

ВВЕДЕНИЕ В ТЕОРИЮ МАТРИЦ

Р. БЕЛЛМАН



Р. БЕЛЛМАН

ВВЕДЕНИЕ В ТЕОРИЮ МАТРИЦ

Перевод с английского под редакцией
В. Б. ЛИДСКОГО

ИЗДАНИЕ ВТОРОЕ



ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1976

517.1
Б 43
УДК 512.83

INTRODUCTION TO MATRIX ANALYSIS

by *Richard Bellman*

MATHEMATICIAN,
THE RAND CORPORATION

MCGRAW-HILL BOOK COMPANY, INC.

New York Toronto London

1960

Введение в теорию матриц. Беллман Р.

Книга посвящена изложению теории матриц и ее приложениям к теории дифференциальных уравнений, математической экономике, теории вероятностей. Монография написана так, что ее может читать студент, не изучавший ранее линейную алгебру. В книге имеется более 600 задач; многие из них подводят читателя к самостоятельной научной деятельности в области теории матриц. Ценность книги увеличивают приводимые в конце каждой главы обзоры оригинальных работ в соответствующей области.

Книга рассчитана на студентов университетов и вузов, на инженеров, физиков, механиков, использующих матричный аппарат. Много привлекательного найдет в ней и математик, интересующийся собственно теорией матриц. Страниц 352. Иллюстраций 2.

20203—058
Б $\frac{053(02)-76}{32-76}$

ОГЛАВЛЕНИЕ

Предисловие редактора перевода	8
Предисловие автора к русскому изданию	10
Предисловие автора к английскому изданию	11
Глава 1. Максимизация и минимизация. Обоснование	21
§ 1. Введение (21). § 2. Максимизация функции одной переменной (21). § 3. Максимизация функции двух переменных (22). § 4. Алгебраический подход (23). § 5. Аналитический подход—I (25). § 6. Аналитический подход—II (26). § 7. Упрощающее преобразование (28). § 8. Другое необходимое и достаточное условие (29). § 9. Определенные и не- определенные формы (29). § 10. Геометрический подход (29). § 11. Об- суждения (31).	
Упражнения к гл. 1	31
Библиография и комментарий	32
Глава 2. Векторы и матрицы	33
§ 1. Введение (33). § 2. Векторы (33). § 3. Сложение векторов (34). § 4. Умножение вектора на скаляр (35). § 5. Скалярное произведение двух векторов (35). § 6. Ортогональность (36). § 7. Матрицы (37). § 8. Умно- жение вектора на матрицу (38). § 9. Умножение матрицы на матрицу (39). § 10. Некоммутативность (41). § 11. Ассоциативность (42). § 12. Ин- вариантные векторы (43). § 13. Квадратичная форма как скалярное произведение (43). § 14. Транспонированная матрица (44). § 15. Симме- трические матрицы (45). § 16. Эрмитовы матрицы (45). § 17. Ортого- нальные матрицы. Инвариантность расстояний (46). § 18. Унитар- ные матрицы (47).	
Упражнения к гл. 2	48
Библиография и комментарий	52
Глава 3. Диагонализация и канонические формы симметрических матриц	54
§ 1. Резюме (54). § 2. Решение системы линейных однородных урав- нений (54). § 3. Собственные векторы и собственные значения (56). § 4. Два фундаментальных свойства симметрических матриц (57). § 5. Приведение к диагональной форме. Различные собственные значе- ния (59). § 6. Приведение квадратичной формы к каноническому ви- ду (60). § 7. Положительно определенные квадратичные формы и матрицы (62).	
Упражнения к гл. 3	63
Библиография и комментарий	64
Глава 4. Приведение симметрических матриц к диагональной форме в общем случае	65
§ 1. Введение (65). § 2. Линейная зависимость (65). § 3. Ортогонализа- ция Грама—Шмидта (65). § 4. Положительность определителей Грама	

