x \in \mathcal{D}$ и $\sigma \geq 0$ множество $S(x, \sigma) \neq \emptyset$ и задача линейного программирования $(S(x, \sigma), g) : g(s, z) = z \rightarrow \max, (s, z) \in S(x, \sigma)$ имеет решение (ср. с таким же утверждением п. 4.2). Обозначим какое-нибудь решение задачи $(S(x, \sigma), g)$ через $(\bar{s}(x, \sigma), \bar{z}(x, \sigma))$. Из утверждений 1° и 2° следует, что $S(x, 0) = S(x)$ и $S(x, \sigma') \supset S(x, \sigma'')$, если $\sigma' \leq \sigma''$. Поэтому $\bar{z}(x, \sigma) \leq \bar{z}(x)$ для любого $x \in \mathcal{D}$ и любого $\sigma \geq 0$, и $\bar{z}(x, \sigma') \geq \bar{z}(x, \sigma'')$, если $\sigma' \leq \sigma''$.

Возьмем число $\sigma > 0$. Для каждого $m = 0, 1, 2, \dots$ положим $\sigma_m = (1/2^m)\sigma$.

5°. Если точка $x \in \mathcal{D}$ не является приемлемой (т. е. если точка $x \in \mathcal{D}$ не является решением задачи $(\mathcal{D}, f)$, см. теорему п. 4.2), то найдется такой номер $\bar{m}$, что $\bar{z}(x, \sigma_m) \geq \sigma_m$ для всех $m \geq \bar{m}$.

Доказательство очевидно следует из утверждения 3°. Действительно, найдется такой номер m' , что $I(x, \sigma_m) = I(x)$ и $J(x, \sigma_m) = J(x)$ для $m \geq m'$. Но тогда $S(x, \sigma_m) = S(x)$ и $\bar{z}(x, \sigma_m) = \bar{z}(x) > 0$ при $m \geq m'$. Поэтому если возьмем такое $\bar{m}$, что $\bar{m} \geq m'$ и $\sigma_{\bar{m}} \leq \bar{z}(x)$, то $\bar{z}(x, \sigma_m) \geq \sigma_m$ при $m \geq \bar{m}$. ■

Наименьший из таких номеров $\bar{m}$, что $\bar{z}(x, \sigma_m) \geq \sigma_m$ для $m \geq \bar{m}$, обозначим через $m(x)$.

4.4. Корректирующее отображение. В дальнейшем будем считать число σ фиксированным. Построение корректирующего отображения состоит из двух этапов: выбора допустимого прогрессивного направления и выбора шага.

Выбор допустимого прогрессивного направления. Пусть $x \in \mathcal{D}$. Решим задачу линейного программирования $(S(x), g)$. Если $\bar{z}(x) = 0$, то положим $\text{сог } x = x$. Иначе если $\bar{z}(x) > 0$, последовательно решаем задачи $(S(x, \sigma), g)$, $(S(x_1, \sigma_1), g)$, ... За конечное число шагов (см. утверждение 5° п. 4.3) обнаружим такой номер $m(x)$, что $\bar{z}(x, \sigma_{m(x)}) \geq \sigma_{m(x)}$. Вектор $\bar{s}(x, \sigma_{m(x)})$ является допустимым прогрессивным направлением в точке x .

Выбор шага. После построения вектора $\bar{s}(x, \sigma_{m(x)})$ решаем задачу одномерной оптимизации: выбираем число $\lambda(x)$ — наименьшее из таких чисел $\lambda > 0$, что

$$f(x + \lambda \bar{s}(x, \sigma_{m(x)})) = \max_{0 < \mu \leq \mu'} \{f(x + \mu \bar{s}(x, \sigma_{m(x)}))\},$$

где μ' — наибольшее из положительных чисел μ , при которых $x + \mu \bar{s}(x, \sigma_{m(x)}) \in \mathcal{D}$. Существование числа $\lambda(x)$, как было установлено в п. 2.3, следует из компактности множества $L_{f(x)}$.

Полагаем $\text{сог } x = x + \lambda(x) \bar{s}(x, \sigma_{m(x)})$. Если $\bar{z}(x) > 0$, то $f(\text{сог } x) > f(x)$.

Задача выбора направления $(S(x, \sigma_{m(x)}), g)$ может иметь неединственное решение. Поэтому фактически определяется множество $\text{Сог } x$ (см. п. 1.4), состоящее из векторов $x + \lambda(x) \bar{s}(x, \sigma_{m(x)})$ ($\lambda(x)$ очевидно однозначно определяется вектором $\bar{s}(x, \sigma_{m(x)})$). Выбор $\text{сог } x \in \text{Сог } x$ связан с такими субъективными факторами, как выбор алгоритма решения задачи линейного программирования, выбор начального плана при решении задачи и т. п., приводящими к построению определенного решения $(\bar{s}(x, \sigma_{m(x)}), \bar{z}(x, \sigma_{m(x)}))$ задачи $(S(x, \sigma_{m(x)}), g)$. Поэтому при доказательстве сходимости алгоритма следует проверить выполнение условий соответствующего утверждения п. 1.4.

4.5. Вычислительные аспекты. Построение итеративной последовательности по методу допустимых направлений начинается после вычисления допустимого плана x_1 задачи $(\mathcal{D}, f)$. Задачу нахождения вектора x_1 обсудим в п. 4.7.

Зададимся числом $\sigma > 0$. Выбор числа σ связан с конкретными значениями параметров задачи. Если σ выбрать большим, то может значительно увеличиться число задач, которые нужно будет решать при выборе допустимого прогрессивного направления. При малом значении σ может замедлиться процесс сходимости итеративной последовательности. Эти обстоятельства следует иметь в виду.

Предположим, что уже построены векторы $x_1, x_2, \dots, x_k$. Вычисляем значения $f_i(x_k)$, $i \in I_n$, и формируем множества $I(x_k)$, $J(x_k)$. Затем вычисляем значения $\nabla f_i(x_k)$, $i \in I_n \cap I(x_k)$, $\nabla f(x_k)$

и решаем задачу линейного программирования $(S(x_k), g)$. Если $\bar{z}(x_k) = 0$, то x_k — решение задачи $(\mathcal{D}, f)$. В противном случае если $\bar{z}(x_k) > 0$, последовательно решаем задачи $(S(x_k, \sigma_m), g)$ при $m = 0, 1, 2, \dots$ до тех пор, пока не будут найдены такой номер $m(x_k)$ и такое число $\sigma_{m(x_k)} = (1/2^{m(x_k)})\sigma$, что $\bar{z}(x_k, \sigma_{m(x_k)}) \geq \sigma_{m(x_k)}$. Эта процедура закончится через конечное число шагов в силу утверждения 5° п. 4.3. Соответствующий вектор $\bar{s}(x_k, \sigma_{m(x_k)})$ — допустимое прогрессивное направление в точке x_k . При составлении каждой задачи $(S(x_k, \sigma_m), g)$ предварительно определяются σ_m -активные множества индексов $I(x_k, \sigma_m)$, $J(x_k, \sigma_m)$ и вычисляются векторы $\nabla f_i(x_k)$, $i \in I(x_k, \sigma_m) \cap I_B$. Здесь обращаем внимание на два обстоятельства. Во-первых, множество ограничений задачи $(S(x_k, \sigma_{m+1}), g)$ содержится в множестве ограничений задачи $(S(x_k, \sigma_m), g)$ (см. утверждение 2° п. 4.3), и, следовательно, оптимальный план задачи $(S(x_k, \sigma_m), g)$ является допустимым для задачи $(S(x_k, \sigma_{m+1}), g)$, а система ограничений последней задачи получается из системы ограничений предыдущей исключением «лишних» ограничений. Во-вторых, решение серии задач $(S(x_k, \sigma_m), g)$ имеет смысл начинать только с такого номера m' , при котором $\sigma_{m'} = (1/2^{m'})\sigma \leq \bar{z}(x_k)$, так как $\bar{z}(x_k, \sigma_m) \leq \bar{z}(x_k)$ при $m = 0, 1, 2, \dots$

После построения вектора $\bar{s}(x_k, \sigma_{m(x_k)})$ решается задача выбора шага. Для этого сначала определяются положительные параметры α_i , $i \in I_B \cup (I_B \setminus (I_B \cup I(x_k)))$, и β_j , $j \in J \setminus J(x_k)$, точек пересечения луча $\{x_k + \mu \bar{s}(x_k, \sigma_{m(x_k)}), \mu \geq 0\}$ с граничными гиперповерхностями $f_i(x) = 0$, $i \in I_B$, $f_i(x) = a_i x - b[i] = 0$, $i \in I_B \setminus (I_B \cup I(x_k))$, $x[j] = 0$, $j \in J \setminus J(x_k)$. Отметим, что среди аффинных ограничений рассматриваются только те гиперплоскости, которые не содержат точку x_k , а выпуклые ограничения рассматриваются все, так как луч может пересекать такую гиперповерхность в двух точках.