D_n (68). § 5. Одно тождество (70). § 6. Диагонализация симметрической матрицы второго порядка (72). § 7. N -мерный случай (73). § 8. Необходимое и достаточное условие положительной определенности (76). § 9. Собственные векторы, соответствующие кратным собственным значениям (76). § 10. Теорема Гамильтона—Кэли для симметрических матриц (77). § 11. Одновременное приведение к диагональной форме (78). § 12. Одновременное приведение к сумме квадратов (80). § 13. Эрмитовы матрицы (81). § 14. Исходная проблема максимизации (82). § 15. Теория возмущений — I (82). § 16. Теория возмущений — II (83).	
Упражнения к гл. 4	86
Библиография и комментарий	92
Глава 5. Условные экстремумы	95
§ 1. Введение (95). § 2. Детерминантный критерий положительной определенности (критерий Сильвестра) (95). § 3. Представление в виде суммы квадратов (97). § 4. Связанные вариации и теорема Финслера (98). § 5. Случай $k=1$ (100). § 6. Задача о минимизации (103). § 7. Общий случай (105). § 8. Прямоугольные матрицы (105). § 9. Клеточные матрицы (106). § 10. Решение задачи в общем случае (107).	
Упражнения к гл. 5	108
Библиография и комментарий	110
Глава 6. Функции от матрицы	112
§ 1. Введение (112). § 2. Функции от симметрической матрицы (112). § 3. Обратная матрица (113). § 4. Единственность обратной матрицы (113). § 5. Квадратные корни (116). § 6. Параметрическое представление (117). § 7. Результат Шура (117). § 8. Основные скалярные функции (118). § 9. Несобственный интеграл $\int_{-\infty}^{\infty} e^{-(x, Ax)} dx$ (119).	
§ 10. Аналог для эрмитовых матриц (121). § 11. Связь между $J(H)$ и $ H $ (122).	
Упражнения к гл. 6	122
Библиография и комментарий	128
Глава 7. Вариационное описание характеристических чисел	134
§ 1. Введение (134). § 2. Отношение Релея (134). § 3. Вариационные свойства характеристических чисел (135). § 4. Обсуждение (136). § 5. Геометрические предпосылки (137). § 6. Теорема Куранта—Фишера о минимаксном представлении характеристических чисел (137). § 7. Монотонное поведение $\lambda_k(A)$ (140). § 8. Теорема отделения Штурма (140). § 9. Необходимое и достаточное условие положительной определенности матрицы A (141). § 10. Теорема отделения Пуанкаре (141). § 11. Теорема о представлении (142). § 12. Приближенные методы (143).	
Упражнения к гл. 7	145
Библиография и комментарий	146
Глава 8. Неравенства	149
§ 1. Введение (149). § 2. Неравенство Коши—Шварца (149). § 3. Интегральный вариант (150). § 4. Неравенство Гёльдера (150). § 5. Вогнутость $ A $ (152). § 6. Одно полезное неравенство (152). § 7. Неравенство Адамара (153). § 8. Вогнутость произведения $\lambda_N \lambda_{N-1} \dots \lambda_k$ (154). § 9. Аддитивные неравенства, вытекающие	

из мультипликативных (155). § 10. Другой путь (156). § 11. Более простое выражение для $\lambda_1 \lambda_2 \dots \lambda_k$ (157). § 12. Неравенство между средним арифметическим и средним геометрическим (158). § 13. Мультипликативные неравенства, вытекающие из аддитивных (159).	
Упражнения к гл. 8	160
Библиография и комментарий	162