Число α_i (β_j) определяется из уравнения

$$f_i(x_k + \alpha \bar{s}(x_k, \sigma_{m(x_k)})) = 0 \quad (x_k[j] + \beta \bar{s}(x_k, \sigma_{m(x_k)})[j] = 0),$$

в котором неизвестным является значение числового параметра α (β). Если это уравнение не имеет положительного корня, то соответствующее значение α_i (β_j) полагается равным $+\infty$. После вычислений α_i и β_j полагаем

$$\mu' = \min \{\alpha_i, i \in I_B \cup (I_B \setminus (I_B \cup I(x_k))), \beta_j, j \in J \setminus J(x_k)\}.$$

Не исключается случай, когда все α_i и β_j равны $+\infty$: тогда и $\mu' = +\infty$.

После определения μ' решаем задачу вычисления числа λ' — наименьшего из положительных чисел λ , для которых

$$f(x_k + \lambda \bar{s}(x_k, \sigma_{m(x_k)})) = \max_{\mu > 0} f(x_k + \mu \bar{s}(x_k, \sigma_{m(x_k)})).$$

Эта задача имеет решение, так как $\bar{z}(x_k) > 0$ и множество $L_f(x_k)$ компактно (см. лемму 1 п. 2.3). Число λ' может быть найдено методами, рассмотренными в п. 2.6. Иногда, как, например, в задаче квадратичного программирования, его можно найти из уравнения

$$\nabla f(x_k + \lambda \bar{s}(x_k, \sigma_m(x_k))) \bar{s}(x_k, \sigma_m(x_k)) = 0.$$

Положим $\lambda(x_k) = \min\{\mu', \lambda'\}$ и $x_{k+1} = x_k + \lambda(x_k) \bar{s}(x_k, \sigma_m(x_k))$.

4.6. Обоснование алгоритма метода допустимых направлений.

Теорема. Алгоритм метода допустимых направлений приводит к построению итеративной последовательности $x_1, x_2, \dots$ точек множества $\mathcal{D}$. Если эта последовательность конечна, то ее последний элемент — решение задачи $(\mathcal{D}, f)$. Если последовательность $x_1, x_2, \dots$ бесконечна, то она содержит подпоследовательность, сходящуюся к решению задачи $(\mathcal{D}, f)$, и всякая сходящаяся подпоследовательность последовательности $\{x_k\}_{k=1}^\infty$ сходится к решению задачи $(\mathcal{D}, f)$. Если к тому же задача $(\mathcal{D}, f)$ имеет единственное решение (например, в случае строгой вогнутости функции f), то последовательность $\{x_k\}_{k=1}^\infty$ сходится к этому решению. Последовательность $\{f(x_k)\}_{k=1}^\infty$ всегда сходится к оптимальному значению целевой функции задачи $(\mathcal{D}, f)$.

Доказательство. При описании алгоритма метода (см. пп. 4.2—4.5) было установлено, что он формирует итеративную последовательность в соответствии со схемой алгоритма п. 1.4. Было также установлено, что вычисления заканчиваются после конечного числа итераций в том и только том случае, если на некоторой итерации будет получена приемлемая точка, которая в силу теоремы п. 4.2 является решением задачи $(\mathcal{D}, f)$. В силу компактности множества $\hat{L}_f(x_1)$ последовательность $\{x_k\}_{k=1}^\infty$ содержит сходящуюся подпоследовательность. Поэтому достаточно установить выполнение условий утверждения п. 1.4. Функция f , выбранная в качестве оценочной функции, непрерывно дифференцируема на R^n и, следовательно, непрерывна. Множество $\mathcal{D}$ замкнуто в силу непрерывности функций $f_i, i \in 1:m$. Остается показать, что если точка $\bar{x} \in \mathcal{D}$ не является приемлемой, т. е. если $\bar{z}(\bar{x}) > 0$, то найдутся такие числа $\varepsilon > 0$ и $\delta > 0$, что $f(x') - f(x) \geq \varepsilon$ для любого $x \in \bar{\mathcal{D}}(\bar{x}, \varepsilon) \cap \mathcal{D}$ и любого $x' \in \text{Cог } x$.

Доказательство этого утверждения достаточно громоздко. Чтобы облегчить изложение, докажем последовательно несколько более простых вспомогательных утверждений.

А. Пусть $\bar{x} \in \mathcal{D}$, $\sigma' > 0$, $(s, z) \in S(\bar{x}, \sigma')$ и $z > 0$. Тогда существует такое число $\delta > 0$, что $(s, \frac{1}{2}z) \in S(x, \sigma')$ при любом $x \in \bar{\mathcal{D}}(\bar{x}, \delta) \cap \mathcal{D}$.

Доказательство. В силу утверждения 4° п. 4.3 существует такое $\delta' > 0$, что $I(x, \sigma') \subset I(\bar{x}, \sigma')$, $I(x, \sigma') \subset I(\bar{x}, \sigma')$ для любого $x \in \mathcal{D}(\bar{x}, \delta') \cap \mathcal{D}$. Выберем такое число δ'' , чтобы для любого $x \in \mathcal{D}(\bar{x}, \delta'')$ выполнялись неравенства

$$\|\nabla f(x) - \nabla f(\bar{x})\| \leq z/(2\sqrt{n}), \quad (7)$$

$$\|\nabla f_i(x) - \nabla f_i(\bar{x})\| \leq z/(2\sqrt{n}), \quad i \in I(\bar{x}, \sigma') \cap I_{\bar{x}} \quad (8)$$

(это возможно в силу непрерывной дифференцируемости функции f и функций f_i , $i \in 1:m$). Положим $\delta = \min\{\delta', \delta''\}$ и покажем, что $(s, \frac{1}{2}z) \in S(x, \sigma')$, если $x \in \mathcal{D}(\bar{x}, \delta) \cap \mathcal{D}$.

Так как $(s, z) \in S(\bar{x}, \sigma')$ и $I(x, \sigma') \subset I(\bar{x}, \sigma')$, $I(x, \sigma') \subset I(\bar{x}, \sigma')$, то для вектора s выполняются неравенства $a_i s \leq 0$, $i \in I(x, \sigma') \setminus I_{\bar{x}}$, $s[j] \geq 0$, $j \in I(x, \sigma')$, $-1 \leq s[j] \leq 1$, $j \in 1:n$, и равенства $a_i s = 0$, $i \in I_{\bar{x}}$. Из неравенств $-1 \leq s[j] \leq 1$, $j \in 1:n$, следует, что $\|s\| \leq 1$. Поэтому и в силу (7), (8)

$$|\nabla f(x)s - \nabla f(\bar{x})s| \leq \|\nabla f(\bar{x}) - \nabla f(x)\| \cdot \|s\| \leq z/2,$$

$$|\nabla f_i(x)s - \nabla f_i(\bar{x})s| \leq \|\nabla f_i(x) - \nabla f_i(\bar{x})\| \cdot \|s\| \leq z/2, \quad i \in I(\bar{x}, \sigma') \cap I_{\bar{x}}.$$

Но тогда

$$\nabla f(x)s \geq \nabla f(\bar{x})s - |\nabla f(x)s - \nabla f(\bar{x})s| \geq \nabla f(\bar{x})s - z/2,$$

$$-\nabla f_i(x)s \geq -\nabla f_i(\bar{x})s - |\nabla f_i(x)s - \nabla f_i(\bar{x})s| \geq -\nabla f_i(\bar{x})s - z/2,$$

$$i \in I(\bar{x}, \sigma') \cap I_{\bar{x}}.$$

Так как $(s, z) \in S(\bar{x}, \sigma')$, то $z \leq \nabla f(\bar{x})s$ и $z \leq -\nabla f_i(\bar{x})s$, $i \in I(\bar{x}, \sigma') \cap I_{\bar{x}}$. Следовательно, $\nabla f(x)s \geq z/2$ и $-\nabla f_i(x)s \geq z/2$, если $i \in I(x, \sigma') \cap I_{\bar{x}}$, поскольку $I(x, \sigma') \subset I(\bar{x}, \sigma')$. ■

Б. Пусть $\bar{x} \in \mathcal{D} \setminus \pi$. Тогда существует такое $\delta_1 > 0$, что $\sigma_{m(x)} \geq \frac{1}{2}\sigma_{m(\bar{x})}$ для любого $x \in \mathcal{D}(\bar{x}, \delta_1) \cap \mathcal{D}$.

Доказательство. Пусть $(\bar{s}(\bar{x}, \sigma_{m(\bar{x})}), \bar{z}(\bar{x}, \sigma_{m(\bar{x})}))$ — какое-нибудь решение задачи $(S(\bar{x}, \sigma_{m(\bar{x})}), g)$. Тогда $\bar{z}(\bar{x}, \sigma_{m(\bar{x})}) \geq \sigma_{m(\bar{x})}$. В силу утверждения А существует такое $\delta_1 > 0$, что для любого $x \in \mathcal{D}(\bar{x}, \delta_1) \cap \mathcal{D}$ вектор $(\bar{s}(\bar{x}, \sigma_{m(\bar{x})}), \frac{1}{2}\bar{z}(\bar{x}, \sigma_{m(\bar{x})})) \in S(x, \sigma_{m(\bar{x})})$. Так как $S(x, \sigma_{m(\bar{x})}) \subset S(x, \frac{1}{2}\sigma_{m(\bar{x})})$, то $(\bar{s}(\bar{x}, \sigma_{m(\bar{x})}), \frac{1}{2}\bar{z}(\bar{x}, \sigma_{m(\bar{x})})) \in S(x, \frac{1}{2}\sigma_{m(\bar{x})})$ для любого $x \in \mathcal{D}(\bar{x}, \delta_1) \cap \mathcal{D}$. Поэтому если $x \in \mathcal{D}(\bar{x}, \delta_1) \cap \mathcal{D}$, то наибольшее значение функции g на $S(x, \frac{1}{2}\sigma_{m(\bar{x})})$, которое обозначим через $\bar{z}(x, \frac{1}{2}\sigma_{m(\bar{x})})$, не меньше $\frac{1}{2}\bar{z}(\bar{x}, \sigma_{m(\bar{x})})$. Следовательно, $\bar{z}(x, \frac{1}{2}\sigma_{m(\bar{x})}) \geq \frac{1}{2}\sigma_{m(\bar{x})}$. Так как $\sigma_{m(x)}$ — наибольшее из σ_k , $k = 0, 1, 2, \dots$, для которых $\bar{z}(x, \sigma_k) \geq \sigma_k$, то $\sigma_{m(x)} \geq \frac{1}{2}\sigma_{m(\bar{x})}$ для любого $x \in \mathcal{D}(\bar{x}, \delta_1) \cap \mathcal{D}$. ■

В. Пусть $\bar{x} \in \mathcal{D} \setminus \pi$ и δ_1 — такое положительное число, что $\sigma_{m(x)} \geq \frac{1}{2}\sigma_{m(\bar{x})}$ для любого $x \in \mathcal{D}(\bar{x}, \delta_1) \cap \mathcal{D}$ (см. утверждение Б). Тогда существует такое число $\bar{\mu} > 0$, что для любого $x \in \mathcal{D}(\bar{x}, \frac{1}{2}\delta_1)$ и любого решения $(\bar{s}(x, \sigma_{m(x)}), \bar{z}(x, \sigma_{m(x)}))$ задачи $(S(x, \sigma_{m(x)}), g)$ вектор $x + \bar{\mu}\bar{s}(x, \sigma_{m(x)}) \in \mathcal{D}$ и $f(x + \bar{\mu}\bar{s}(x, \sigma_{m(x)})) - f(x) \geq \frac{1}{4}\bar{\mu}\sigma_{m(x)}$.

Доказательство. Функция f и функции $f_i, i \in 1:m$, равномерно непрерывны на компактном множестве $\mathcal{D}(\bar{x}, \delta_1)$, и, следовательно, найдется такое $\delta_2 > 0$, что

$$|f(x_1) - f(x_2)| < 1/2 \sigma_{m(\bar{x})}, |f_i(x_1) - f_i(x_2)| < 1/2 \sigma_{m(\bar{x})}, i \in 1:m,$$

если $x_1, x_2 \in \mathcal{D}(\bar{x}, \delta_1)$ и $\|x_1 - x_2\| \leq \delta_2$. В силу непрерывной дифференцируемости функции f и функций $f_i, i \in 1:m$, отображения ∇f и $\nabla f_i, i \in I_B$ (из R^n в R^n), равномерно непрерывны на $\mathcal{D}(\bar{x}, \delta_1)$, так что найдется такое $\delta_3 > 0$, что

$$\|\nabla f(x_1) - \nabla f(x_2)\| \leq \frac{1}{4\sqrt{n}} \sigma_{m(\bar{x})},$$

$$\|\nabla f_i(x_1) - \nabla f_i(x_2)\| \leq \frac{1}{2\sqrt{n}} \sigma_{m(\bar{x})}, i \in I_B,$$

если x_1, x_2 из $\mathcal{D}(\bar{x}, \delta_1)$ и $\|x_1 - x_2\| \leq \delta_2$. Положим

$$\bar{\mu} = \min \left\{ \frac{1}{2\sqrt{n}} \delta_1, \frac{1}{\sqrt{n}} \delta_2, \frac{1}{\sqrt{n}} \delta_3, \frac{1}{2} \sigma_{m(\bar{x})} \right\}.$$

Пусть $x \in \mathcal{D}(\bar{x}, 1/2 \delta_1) \cap \mathcal{D}$ и $(\bar{s}(x, \sigma_{m(x)}), \bar{z}(x, \sigma_{m(x)}))$ — какое-нибудь решение задачи $(S(x, \sigma_{m(x)}), g)$. Проверим, что $x + \bar{\mu} \bar{s}(x, \sigma_{m(x)}) \in \mathcal{D}$ и

$$f(x + \bar{\mu} \bar{s}(x, \sigma_{m(x)})) \geq f(x) + (\bar{\mu}/4) \sigma_{m(\bar{x})}. \quad (9)$$

Для краткости будем писать $\bar{s}$ вместо $\bar{s}(x, \sigma_{m(x)})$ и $\bar{z}$ вместо $\bar{z}(x, \sigma_{m(x)})$.

Прежде всего отметим, что $\|\bar{x} - (x + \bar{\mu} \bar{s})\| \leq \|\bar{x} - x\| + \bar{\mu} \|\bar{s}\|$, и так как $x \in \mathcal{D}(\bar{x}, 1/2 \delta_1)$, $\|\bar{s}\| \leq \sqrt{n}$ и $\bar{\mu} \leq \frac{1}{2\sqrt{n}} \delta_1$, то $x + \bar{\mu} \bar{s} \in \mathcal{D}(\bar{x}, \delta_1)$. Кроме того, $\|x - (x + \bar{\mu} \bar{s})\| = \bar{\mu} \|\bar{s}\| \leq \min \{\delta_2, \delta_3\}$, так что

$$|f_i(x) - f_i(x + \bar{\mu} \bar{s})| < 1/2 \sigma_{m(\bar{x})}, i \in 1:m, \quad (10)$$

$$|\nabla f(x) \bar{s} - \nabla f(x + \bar{\mu} \bar{s}) \bar{s}| \leq \|\nabla f(x) - \nabla f(x + \bar{\mu} \bar{s})\| \cdot \|\bar{s}\| < 1/4 \sigma_{m(\bar{x})}, \quad (11)$$

$$|\nabla f_i(x) \bar{s} - \nabla f_i(x + \bar{\mu} \bar{s}) \bar{s}| < 1/2 \sigma_{m(\bar{x})}, i \in I_B. \quad (12)$$

Пусть $i \in I_B \cap I(x, \sigma_{m(x)})$. Так как функция f_i выпукла, то (см. п. 1.8.8)

$$f_i(x) - f_i(x + \bar{\mu} \bar{s}) \geq -\bar{\mu} \nabla f_i(x + \bar{\mu} \bar{s}) \bar{s},$$

откуда в силу (12) получаем

$$f_i(x + \bar{\mu} \bar{s}) \leq f_i(x) + \bar{\mu} \nabla f_i(x + \bar{\mu} \bar{s}) \bar{s} \leq f_i(x) + \bar{\mu} (\nabla f_i(x) \bar{s} + 1/2 \sigma_{m(\bar{x})}).$$

Так как $x \in \mathcal{D}$, то $f_i(x) \leq 0$. Кроме того, $1/2 \sigma_{m(\bar{x})} \leq \sigma_{m(x)} \leq \bar{z}$, так как $x \in \mathcal{D}(\bar{x}, \delta_1)$, и $\nabla f_i(x) \bar{s} + \bar{z} \leq 0$, поскольку $(\bar{s}, \bar{z}) \in S(x, \sigma_{m(x)})$. Следовательно, $\nabla f_i(x + \bar{\mu} \bar{s}) \bar{s} \leq 0$.

Пусть $i \in (I_H \setminus I_V) \cap I(x, \sigma_m(x))$. Тогда $f_i(x) < -\sigma_m(x)$ и в силу (10)

$$f_i(x + \bar{\mu}\bar{s}) < f_i(x) + 1/2 \sigma_m(x) < -\sigma_m(x) + 1/2 \sigma_m(x) \leq 0.$$

Если $i \in I_y$, то $a_i(x + \bar{\mu}\bar{s}) = a_i x + \bar{\mu} a_i \bar{s} = b[i]$, так как $a_i x = b[i]$ и $a_i \bar{s} = 0$, $i \in I_y$. Если $j \in J(x, \sigma_m(x))$, то $(x + \bar{\mu}\bar{s})[j] = x[j] + \bar{\mu}\bar{s}[j] \geq 0$, так как $x[j] \geq 0$ и $\bar{s}[j] \geq 0$ для $j \in J(x, \sigma_m(x))$.

Если $j \in J \setminus J(x, \sigma_m(x))$, то $x[j] \geq \sigma_m(x)$, и так как $\bar{s}[j] \geq -1$, а $\bar{\mu} \leq 1/2 \sigma_m(x)$, то $\bar{\mu}\bar{s}[j] \geq -1/2 \sigma_m(x)$ и

$$(x + \bar{\mu}\bar{s})[j] = x[j] + \bar{\mu}\bar{s}[j] \geq \sigma_m(x) - 1/2 \sigma_m(x) \geq 0.$$

Докажем теперь неравенство (9). Так как f — вогнутая функция, то (см. замечание 2 п. 1.8.8)

$$f(x) - f(x + \bar{\mu}\bar{s}) \leq -\bar{\mu} \nabla f(x + \bar{\mu}\bar{s}) \bar{s},$$

откуда с учетом неравенства (11) получаем

$$f(x + \bar{\mu}\bar{s}) \geq f(x) + \bar{\mu} (\nabla f(x) \bar{s} - 1/4 \sigma_m(x)). \quad (13)$$

Так как $\nabla f(x) \bar{s} \geq \bar{z} \geq 1/2 \sigma_m(x)$, то из (13) следует (9).