Глава 9. Динамическое программирование 167

§ 1. Введение (167). § 2. Задача наименьшего отклонения (167). § 3. Функциональное уравнение (168). § 4. Рекуррентные соотношения (169). § 5. Более сложный пример (169). § 6. Проблема Штурма—Лиувилля (170). § 7. Функциональные уравнения (172). § 8. Матрицы Якоби (173). § 9. Аналитическое продолжение (174). § 10. Несимметрические матрицы (175). § 11. Случай комплексной матрицы A (175). § 12. Слабо связанные системы (177). § 13. Упрощения—I (178). § 14. Упрощения—II (178). § 15. Уравнение $Ax = y$ (178). § 16. Квадратичное уклонение (180). § 17. Результат Стильбеса (182).	
Упражнения к гл. 9	183
Библиография и комментарий	184

Глава 10. Матрицы и дифференциальные уравнения 187

§ 1. Обоснование (187). § 2. Векторно-матричные обозначения (188). § 3. Нормы векторов и матриц (189). § 4. Бесконечные ряды векторов и матриц (191). § 5. Существование и единственность решений линейной системы уравнений (191). § 6. Матричная экспонента (194). § 7. Функциональные уравнения—I (194). § 8. Функциональные уравнения—II (195). § 9. Функциональные уравнения—III (195). § 10. невырожденность решения (196). § 11. Решение неоднородного уравнения. Постоянные коэффициенты (197). § 12. Неоднородное уравнение. Переменные коэффициенты (198). § 13. Неоднородное уравнение. Сопряженная система (198). § 14. Теория возмущений (199). § 15. Неотрицательность решения (201). § 16. Функциональное уравнение Поля (202). § 17. Уравнение $\frac{dX}{dt} = AX + XB$ (204). § 18. Уравнение $AX + XB = C$ (205).	
Упражнение к гл. 10	206
Библиография и комментарий	206

Глава 11. Явные решения и канонические формы матриц 212

§ 1. Введение (212). § 2. Метод Эйлера (212). § 3. Построение решения (213). § 4. невырожденность матрицы C (214). § 5. Другой метод (214). § 6. Определитель Вандермонда (215). § 7. Явная форма решения линейного дифференциального уравнения. Диагональные матрицы (216). § 8. Диагонализация матрицы (217). § 9. Связь между двумя подходами (218). § 10. Кратные характеристические числа (219). § 11. Каноническая форма Жордана (220). § 12. Кратные характеристические числа (другой метод) (221). § 13. Треугольная форма матрицы. Теорема Шура (224). § 14. Нормальные матрицы (226). § 15. Теорема об аппроксимации (227). § 16. Другая теорема об аппроксимации (229). § 17. Теорема Гамильтона—Кэли (229). § 18. Другое доказательство теоремы Гамильтона—Кэли (230). § 19. Линейные уравнения с периодическими коэффициентами (231). § 20. Представление невырожденной матрицы α	
---	--

виде экспоненты (232). § 21. Другое доказательство (234). § 22. Некоторые интересные преобразования (235). § 23. Биортогональность (235). § 24. Преобразование Лапласа (237). § 25. Пример (238). § 26. Обсуждение результата (240). § 27. Матричный случай (240).

Упражнения к гл. 11 241
Библиография и комментарий 249

Глава 12. Симметрические функции, кронекеровские произведения и циркулянты 252

§ 1. Введение (252). § 2. Степени собственных значений (252). § 3. Полиномы и характеристические уравнения (254). § 4. Симметрические функции (254). § 5. Кронекеровские произведения (255). § 6. Алгебра кронекеровских произведений (256). § 7. Кронекеровские степени—I (256). § 8. Кронекеровские степени—II (257). § 9. Кронекеровские степени—III (257). § 10. Кронекеровский логарифм (258). § 11. Кронекеровская сумма—I (259). § 12. Кронекеровская сумма—II (259). § 13. Уравнение $AX + XB = C$ (259). § 14. Другое доказательство (261). § 15. Циркулянты (263).

Упражнения к гл. 12 264
Библиография и комментарий 266

Глава 13. Теория устойчивости 269

§ 1. Введение (269). § 2. Необходимые и достаточные условия устойчивости (270). § 3. Устойчивые матрицы (271). § 4. Метод Ляпунова (271). § 5. Среднеквадратичное отклонение (273). § 6. Некоторые эффективные критерии устойчивости (273). § 7. Необходимое и достаточное условие устойчивости матриц (275). § 8. Дифференциальные уравнения и собственные значения (275). § 9. Эффективные условия устойчивости матриц (277).