Г. Покажем, наконец, что для $\bar{x} \in \mathcal{D} \setminus \pi$ существуют такие $\varepsilon > 0$ и $\delta > 0$, что $f(x') - f(x) > \varepsilon$ для любого $x \in \mathcal{D}(\bar{x}, \delta) \cap \mathcal{D}$. Возьмем такое $\delta_1 > 0$, что $\sigma_m(x) \geq 1/2 \sigma_m(\bar{x})$ для любого $x \in \mathcal{D}(\bar{x}, \delta_1) \cap \mathcal{D}$, и такое $\bar{\mu} > 0$, что для любого $x \in \mathcal{D}(\bar{x}, 1/2 \delta_1)$ и любого решения $(\bar{s}(x, \sigma_m(x)), \bar{z}(x, \sigma_m(x)))$ задачи $(S(x, \sigma_m(x)), g)$ вектор $x + \bar{\mu}\bar{s}(x, \sigma_m(x)) \in \mathcal{D}$ и

$$f(x + \bar{\mu}\bar{s}(x, \sigma_m(x))) \geq f(x) + 1/4 \bar{\mu} \sigma_m(x)$$

(см. утверждение В). Положим $\delta = 1/2 \delta_1$, $\varepsilon = 1/4 \bar{\mu} \sigma_m(\bar{x})$. Тогда если $x \in \mathcal{D}(\bar{x}, \delta)$, $x' \in \text{Cor } x$, то $x' = x + \lambda(x) \bar{s}(x, \sigma_m(x))$, где $(\bar{s}(x, \sigma_m(x)), \bar{z}(x, \sigma_m(x)))$ — некоторое решение задачи $(S(x, \sigma_m(x)), g)$, и так как $f(x') = f(x + \lambda(x) \bar{s}(x, \sigma_m(x))) = \max \{f(x + \mu \bar{s}(x, \sigma_m(x))), \mu > 0, x + \mu \bar{s}(x, \sigma_m(x)) \in \mathcal{D}\}$, то

$$f(x') \geq f(x + \bar{\mu}\bar{s}(x, \sigma_m(x))) \geq f(x) + \varepsilon. \blacksquare$$

4.7. Вычисление допустимого плана задачи выпуклого программирования. В рассмотренных выше итеративных алгоритмах решения задачи $(\mathcal{D}, f)$ выпуклого программирования расчеты начинались с некоторого плана $x_1 \in \mathcal{D}$. Задача нахождения этого вектора является, вообще говоря, самостоятельной вычислительной задачей (иногда его удастся определить из соображений, связанных с реальной ситуацией, на основе анализа которой возникла экстремальная задача).

Рассмотрим следующую экстремальную задачу, связанную с задачей $(\mathcal{D}, f)$:

$$\begin{aligned} (\mathcal{D}', f') : f'(x, z) = z \rightarrow \max, \quad \mathcal{D}' = \{(x, z) \in R^{n+1} : \\ f_i(x) + z \leq 0, \quad i \in I_n, \quad a_i x - b[i] \leq 0, \quad i \in I_n \setminus I_n, \\ a_i x - b[i] = 0, \quad i \in I_y, \quad x[j] \geq 0, \quad j \in J\}. \end{aligned}$$

Нетрудно убедиться, что $(\mathcal{D}', f')$ — задача выпуклого программирования и множество $\mathcal{D}' \neq \emptyset$ в том и только том случае, если система уравнений и неравенств

$$\begin{aligned} a_i x &= b[i], \quad i \in I_y, \\ a_i x &\leq b[i], \quad i \in I_n \setminus I_n, \\ x[j] &\geq 0, \quad j \in J, \end{aligned} \quad (14)$$

совместна. Более того, для любого вектора $x' \in R^n$, удовлетворяющего соотношениям (14), и $z' = \min_{i \in I_n} \{-f_i(x')\}$ вектор

(x', z') очевидно принадлежит $\mathcal{D}'$. Очевидно также, что если $(x', z') \in \mathcal{D}'$ и $z' \geq 0$, то $x' \in \mathcal{D}$. Эти соображения позволяют наметить следующую схему вычисления начального допустимого плана задачи $(\mathcal{D}, f)$.

Методом минимизации невязок (см. пп. 4.2.2, 1° или 4.2.3, 1°) вычисляется вектор x'_1 , удовлетворяющий системе (14) (если система несовместна, то $\mathcal{D} = \emptyset$). Полагаем $z'_1 = \min_{i \in I_n} \{-f_i(x'_1)\}$.

Вектор $(x'_1, z'_1) \in \mathcal{D}'$, и для решения задачи выпуклого программирования $(\mathcal{D}', f')$ можно применить какой-либо из итеративных алгоритмов, например, алгоритм настоящего параграфа.

Пусть $(\bar{x}', \bar{z}')$ — решение задачи $(\mathcal{D}', f')$. Тогда если $\bar{z}' \geq 0$, то $\bar{x}' \in \mathcal{D}$. Очевидно, что если для системы ограничений I_n выполняются условия регулярности С, то $\bar{z}' > 0$. Поэтому (см. п. 1.3) в случае выполнения условия С для множества ограничений I_n в процессе решения задачи $(\mathcal{D}', f')$ за конечное число итераций будет построен такой вектор (x'_k, z'_k) , что $z'_k > 0$. Но тогда $x'_k \in \mathcal{D}$.

§ 5. Итеративный метод решения задачи квадратичного программирования — метод сопряженных направлений

В настоящем параграфе рассматривается итеративный метод решения задачи квадратичного программирования

$$\begin{aligned} (\mathcal{D}, f) : f(x) = cx + x^T D x \rightarrow \max, \quad \mathcal{D} = \{x \in R^n : \\ a_i x \leq b[i], \quad i \in I_n; \quad a_i x = b[i], \quad i \in I_y; \quad x[j] \geq 0, \quad j \in J\} \end{aligned} \quad (1)$$

(см. п. 2.2.3 и § 2 гл. 5) с неположительно определенной матрицей D . Алгоритм метода укладывается в общую схему методов

допустимых направлений, обсуждавшуюся в § 4, и является, по-видимому, одним из наиболее эффективных методов решения задач квадратичного программирования. Конечная сходимость метода обеспечивается специальными ограничениями на выбор допустимого прогрессивного направления — условиями сопряженности.

5.1. Предварительные замечания. Построение итеративной последовательности $x_0, x_1, \dots$ начинается после вычисления вектора $x_0 \in \mathcal{D}$. Если $\mathcal{D} \neq \emptyset$, то $\mathcal{D}$ — многогранное выпуклое множество и такой вектор может быть найден методами линейного программирования (например, методом минимизации невязок, см. пп. 4.2.2, 1° или 4.3.8, 4.4.8).

В соответствии с рекомендациями, содержащимися в § 4, для проверки оптимальности (приемлемости) вектора $x \in \mathcal{D}$ составляется задача линейного программирования $(S(x), g)$, формулировка которой приведена в п. 4.2. Легко видеть, что для задачи квадратичного программирования (1) задача $[(S(x), g)]$ записывается следующим образом:

$$(S(x), g): g(s, z) = z \rightarrow \max, S(x) = \{(s, z) \in R^{n+1}, s \in R^n, z \in R:$$

$$a_i s \leq 0, i \in I(x), \quad (2)$$

$$a_i s = 0, i \in I_y, \quad (3)$$

$$s[j] \geq 0, j \in I(x), \quad (4)$$

$$-1 \leq s[j] \leq 1, j \in 1:n, \quad (5)$$

$$z \leq \nabla f(x) s, \quad (6)$$

причем ограничение (6) можно записать в виде

$$z \leq (c + 2Dx) s, \quad (6')$$

поскольку

$$\nabla f(x) = c + 2Dx. \quad (7)$$

Задачу максимизации функции g на множестве, заданном системой ограничений (2) — (6), а также дополнительными ограничениями, о которых будет сказано ниже, назовем *задачей выбора направления*.

В п. 4.2 было установлено, что задача $(S(x), g)$ всегда имеет решение $(\bar{s}(x), \bar{z}(x))$, и равенство $\bar{z}(x) = \nabla f(x) \bar{s}(x) = 0$ необходимо и достаточно для оптимальности вектора $x \in \mathcal{D}$.

Если $\bar{z}(x) = \nabla f(x) \bar{s}(x) > 0$, то вектор $\bar{s}(x)$ — допустимое прогрессивное направление в точке x (см. доказательство теоремы п. 4.2) и для построения $\text{сог } x = x + \lambda(x) \bar{s}(x)$ нужно решить задачу одномерной оптимизации: $\lambda(x)$ — наименьшее из таких чисел $\lambda > 0$, что $f(x + \lambda \bar{s}(x)) = \max\{f(x + \mu \bar{s}(x)), 0 < \mu \leq \mu'\}$, где μ' — наибольшее из положительных чисел μ , при которых $x + \mu \bar{s}(x) \in \mathcal{D}$. Покажем, что в задаче квадратичного программирования $\lambda(x)$ может быть получено применением следующей простой вычислительной процедуры.

1) **Вычисление** $\mu' = \mu'(x)$. Очевидно, что $x + \mu\bar{s}(x) \in \mathcal{D}$ в том и только том случае, если

$$a_i x + \mu a_i \bar{s}(x) \leq b[i], \quad i \in I_n, \quad a_i x + \mu a_i \bar{s}(x) = b[i], \quad i \in I_y,$$

и $x[j] + \mu\bar{s}(x)[j] \geq 0, j \in J$. Отметим, что $b[i] - a_i x > 0$ для $i \in I_n \setminus I(x)$ и $x[j] > 0$ для $j \in J \setminus J(x)$. Так как вектор $\bar{s}(x)$ удовлетворяет условиям (2) — (4), то $\mu' = \mu'(x)$ — наибольшее из таких чисел μ , для которых $\mu a_i \bar{s}(x) \leq b[i] - a_i x, i \in I_n \setminus I(x)$, и $\mu\bar{s}(x)[j] \geq -x[j], j \in J \setminus J(x)$. Если $a_i \bar{s}(x) \leq 0, i \in I_n \setminus I(x)$, и $\bar{s}(x)[j] \geq 0, j \in J \setminus J(x)$, то полагаем $\mu' = \mu'(x) = +\infty$ (естественно, $\mu' = +\infty$, если $I_n \setminus I(x) = J \setminus J(x) = \emptyset$).

2) **Вычисление** $\lambda' = \lambda'(x)$. Для задачи квадратичного программирования имеет место следующее утверждение.

Пусть $\nabla f(x) \bar{s}(x) > 0$. Если $\bar{s}(x)^T D \bar{s}(x) = 0$, то функция $\phi(\mu) = f(x + \mu\bar{s}(x))$ монотонно возрастает и не ограничена сверху на промежутке $[0, +\infty)$; если $\bar{s}(x)^T D \bar{s}(x) < 0$, то существует единственное положительное число λ' , для которого $\phi(\lambda') = \max\{\phi(\mu), \mu \geq 0\}$, и функция ϕ монотонно возрастает на промежутке $[0, \lambda']$.

В самом деле (см. (1) и (7)),

$$\begin{aligned} \phi(\mu) &= c(x + \mu\bar{s}(x)) + (x + \mu\bar{s}(x))^T D(x + \mu\bar{s}(x)) = \\ &= cx + \mu c\bar{s}(x) + x^T D x + 2\mu x^T D \bar{s}(x) + \mu^2 \bar{s}(x)^T D \bar{s}(x) = \\ &= f(x) + \mu \nabla f(x) \bar{s}(x) + \mu^2 \bar{s}(x)^T D \bar{s}(x), \end{aligned}$$

так что утверждение очевидно следует из известных свойств квадратичной функции одной переменной. Если $\bar{s}(x)^T D \bar{s}(x) < 0$, то функция ϕ принимает наибольшее значение при $\mu = -\nabla f(x) \bar{s}(x) / \bar{s}(x)^T D \bar{s}(x)$. ■

Полагаем $\lambda' = \lambda'(x) = +\infty$, если $\bar{s}(x)^T D \bar{s}(x) = 0$, и

$$\lambda' = \lambda'(x) = -\nabla f(x) \bar{s}(x) / \bar{s}(x)^T D \bar{s}(x), \quad (8)$$

если $\bar{s}(x)^T D \bar{s}(x) < 0$.

3) **Вычисление** $\lambda(x)$. Полагаем $\lambda(x) = \min\{\mu'(x), \lambda'(x)\}$.

Как уже отмечалось, конечная сходимость алгоритма настоящего параграфа обеспечивается специальными ограничениями на выбор допустимого прогрессивного направления. Содержание этих ограничений состоит в следующем. Предположим, что уже построены векторы $x_0, x_1, \dots, x_m$ итеративной последовательности и $x_k = x_{k-1} + \lambda_{k-1} \bar{s}_{k-1}, k \in 1:m$ (для удобства будем писать λ_k и $\bar{s}_k$ вместо $\lambda(x_k)$ и $\bar{s}(x_k)$, а также λ'_k, μ'_k вместо $\lambda'(x_k), \mu'(x_k)$). Возможны два случая.

а) $\lambda_{m-1} = \mu'_{m-1}$. Тогда точка x_m принадлежит некоторой ограничивающей гиперплоскости многогранного выпуклого множества $\mathcal{D}$, не содержащей x_{m-1} . Если этой гиперплоскости принадлежала хотя бы одна из точек $x_k, k \in 0:m-2$, то в задаче

выбора направления соответствующее неравенство в (2) или (4) заменяется на уравнение. Тогда x_{m+1} — следующая за x_m точка итеративной последовательности — будет принадлежать той же гиперплоскости. Такая замена неравенства уравнением в задачах выбора направления сохраняется и на последующих итерациях, а если к тому же обнаружится повторное попадание точек итеративной последовательности еще на какую-нибудь ограничивающую гиперплоскость, то в следующих задачах выбора направления еще одно (соответствующее) неравенство в (2) или (4) заменяется на уравнение. Все эти замены сохраняются до тех пор, пока не будет найдено максимальное значение целевой функции на некоторой грани множества $\mathcal{D}$. После того как такое значение будет найдено, т. е. на некоторой итерации будет обнаружено, что $\bar{z}(x_k) = 0$, в задаче выбора направления все измененные ограничения — уравнения — снова заменяются на неравенства. Здесь полезно напомнить, что число граней многогранного выпуклого множества конечно.

б) $\lambda_{m-1} = \lambda'_{m-1} < \mu'_{m-1}$. Пусть $\lambda_k = \lambda'_k < \mu'_k$ для всех $k \in \varepsilon: m-1$. В этом случае к задаче выбора направления добавляются дополнительные ограничения вида $s^T D \bar{s}_k = 0$, $k \in \varepsilon: m-1$, которые называются *условиями сопряженности* векторов s и $\bar{s}_k$ (векторы a и b называются *сопряженными относительно матрицы* D , если $a^T D b = 0$). Будет показано (см. 1° п. 5.4), что процесс построения сопряженных направлений конечен, поскольку попарно сопряженные векторы $s_1, s_2, \dots, s_k$, для которых $s_i^T D s_i \neq 0$, $i \in 1:k$, образуют линейно независимую систему. В то же время (это будет установлено в 4° п. 5.4) дополнительные условия сопряженности не мешают построению допустимого прогрессивного направления, если такое направление существует.

5.2. Вычислительная процедура. Как уже отмечалось (см. п. 5.1), построение итеративной последовательности начинается с вычисления вектора $x_0 \in \mathcal{D}$, которое можно осуществить методом минимизации невязок. Предположим, что уже построены векторы $x_0, x_1, \dots, x_m, x_k = x_{k-1} + \lambda_{k-1} \bar{s}_{k-1}$, $k \in 1:m$, и заданы множества индексов $K \subset I_n$, $L \subset J$ (эти множества могут быть пустыми). При вычислении шага λ_{m-1} (если $m \geq 1$) имел место один и только один из следующих случаев: а) $\lambda_{m-1} = \mu'_{m-1}$ б) $\lambda_{m-1} = \lambda'_{m-1} < \mu'_{m-1}$ (см. правила 1) — 3) вычисления шага в п. 5.1).

1. В случае а) определяем (см. п. 5.1, 1)) индекс $i_0 \in \varepsilon \cap I(x_m)$, для которого $\lambda_{m-1} a_{i_0} \bar{s}_{m-1} = b[i_0] - a_{i_0} x_m$ (или индекс $j_0 \in J \setminus I(x_m)$, для которого $\lambda_{m-1} \bar{s}_{m-1}[j_0] = -x_{m-1}[j_0]$). Проверяем, принадлежал ли этот индекс i_0 множеству $I(x_k)$ при каком-либо $k \in 0:m-2$ (соответственно принадлежал ли индекс j_0 множеству $J(x_k)$ при каком-либо $k \in 0:m-2$). Если принадлежал, то полагаем $K = K \cup \{i_0\}$ (соответственно

$L = L \cup \{j_0\}$). Составляется задача выбора направления

$$\begin{aligned} (S_m, g) : g(s, z) = z \rightarrow \max, \quad S_m = \{(s, z) \in R^{n+1}, s \in R^n, z \in R : \\ a_i s = 0, i \in K; a_i s \leq 0, i \in I(x_m) \setminus K; a_i s = 0, i \in I_y; \\ s[j] = 0, j \in L; s[j] \geq 0, j \in J(x_m) \setminus L; \\ -1 \leq s[j] \leq 1, j \in 1:n; z \leq (c + 2Dx_m)s\} \end{aligned} \quad (9)$$

(ср. с условиями (2)–(5), (6')). Переходим к рекомендации 3.