Упражнения к гл. 13 278
Библиография и комментарий 279

Глава 14. Марковские матрицы и теория вероятностей 283

§ 1. Введение (283). § 2. Простой стохастический процесс (283). § 3. Марковские матрицы и вероятностные векторы (285). § 4. Аналитическое описание дискретных марковских процессов (285). § 5. Асимптотическое поведение (286). § 6. Первое доказательство (287). § 7. Второе доказательство независимости от начального состояния (289). § 8. Некоторые свойства положительных марковских матриц (289). § 9. Второе доказательство сходимости (290). § 10. Марковские матрицы общего вида (292). § 11. Непрерывный стохастический процесс (293). § 12. Доказательство вероятностных свойств (294). § 13. Обобщенные вероятности: унитарные преобразования (295). § 14. Обобщенные вероятности: матричные преобразования (296).

Упражнения к гл. 14 298
Библиография и комментарий 300

Глава 15. Случайные матрицы 302

§ 1. Введение (302). § 2. Предельное поведение физических систем (302). § 3. Средние значения (303). § 4. Средние значения квадратов (304).

Упражнения к гл. 15 304
Библиография и комментарий 304

Глава 16. Положительные матрицы, теорема Перрона и математическая экономика	307
§ 1. Введение (307). § 2. Некоторые процессы простого роста (307). § 3. Обозначения и определения (308). § 4. Теорема Перрона (309). § 5. Доказательство теоремы 1 (309). § 6. Второе доказательство простоты $\lambda(A)$ (311). § 7. Доказательство свойства минимальности $\lambda(A)$ (312). § 8. Эквивалентное определение $\lambda(A)$ (313). § 9. Предельная теорема (313). § 10. Стационарный рост (313). § 11. Непрерывные процессы роста (314). § 12. Аналог теоремы Перрона (315). § 13. Ядерный распад (315). § 14. Математическая экономика (316). § 15. Матрицы Минковского—Леонтьева (319). § 16. Положительность определителя $ I - A $ (320). § 17. Усиление теоремы 6 (320). § 18. Линейное программирование (321). § 19. Теория игр (322). § 20. Марковские процессы принятия решений (323). § 21. Экономическая модель (324).	
Упражнения к гл. 16	325
Библиография и комментарий	328
Приложение А. Линейные уравнения и ранг	334
§ 1. Введение (334). § 2. Определители (334). § 3. Свойство алгебраических дополнений (335). § 4. Правило Крамера (335). § 5. Однородные системы (335). § 6. Ранг (339). § 7. Ранг квадратичной формы (339). § 8. Закон инерции (Якоби—Сильвестра) (339). § 9. Сигнатура (340).	
Упражнения к приложению А	340
Библиография и комментарий	341
Приложение Б. Метод Эрмита	342
Приложение В. Моменты и квадратичные формы	344
§ 1. Введение (344). § 2. Обозначения (344). § 3. Метод Фишера (345). § 4. Моментное представление (346). § 5. Результат Герглотца (347).	
Библиография и комментарий	348
Предметный указатель	349

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Предлагаемый читателю перевод книги известного американского математика Р. Беллмана «Введение в теорию матриц» вызовет безусловный интерес не только у физиков, механиков и инженеров, использующих матричный аппарат, но и у специалистов-математиков, интересующихся математическими аспектами теории.

В книге излагается собственно теория матриц и ее приложения к теории дифференциальных уравнений, теории вероятностей, математической экономике, проблеме отыскания экстремума в случае большого числа переменных и другим вопросам.

Помимо основного текста, в котором, как правило, приводятся доказательства сравнительно несложных утверждений, автор в конце каждой главы помещает значительное число задач, углубляющих и развивающих затрагиваемые вопросы. Часть задач уводит читателя далеко за рамки основного текста.