2. В случае б) определяем такое $r \in 0:m-2$, что $\lambda_k = \lambda$ при $k \in r:m-1$ и $\lambda_{r-1} = \mu'_{r-1}$. Составляется задача (S_m, g) выбора направления, в которой $g(s, z) = z$, а множество S_m определяется системой ограничений (9) и условиями сопряженности

$$\bar{s}_k Ds = 0, k \in r:m-1. \quad (10)$$

Переходим к рекомендации 3.

3. Решается задача (S_m, g) . Пусть $(\bar{s}_m, \bar{z}_m)$ — ее решение. Если $\bar{z}_m > 0$, переходим к рекомендации 4. Если $\bar{z}_m = 0$ и хотя бы одно из множеств K или L непусто, то полагаем $K = \emptyset$, $L = \emptyset$ и снова решаем задачу (S_m, g) (с измененной системой ограничений). Пусть $(\bar{s}_m, \bar{z}_m)$ — ее решение. Если $\bar{z}_m = 0$, то x_m — решение задачи $(\mathcal{D}, f)$. Если $\bar{z}_m > 0$, то переходим к рекомендации 4.

4. Так как $\bar{z}_m > 0$, то $\bar{s}_m$ — допустимое прогрессивное направление в точке x_m . С помощью процедуры 1)–3), описанной в п. 5.1, вычисляем $\lambda_m = \lambda(x_m)$. Если $\lambda_m = +\infty$, то задача $(\mathcal{D}, f)$ не имеет решения — целевая функция f не ограничена сверху на $\mathcal{D}$. В противном случае полагаем $x_{m+1} = x_m + \lambda_m \bar{s}_m$, $m = m + 1$. Замечаем, какой из случаев а) или б) имеет место, и в зависимости от этого переходим к рекомендации 1 или 2.

5.3. Пример. Рассмотрим следующую задачу квадратичного программирования:

$$\begin{aligned} (\mathcal{D}, f) : f(x) = 2x[1] + 2x[2] - x[1]^2 - 2x[2]^2 + 2x[1]x[2] \rightarrow \max, \\ \mathcal{D} = \{x \in R^2 : 2x[1] + x[2] \leq 4, -x[1] + x[2] \leq 2, \\ x[1] \geq 0, x[2] \geq 0\}. \end{aligned}$$

Здесь $a_1 = (2, 1)$, $b[1] = 4$, $a_2 = (-1, 1)$, $b[2] = 2$, $c = (2, 2)$,

$$D = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix}, \quad \nabla f(x) = (2 - 2x[1] + 2x[2], 2 + 2x[1] - 4x[2]),$$

$$I_u = \{1, 2\}, \quad I_y = \emptyset, \quad J = \{1, 2\}.$$

Очевидно, что $x_0 = (0, 0) \in \mathcal{D}$. Тогда $I(x_0) = \emptyset$, $J(x_0) = \{1, 2\}$; $\nabla f(x_0) = (2, 2)$. Полагаем $K = L = \emptyset$.

Первая итерация. $(S_0, g): g(s, z) = z \rightarrow \max, S_0 = \{(s, z) \in R^3:$

$$s[j] \geq 0, -1 \leq s[j] \leq 1, j \in 1:2; z \leq 2s[1] + s[2]\};$$

$$\bar{s}_0 = (1, 1), \bar{z}_0 = \nabla f(x_0) \bar{s}_0 = 4; \mu'_0 = 4/3, \lambda'_0 = 2, \lambda_0 = 4/3 = \mu'_0;$$

$$x_1 = (4/3, 4/3); \nabla f(x_1) = (2, -2/3); I(x_1) = \{1\}, J(x_1) = K = L = \emptyset.$$

Вторая итерация. $(S_1, g): g(s, z) = z \rightarrow \max, S_1 = \{(s, z) \in R^3:$

$$2s[1] + s[2] \leq 0; -1 \leq s[j] \leq 1, j \in 1:2; z \leq 2s[1] - 2/3 s[2]\};$$

$$\bar{s}_1 = (1/2, -1), \bar{z}_1 = 5/3; \mu'_1 = 4/3, \lambda'_1 = 10/39, \lambda_1 = 10/39 = \lambda'_1;$$

$$x_2 = (19/13, 14/13); \nabla f(x_2) = (16/13, 12/13); I(x_2) = \{1\}, J(x_2) = K = L = \emptyset.$$

Третья итерация. $(S_3, g): g(s, z) = z \rightarrow \max, S_3 = \{(s, z) \in R^3:$

$$2s[1] + s[2] \leq 0; -1 \leq s[j] \leq 1, j \in 1:2; z \leq 16/13 s[1] + 12/13 s[2];$$

$$1/2 s[1] - s[2] = 0\};$$

$$s_2 = (0, 0), \bar{z}_2 = 0.$$

Итак, $x_2 = (19/13, 14/13)$ — решение задачи $(\mathcal{D}, f)$, $f(x_2) = 3,77$.

5.4. Обоснование алгоритма. Докажем последовательно несколько утверждений.

1°. Система $d_1, d_2, \dots, d_k$ попарно сопряженных относительно матрицы D векторов, для которых $d_i^T D d_i \neq 0, i \in 1:k$, линейно независима.

Действительно, пусть $\sum_{i=1}^k \alpha_i d_i = 0$. Тогда при любом $l \in 1:k$

$$0 = \left(\sum_{i=1}^k \alpha_i d_i \right)^T D d_l = \sum_{i=1}^k \alpha_i d_i^T D d_l = \alpha_l d_l^T D d_l.$$

Следовательно, $\alpha_l = 0, l \in 1:k$. ■

2°. Пусть $x_1, x_2, \dots, x_m$ — элементы итеративной последовательности, $x_k = x_{k-1} + \lambda_{k-1} \bar{s}_{k-1}, k \in 1:m$. Тогда для любого $k \in 0:m-1$

$$\nabla f(x_m) = \nabla f(x_k) + 2 \sum_{i=k}^{m-1} \lambda_i D \bar{s}_i. \quad (11)$$

Действительно, $x_m = x_k + \sum_{i=k}^{m-1} \lambda_i \bar{s}_i$ для любого $k \in 0:m-1$ и (11) следует из (7). ■

3°. Пусть в процессе построения итеративной последовательности при некотором k окажется $\nabla f(x_k) \bar{s}_k > 0$ и $\bar{s}_k^T D \bar{s}_k = 0$. Тогда $\lambda_k = \mu'_k$ и f не ограничена сверху на $\mathcal{D}$, если $\mu'_k = +\infty$.

В самом деле, как установлено в рекомендации 2 вычисления шага (см. п. 5.1), $\lambda'_k = +\infty$, если $\bar{s}_k^T D \bar{s}_k = 0$, и функция f строго возрастает на луче $\{x_k + \mu \bar{s}_k, \mu \geq 0\}$. Поэтому $\lambda_k = \mu'_k$ и f не ограничена сверху на $\mathcal{D}$ при $\mu'_k = +\infty$. ■

4°. Пусть $x_0, x_1, \dots, x_m$ — члены итеративной последовательности, построенные в соответствии с рекомендациями п. 5.2, и пусть $\lambda_k = \lambda'_k$ для $k \in r:m-1$ (так что векторы $\bar{s}_r, \bar{s}_{r+1}, \dots$

..., $\bar{s}_{m-1}$ попарно сопряжены и $\bar{s}_k^T D \bar{s}_k \neq 0$, $k \in r:m-1$ (см. 3°)). Если множество S_m определяется системой ограничений (9) и условиями сопряженности (10) (см. рекомендацию 2 п. 5.2), а $\nabla f(x_m)s \leq 0$ для любого $s \in S_m$, то $\nabla f(x_m)s \leq 0$ и для любого s , удовлетворяющего только системе ограничений (9).

Доказательство. Предположим противное: пусть существует вектор t , удовлетворяющий системе ограничений (9), для которого $\nabla f(x_m)t > 0$. Положим $s = t - \sum_{k=r}^{m-1} \alpha_k \bar{s}_k$, где $\alpha_k = \bar{s}_k^T D t / \bar{s}_k^T D \bar{s}_k$, и покажем, что $\alpha s \in S_m$ при некотором $\alpha > 0$ и $\nabla f(x_m)s > 0$ (множитель α нужен, чтобы обеспечить выполнение неравенств $-1 \leq \alpha s[j] \leq 1$, $j \in 1:n$).

Очевидно, что

$$\bar{s}_i^T D s = \bar{s}_i^T D t - \sum_{k=r}^{m-1} (\bar{s}_k^T D t) (\bar{s}_i^T D \bar{s}_k) / \bar{s}_k^T D \bar{s}_k = \bar{s}_i^T D t - \bar{s}_i^T D t = 0$$

при $i \in r:m-1$, т. е. вектор s сопряжен со всеми векторами $\bar{s}_k$, $k \in r:m-1$. Ясно также, что $I(x_k) \supset I(x_m)$ и $J(x_k) \supset J(x_m)$ при $k \in r:m-1$. Кроме того, $a_i \bar{s}_k = 0$ для $i \in I(x_m)$ и $\bar{s}_k[j] = 0$ для $j \in J(x_m)$ при каждом $k \in r:m-1$, так как в противном случае на k -й итерации мы покинули бы одну из ограничивающих гиперплоскостей и вернулись бы на нее на l -й при некотором $l \in k+1:m$, что невозможно при $k \geq r$. Следовательно, $a_i s = a_i t$ при $i \in I(x_m)$ и $s[j] = t[j]$ при $j \in J(x_m)$.

Итак, вектор s удовлетворяет всем ограничениям (9), кроме, быть может, условий $-1 \leq s[j] \leq 1$, $j \in 1:n$. Поэтому $\alpha s \in S_m$ при некотором положительном α .

Покажем теперь, что $\nabla f(x_m)s = \nabla f(x_m)t > 0$. Так как $\lambda_k = \lambda'_k$ при $k \in r:m-1$, то (см. (8) и 2°)

$$\nabla f(x_m) = \nabla f(x_k) - 2 \sum_{i=k}^{m-1} (\nabla f(x_i) \bar{s}_i / \bar{s}_i^T D \bar{s}_i) D \bar{s}_i.$$

Следовательно, $\nabla f(x_m) \bar{s}_k = 0$ для любого $k \in r:m-1$ и

$$\nabla f(x_m)s = \nabla f(x_m)t - \sum_{k=r}^m \alpha_k \nabla f(x_m) \bar{s}_k = \nabla f(x_m)t > 0. \blacksquare$$

5°. Покажем теперь, что алгоритм п. 5.2 позволяет, если $\mathcal{D} \neq \emptyset$, вычислить оптимальный план задачи (1) за конечное число итераций или обнаружить неограниченность f на $\mathcal{D}$.

Доказательство проведем индукцией по размерности многогранного выпуклого множества $\mathcal{D}$. Утверждение очевидно, если $\dim \mathcal{D} = 0$. Предположим, что оно верно, если $\dim \mathcal{D} = p$, и пусть $\dim \mathcal{D} = p+1$. Тогда (см. рекомендацию 2 и 3 алгоритма, а также 1° и 4°) после конечного числа итераций будет либо построен оптимальный план, либо (см. 3°) обнаружена неограниченность f на $\mathcal{D}$, либо точка итеративной последовательности окажется на некоторой ограничивающей гиперплоскости. Поскольку ограничивающих гиперплоскостей конечное чис-

ло, то после конечного числа итераций мы вторично окажемся на одной из них. Тогда в силу рекомендации 1 алгоритма будет осуществлен поиск максимального значения квадратичной функции на грани многогранного выпуклого множества $\mathcal{D}$ размерности, не превосходящей p , который в силу предположения индукции будет завершен после конечного числа итераций. Остается отметить, что число граней множества $\mathcal{D}$ конечно, а повторное возвращение на грань, для которой уже было найдено максимальное значение, невозможно.

Следствие. Если $\mathcal{D} \neq \emptyset$, то задача квадратичного программирования (1) либо имеет решение, либо f не ограничена сверху на $\mathcal{D}$ (ср. с теоремой п. 5.2.1).

УПРАЖНЕНИЯ

- 9.1. См. с. 295. 9.2. См. с. 299. 9.3. См. с. 299. 9.4. См. с. 303. 9.5. См. с. 304.

УКАЗАТЕЛЬ ЛИТЕРАТУРЫ

1. Абрамов Л. М., Капустин В. Ф. Математическое программирование. Л., 1976. 184 с.
2. Белоусов Е. Г. Введение в выпуклый анализ и целочисленное программирование. М., 1977. 196 с.
3. Берг К. Теория графов и ее применения. М., 1962. 320 с.
4. Болтянский В. Г. Оптимальное управление дискретными системами. М., 1973. 448 с.
5. Булавский В. А., Звягина Р. А., Яковлева М. А. Численные методы линейного программирования (специальные задачи)/Под ред. Л. В. Канторовича. М., 1977. 368 с.
6. Гасс С. Линейное программирование (методы и приложения). М., 1961. 304 с.
7. Гейл Д. Теория линейных экономических моделей. М., 1963. 418 с.
8. Гольштейн Е. Г. Выпуклое программирование. Элементы теории. М., 1970. 68 с.
9. Гольштейн Е. Г. Теория двойственности в математическом программировании и ее приложения. М., 1971. 352 с.
10. Гольштейн Е. Г., Юдин Д. Б. Задачи линейного программирования транспортного типа. М., 1969. 384 с.
11. Данил Дж. Линейное программирование, его обобщения и применения. М., 1966. 600 с.
12. Еремин И. И., Астафьев Н. Н. Введение в теорию линейного и выпуклого программирования. М., 1976. 192 с.
13. Зангвилл У. И. Нелинейное программирование. Единый подход. М., 1973. 310 с.
14. Заславский Ю. Л. Сборник задач по линейному программированию. М., 1969. 256 с.
15. Зойтендейк Г. Методы возможных направлений. М., 1963. 176 с.
16. Зуховицкий С. И., Авдеева Л. И. Линейное и выпуклое программирование. М., 1964. 348 с.
17. Калихман И. Л. Сборник задач по математическому программированию. М., 1975. 270 с.

18. Канторович Л. В. Математические методы организации и планирования производства. Л., 1939. 67 с.
19. Канторович Л. В. О перемещении масс. — Докл. АН СССР, 1942, т. 37, № 7—8, с. 227—229.
20. Канторович Л. В. Экономический расчет наилучшего использования ресурсов. М., 1959. 348 с.
21. Канторович Л. В., Гавурин М. К. Применение математических методов в вопросах анализа грузопотоков. — В кн.: Проблемы повышения эффективности работы транспорта. М., 1949. с. 110—138.
22. Канторович Л. В., Горстко А. Б. Оптимальные решения в экономике. М., 1972. 232 с.
23. Канторович Л. В., Залгаллер В. А. Рациональный раскрой промышленных материалов. Новосибирск, 1971. 300 с.
24. Капустин В. Ф. Практические занятия по курсу математического программирования. Л., 1976. 192 с.
25. Карманов В. Г. Математическое программирование. М., 1975. 272 с.
26. Карлин С. Математические методы в теории игр, программировании и экономике. М., 1964. 839 с.
27. Корбут А. А., Фиякельштейн Ю. Ю. Дискретное программирование. М., 1968. 368 с.
28. Кофман А., Анри-Лабордер А. Методы и модели исследования операций. М., 1977. 432 с.
29. Кюнде Г. П., Крелле В. Нелинейное программирование. М., 1965. 304 с.
30. Линейные неравенства и смежные вопросы/Под ред. Г. У. Куна, А. У. Таккера. М., 1959. 472 с.
31. Ляшенко И. Н., Карагодова Е. А., Черникова Н. В., Шор Н. З. Линейное и нелинейное программирование. Киев, 1975. 372 с.
32. Мухачева Э. А., Рубинштейн Г. Ш. Математическое программирование. Новосибирск, 1977. 320 с.
33. Никайдо Х. Выпуклые структуры и математическая экономика. М., 1972. 520 с.
34. Оре О. Теория графов. М., 1968. 352 с.
35. Полак Э. Численные методы оптимизации. Единый подход. М., 1974. 376 с.
36. Пшеничный Б. Н. Необходимые условия экстремума. М., 1969. 152 с.
37. Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. М., 1975. 320 с.
38. Рокафеллар Р. Выпуклый анализ. М., 1973. 469 с.
39. Романовский И. В. Алгоритмы решения экстремальных задач. М., 1977. 352 с.
40. Рубинштейн Г. Ш. Конечномерные модели оптимизации. Новосибирск, 1970. 229 с.
41. Толстой А. Н. Методы устранения нерациональных перевозок при планировании. — Социалистический транспорт, 1939, № 9, с. 28—51.
42. Трубин В. А. О методе решения задач целочисленного программирования специального вида. — Докл. АН СССР, 1969, т. 189, № 5, с. 552—554.
43. Фиакко А., Мак-Кормик Г. Нелинейное программирование. Методы последовательной безусловной минимизации. М., 1972. 240 с.
44. Хелли Дж. Нелинейное и динамическое программирование. М., 1967. 507 с.
45. Ху Т. Целочисленное программирование и потоки в сетях. М., 1974. 520 с.
46. Шварц Л. Анализ. Т. 1. М., 1972. 824 с.
47. Шилов Г. Е. Математический анализ (конечномерные линейные пространства). М., 1969. 432 с.

48. Эрроу К. Дж., Гурвиц Л., Удзава Х. Исследования по линейному и нелинейному программированию. М., 1962. 335 с.

49. Юдин Д. Б., Гольштейн Е. Г. Задачи и методы линейного программирования. М., 1964. 736 с.

50. Юдин Д. Б., Гольштейн Е. Г. Линейное программирование. Теория и конечные методы. М., 1963. 776 с.

51. Юдин Д. Б., Гольштейн Е. Г. Линейное программирование (теория, методы и приложения). М., 1969. 424 с.