Как отмечает автор книги, задачи не расположены в тексте в порядке возрастающей трудности. Мы сохранили расположение задач, хотя это, на наш взгляд, и вызовет у читателя известные затруднения.

Несмотря на то, что на русском языке имеется прекрасная монография Ф. Р. Гантмахера по теории матриц, перевод книги Р. Беллмана следует признать целесообразным. Уступая книге Ф. Р. Гантмахера в систематичности, стройности и последовательности изложения, книга Р. Беллмана отличается широким охватом новых проблем теории матриц и ее приложений.

Ценность книги повышают приводимые автором в конце каждой главы обзоры, в которых названы и прокомментированы оригинальные статьи и новые результаты в соответствующем направлении.

При переводе книги были устранены опечатки и прочие погрешности, которых в оригинале обнаружилось значительное количество. Некоторые доказательства, оказавшиеся неудачными, были заменены.

При редактировании были сделаны примечания и составлен дополнительный список литературы.

Число работ по теории матриц непрерывно растет. Мы указали лишь те из них, которые имеют непосредственное отношение к вопросам, обсуждаемым в книге.

В проверке задач (их в книге более шестисот) мне помог А. Я. Белянков. Пользуясь случаем, приношу ему свою признательность.

В. Б. Лидский

ПРЕДИСЛОВИЕ АВТОРА К РУССКОМУ ИЗДАНИЮ

Перевод этой книги на русский язык—для меня честь и большая радость. Различие в языках ставит преграды между народами, и переводы служат благородной цели осуществления свободного обмена, который иначе был бы невозможен. Они позволяют узнать все то прекрасное и доброе, что создано другими культурами.

Поскольку теория матриц—одна из самых изящных частей математического анализа, естественно, чтобы русские работы в этой области переводились на английский язык и наоборот. Я хотел бы надеяться, что эту книгу прочтут с тем же удовольствием, с каким я ее писал.

Еще одно замечание. Сейчас трудное время для всех, кто заинтересован в мире и процветании народов. Будем надеяться, что тот факт, что в США и в СССР так много поборников правды и красоты, послужит делу объединения наших народов и что дружба между нашими народами приведет к счастливой жизни, к которой все стремятся.

Ричард Беллман

*Лос-Анджелес, Калифорния
Университет Южной Калифорнии*

ПРЕДИСЛОВИЕ АВТОРА К АНГЛИЙСКОМУ ИЗДАНИЮ

Цель настоящего тома — ввести читателя в круг идей и методов теории матриц, которую можно справедливо считать арифметикой высшей математики.

Постараемся оправдать наше решение отвести теории матриц столь важную роль. Рассматривая в целом любую из классических областей математики, нетрудно заметить, что наиболее интересные и важные ее разделы связаны с анализом взаимодействия различных факторов. Одним из примеров описания этого взаимодействия служат функции от нескольких переменных. Анализ таких функций приводит к преобразованиям многомерного типа.

Довольно быстро становится ясным, что само описание возникающих при этом задач сопряжено со значительными трудностями. Достаточно обратиться к работам, написанным сто лет назад, чтобы убедиться в том, как велика опасность в начале любого исследования потонуть в море арифметических и алгебраических деталей. И это помимо многочисленных понятий и аналитических трудностей.

Таким образом, решительные усилия должны быть направлены прежде всего на создание удобных, гибких и легко воспринимаемых обозначений. Хотя какая-либо количественная оценка зависимости успеха исследования от хорошо придуманных обозначений вряд ли будет иметь смысл, нетрудно привести огромное число примеров, когда решения становились очевидными благодаря удачной формулировке задачи. Напротив, там, где используются неуклюжие и туманные обозначения, для решения потребуются неизмеримо большие усилия. Представьте себе, например, как сложно было бы производить арифметические и алгебраические операции с римскими цифрами.