52. Little I. D. C., Murty K. G., Sweeney D. W., Karel C. An algorithm for the traveling salesman problem. — Operat. Res., 1963, vol. 11, № 6, p. 972—989. Русск. пер.: Литл Дж. Мурти К., Сьюни Д., Керел К. Алгоритм для решения задачи о коммивояжере. — Экономика и математические методы, 1965, т. 1, № 1, с. 94—107.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Алгоритм венгерского метода 247
— генерирования потока 238
— градиентного метода 182
— двойственного симплекс-метода 169
— дробления 273
— метода потенциалов 239
— — решения задачи линейного программирования с ограниченными сверху переменными 195
— наискорейшего спуска Коши 293
— Ньютона 293
— решения параметрической задачи с параметром в целевой функции 257
— симплекс-метода 152
- Базис двойственно допустимый 169
— задачи 149
— — допустимый 151
— канонический 7, 8
— θ -оптимальный 258
Блок задачи 219
Булева переменная 67
- Вектор гиперплоскости нормальный 11
— двойственных оценок 132
— ограничений 122
- Векторы многогранного выпуклого конуса образующие 44
— сопряженные относительно матрицы 319
- Вершина выпуклого многогранника 33
— графа 229
— висячая 233
— многогранного выпуклого множества 47
- Вершины смежные дуги 229
— — ребра 229
- Генерирование 213
Гиперплоскость (10):
— ограничивающая 34
Гиперплоскость опорная 29
- разделяющая 27
Гиперплоскости, независимое множество 34
Градиент 15
Граница относительная 24
Грань многогранного выпуклого множества 47
Граф (229):
антисимметрический 231
конечный 229
неориентированный (неограф) см. Неограф
ориентированный (ограф) см. Орограф
планарный 230
полный 231
симметрический 231
частичный 229
График функции 53
- Дерево 232
— частичное 233
Дуга, вершины 229
— графа 229
— инцидентная вершинам 229
Дуги смежные 229
- Задача безусловной оптимизации 67
— выбора направления 317
— выпуклого программирования 68
— — — целочисленная 70
— вырожденная 151, 169
— вычисления шага 294
— двойственная 120
— двойственно невырожденная 169
— дискретного программирования 70
— Канторовича 77
— —, линейная форма 87
— квадратичного программирования 70
— — — каноническая 102
- Задача коммивояжера 79
—, корректирующее преобразование 84

- линейного программирования (69):
 - блочная (разложимая) 218
 - каноническая 70
 - форма 87
 - общая 69
 - параметрическая 254
 - с параметром в векторе ограничений 255
 - — — целевой функции 255
 - экстремальная 254
 - с ограниченными сверху переменными 90
 - — —, каноническая форма 91
 - целочисленная 70
 - каноническая 268
 - частично целочисленная 70
- максимизации 66
- минимизации 66
- невырожденная 137
- о диете 78
- о назначениях (задача выбора) 78
- о раскрое 82
- с булевыми переменными 70
- сглаживания 83
- станковая (распределительная) 76
- —, линейная форма 86
- транспортная:
 - в матричной постановке 73
 - — — с ограниченными сверху переменными 74
 - в сетевой постановке 74
 - — — переменными 74, 234
 - — — с ограниченными сверху с нарушенным балансом 246
- угловая 220
- целочисленного программирования 70
- частично целочисленного программирования 70
- эквивалентные преобразования 84
- экстремальная 66
- —, область определения 66
- —, ограничения 67
- Задачи линейного программирования
 - двойственные 122
 - — —, пара 123
 - сопряженные 118
 - эквивалентные 84
- Защипливание 190, 266
- Звенья маршрута (276):
 - запрещенные 276
 - —, цепочки 276
 - фиксированные 276
- Итерация 138
- Комбинация линейная, аффинная 10
 - — выпуклая 18
- Компонента вектора (плана) 8
 - плана базисная 151
 - — небазисная 151
- связности графа 232
- Контур 231
- Конус (37):
 - выпуклый 38
 - многогранный 43
- замкнутый 38
- неограниченности 39
- Коррекция (корректирующее отображение) 284
- неоднозначная 285
- Критерий оптимальности допустимого плана (92):
 - задачи Канторовича 106, 107
 - квадратичного программирования 102
 - — — канонической 102
 - линейного программирования канонической 104
 - — общей 103
 - — — с ограниченными сверху переменными 109
- транспортной задачи в матричной постановке 110
- рабочий 100
- Луч 9
- Лучи многогранного выпуклого конуса образующие 44
- Маршрут 276
- Матрица блочная 218
 - Гессе 16
 - задачи 182
 - неотрицательно определенная 11
 - неположительно определенная 11
 - отрицательно определенная 11
 - положительно определенная 11
 - транспортной задачи 110
 - угловой задачи 220
 - унимодулярная 270
- Метод венгерский 247
 - ветвей и границ 247
 - — — для решения задачи коммивояжера 276
- внешних штрафных функций 300
- внутренних штрафных функций 304
- возмущения 190
- Вулфа 204
- градиентный 180
- декомпозиции (разложения) Данцига — Вулфа 217
- деление пополам (метод Больцано) 296
- допустимых направлений (138):
 - бесконечный 139
 - конечный 139
 - сходящийся 139

Метод золотого сечения поиск Фибоначчи) 295

- исключения Гаусса 181
 - локальных вариаций 297
 - минимального элемента 216
 - минимизации невязок 141
 - нахождения допустимого потока 245
 - потенциалов 239
 - решения задачи линейного программирования с ограниченными сверху переменными 193
 - северо-западного угла 216
 - сопряженных направлений 316
- Методы ветвей и границ 272
- итеративные 138
 - —, принципиальная схема 284
 - последовательного улучшения плана 138

Многогранник выпуклый 33

- — целочисленный 269

Множества индексов σ -активные 308

- концевые 273
- отделимые 27
- строго отделимые 27

Множество аффинное

- выпуклое (17):
 - многогранное 33
 - целочисленное 269
- допустимых планов 67
- относительно открытое 25
- приемлемых точек 284
- экстремальное 31

Модель динамическая 72

- Канторовича 77
- станковая 76
- статическая 72
- транспортная 73

Надграфик функции 52

Направление (38):

- допустимое 93
- неограниченности множества 38
- прогрессивное 93

Неорграф 229

- связный 231

Оболочка аффинная 10

- выпуклая 18

Окрестность относительная 22

Орграф 229

- связанный 231

Остов 236

Отображение аффинное 14

Отрезок 9

Оценка множества дробления 274

План (67):

- базисный 151

План $\ast$ -базисный 194

- вырожденный 194

План невырожденный 194

- целочисленный 268
- вырожденный 151, 210
- двойственно допустимый 169
- допустимый 67
- невырожденный 151, 169
- оптимальный (решение задачи) 151
- целочисленный 268

Полупространство 11

- замкнутое 11

Порядок графа 229

Последовательность внешних штрафных функций 299

- внутренних штрафных функций 303

Поток (233):

- $\ast$ -базисный 237
- вырожденный 237
- невырожденный 237
- допустимый 233
- недопустимый 233

Правило ввода 205

Приведение матрицы 277

Пропускная способность дуги 233

Процесс итеративный 138

Прямая 9

Прямоугольник 12

- замкнутый 12

Псевдоплан 173

Путь (231):

- простой 231
- составной 231
- элементарный 231

Размерность аффинного множества 10

- выпуклого множества 24

- подпространства 8

Ребра смежные 229

Ребро графа 229

- многогранность выпуклого множества 47

Рекорд 275

Семейство экстремальных задач параметрическое 254

Симплекс 21

Симплекс-метод 148

- двойственный 168

Симулятор 215

Симуляция 215

Система ограничений (условий) задачи 67

Стоимость маршрута 276

- потока 233

Столбец бедный 248

- везущий 153

Строка богатая 248
— ведущая 153
Сфера 12
Схема одностороннего ветвления 275
— полного ветвления 274
— с приближенной коррекцией 286
Теорема Каратеодори 19
— сходимости метода внешних штрафных функций 300
— — — внутренних штрафных функций 305
— — — допустимых направлений 312
— Фаркаша 46
— Чарнса 191
Теоремы двойственности 125
— Куна — Таккера 112
Технологический способ 76
Точка локального максимума (минимума) 16
— относительно внутренняя 22
— — граничная 24
— приемлемая, приближенное вычисление 285
— седловая 111, 112
— —, неравенства 112
— стационарная 16
— экстремальная 31
— экстремума 16
Транспортная сеть (233):
 несбалансированная (открытая) 233
 сбалансированная 233
Унимодулярность матрицы ограниченной транспортной задачи 271
Уравнение аффинного множества 10
— гиперплоскости 11
— подпространства 8
Условие Сильвестра 11
— — — Слейтера С 97
— — — относительно системы ограничений 97

Условие регулярности 97
— сопряженности векторов 319
— существования решения задачи линейного программирования 131
— — — канонической задачи линейного программирования 128
— унимодулярности матрицы 270
Условия допустимости направления 94
— прогрессивности направления 96
Функция аффинная 14
— вогнутая 52
— — на множестве 54
— выпуклая 52
— —, эффективная область определения 53
— — на множестве 54
— квазивыпуклая 69
— Лагранжа задачи 111
— строго вогнутая 54
— — выпуклая 54
— целевая, или целевая функция задачи 66
— штрафные внешние 298
— — внутренние 303
Цепь (231):
 простая 231
 элементарная 231
Цикл (231):
 простой 231
 элементарный 231
Число приведения матрицы 277
Шар 12
— замкнутый 12
Элемент ведущий 153
Ядро открытое 13
— — — относительно открытое 22