Хорошо построенными обозначениями нужно стремиться выразить математическую суть задачи, и ни в коем случае не затемнять ее и не отвлекать от нее.

Сделав такое вступление, мы теперь выскажем очень простой силлогизм: матрицы являются наиболее важными преобразованиями

ями — линейными преобразованиями; преобразования образуют сердцевину математики. В этом суть нашего первого утверждения.

Этот том, первый из планируемой к выпуску серии, посвященной результатам и методам современной теории матриц, должен познакомить читателя с основными понятиями теории матриц. Особое внимание в нем будет уделено анализу и приложениям к задачам математической физики, техники и математической экономики.

Теория матриц на том уровне, на котором мы собираемся ее здесь излагать, довольно четко распадается на три части: теория симметрических матриц, которая проникает во все области; матрицы и дифференциальные уравнения — раздел, представляющий интерес в первую очередь для инженера и физика; положительные матрицы, играющие особо важную роль в теории вероятностей и математической экономике.

Хотя мы не пытались связать изложение материала с какими-либо реальными прикладными задачами, мы всегда старались указать на происхождение основных проблем, рассматриваемых в книге.

Изложение начинается с задачи нахождения максимума или минимума функции нескольких переменных. Используя методы дифференциального исчисления, мы показываем, что определение локального максимума или минимума при обычных предположениях о существовании достаточного числа частных производных приводит к соответствующей задаче для функций, значительно более простых, а именно к квадратичным функциям. Таким образом, мы приходим к рассмотрению квадратичных форм, а следовательно, и симметрических матриц.

Сначала рассматривается случай функции двух переменных, где обычных обозначений вполне достаточно для описания всех интересующих нас результатов. Обратившись к случаю более высокого числа измерений, мы убедимся в необходимости новых обозначений. Тем не менее глубокое понимание двумерного случая очень важно, поскольку все важные черты многомерного случая содержатся уже здесь.

Оставив задачу многомерной максимизации, мы обращаемся к введению матричных обозначений. При этом на каждом этапе мы стараемся вводить только те новые символы и понятия, в которых возникает необходимость. Читателя, возможно, удивит, насколько далеко можно продвинуться в изучении теории симметрических матриц, не вводя понятия обратной матрицы.

Следуя этой идее, мы отказались от обычного подхода, при котором новое понятие обратной матрицы заливают потоком результатов, связанных с решением линейных систем уравнений. Никким образом не желая умалить важность такого подхода, отметим все же, что, отказавшись от этого длинного и несколько

утомительного пути, можно все же изложить целый ряд важных и интересных результатов. Понятие линейной независимости вводится в связи с процессом ортогонализации, где оно играет первостепенную роль. В приложении приводится доказательство основного результата, касающегося решений линейных систем, а также обсуждаются некоторые теоремы о ранге матрицы.

Это понятие, очень важное для многих разделов теории матриц, не имеет большого значения в тех разделах, которые мы здесь изучаем.

Очень часто во многих разделах математики от читателя требуют большого количества подготовительного материала, на первый взгляд не имеющего прямого отношения к основным результатам. Мы старались избежать такого положения. Основанием для введения нового понятия для нас служило возникновение конкретной задачи. Это в какой-то мере отражает действительное положение вещей, с которым сталкивается математик в своих исследованиях.

Хотя мы и старались сделать логичным наше изложение, это не было самоцелью. Логика в конце концов является одним из приемов, изобретенных человеческим умом для решения определенных задач. Но математика — это больше, чем логика, это — логика плюс процесс созидания. То, каким образом законы и понятия логики, составляющие орудие математики, используются для получения результатов, вряд ли является логическим процессом, во всяком случае, не более логическим, чем создание симфонии или картины.

Введя квадратные матрицы, центральный объект нашего исследования, мы займемся каноническим представлением действительных квадратичных форм или, что то же, действительных симметрических матриц. Здесь будет установлен очень важный результат теории симметрических матриц, состоящий в том, что всякая действительная симметрическая матрица в определенном смысле эквивалентна диагональной матрице.

Другими словами, многомерные преобразования этого типа можно рассматривать как совокупность одновременно выполняемых одномерных преобразований.

Результаты этих вступительных глав поучительны по ряду причин. Прежде всего они показывают, как существенно упрощаются доказательства, если предположить, что характеристические числа матрицы простые. Во-вторых, на их примере можно убедиться, что трудности, возникающие из-за кратности характеристических чисел, могут быть преодолены при помощи двух мощных методов: индукции и предельного перехода по непрерывности.

Второй из этих методов требует очень осторожного обращения. Без преувеличения можно сказать, что те моменты доказа-

тельств, где мы пользуемся им, являются самыми трудными во всей книге. Именно поэтому мы не везде, где это возможно, пользовались этим приемом и часто лишь ограничивались указанием на его применимость. Читатели могут воспринимать это как вызов или, скорее, призыв провести доказательство этим методом во всех деталях.

Получив диагональное представление, мы приступим к выводу минимаксных свойств характеристических чисел, открытых Курантом и Фишером. Обобщение этих свойств на случай операторов с частными производными, проведенное Курантом, является одним из выдающихся достижений математики.

После этого вполне уместно заняться некоторыми дальнейшими свойствами матриц. В частности, мы изучим некоторые частные функции матриц. Вопрос об общем определении функции от матриц довольно сложен, и мы лишь слегка затронем его.

Далее, мы вернемся к нашей исходной задаче — исследованию области значений квадратичной формы. На этот раз, правда, мы усложним задачу введением линейных ограничений. Это усложнение не только интересно само по себе, но и дает нам прекрасный повод для введения прямоугольных матриц. На этой стадии разумно рассмотреть матрицы, элементы которых сами являются матрицами. Такое обобщение матричных обозначений часто оказывается чрезвычайно полезным.

После этого мы рассмотрим ряд очень интересных матричных неравенств, характеризующих собственные значения, а также различные функции собственных значений. Глава, посвященная неравенствам, является наиболее специальной из всех глав этой книги; она скорее отражает вкусы автора, нежели отвечает потребностям читателя.

В последней из глав, посвященных симметрическим матрицам, рассматривается метод функциональных уравнений динамического программирования. Ряд задач максимизации и минимизации квадратичных форм, а также решение линейных систем изучаются с помощью этого метода. Интересной чертой получаемых аналитических результатов является их явная зависимость от параметров, которые при решении обычно считают постоянными. Вместе с тем рекуррентные соотношения, которые лежат в основе метода динамического программирования, приводят к алгоритмам, нередко оказывающимся удобными с вычислительной точки зрения.

Вторая треть этого тома посвящена приложению теории матриц к решению линейных систем дифференциальных уравнений. У читателя не предполагается предварительного знакомства с дифференциальными уравнениями. Требующиеся теоремы существования и единственности для линейных систем будут доказаны в ходе изложения.

Важную роль в исследовании линейных систем с постоянными коэффициентами играет понятие матричного экспоненциала. Через эту функцию от матрицы выражается решение линейной системы. Случай переменных коэффициентов более сложен. Для того чтобы получить аналогичное выражение для решения в этом случае, нужно ввести понятие матрицанта. Однако рассмотрение этого понятия выходит за рамки этой книги. Заметим, что матрицант играет важную роль в современной квантовой механике.

Хотя экспоненциал и очень удобен для представления решения линейной системы с постоянными коэффициентами, он не дает выражений для отдельных компонент решения. Для этой цели мы пользуемся методом Эйлера нахождения частных решений экспоненциального вида. При этом мы опять приходим к задаче определения характеристических чисел и собственных векторов матрицы. Поскольку здесь матрица уже не является, вообще говоря, симметрической, эта задача много сложнее той, которой мы занимались в предыдущих главах. Хотя и здесь имеется несколько канонических форм, ни одна из них уже не является такой удобной, как диагональная форма, полученная в случае симметрических и эрмитовых матриц.

Представление решения в виде суммы экспонент, а также предельные случаи позволяют нам сформулировать необходимые и достаточные условия того, что все решения однородной системы стремятся к нулевому вектору при стремлении к бесконечности временного параметра. Так мы приходим к задаче устойчивости и формулируем несколько критериев устойчивости. В общем виде эта задача чрезвычайно сложна.

Получив ряд результатов для общего случая не обязательно симметрических матриц, мы обращаемся к задачам, представляющим более специфический интерес. Дана матрица A . Как определить матрицу, собственные значения которой являются определенными функциями от собственных значений A ? Если мы разыскиваем эти функции как некоторые симметрические функции от собственных значений A , то весьма естественным образом мы приходим к одной из важнейших концепций алгебры матриц — к понятию кронекеровского произведения двух матриц. Как мы увидим в заключительной части книги, эта функция от двух матриц появляется также при изучении случайных матриц.

Заключительная часть этого тома посвящена изучению матриц, все элементы которых неотрицательны. Матрицы этого явно специального вида появляются двумя возможными путями: во-первых, при изучении марковских процессов и, во-вторых, при решении различных экономических задач.

Рассмотрение физических истоков этих матриц делает интуитивно ясным большое число важных и интересных предельных теорем, связанных с именами Маркова, Перрона и Фробениуса.

Фундаментальную роль в теории положительных матриц играет, в частности, ряд преобразований, выполненных Виландтом.

Небольшая глава посвящена теории случайных матриц. В этой области мы имеем дело скорее с мультипликативными и, следовательно, некоммутативными операциями, чем с аддитивными и коммутативными. Эти аспекты теории стохастических процессов почти не исследованы. Результаты, которые мы приводим, являются вводными и достаточно элементарными.

Наконец, в серии приложений изложены некоторые дополнительные сведения, которые либо имеют косвенное отношение к основным проблемам книги, либо представляют весьма специальный интерес. Результаты, полученные в теории симметрических матриц Селбергом, Эрмитом и Фишером, являются столь элегантными, что остается только сожалеть, что не было абсолютно никакой возможности включить их в книгу.

Несколько слов к читателю, впервые обращающемуся к этой пленительной области. Как указывалось выше, этот том написан как введение в теорию матриц. Хотя все главы являются вводными в этом смысле, однако, перефразируя слова Оруэла, можно сказать, что некоторые из них являются более вводными, чем другие.

Следовательно, не обязательно читать главы одну за другой. Начинающему настоятельно советуем прочесть главы 1—5 в качестве общего введения в теорию матричных операций и как зачатки теории симметрических матриц. После этого целесообразно перейти к общему изучению квадратных матриц, используя для этой цели гл. 10 и 11. Наконец, для понимания общих основ марковских и неотрицательных матриц могут быть изучены гл. 14 и 16. Совместно с решением ряда упражнений этот материал составит содержание семестрового курса лекций.

Читатель, изучающий теорию матриц самостоятельно, может следовать своей собственной программе.

Поднявшись на этот уровень, можно перейти к следующей важной области—минимаксным свойствам собственных значений матриц. В этой связи предлагается изучить кронекеровское произведение. Здесь кажется подходящей гл. 6. Оставшиеся главы могут читаться после этого в любой последовательности.

Несколько замечаний относительно упражнений. Поскольку цель математиков состоит в том, чтобы решать задачи, то очевидно, что невозможно оценить свой собственный прогресс, не пробуя свои силы на некоторых задачах. Что мы пытались сделать—это предусмотреть задачи разной степени сложности, начиная от чисто иллюстративных и кончая задачами достаточной степени трудности. Эти последние могут быть обычно опущены при беглом ознакомлении с книгой. Задачи, непосредственно следующие за каждым параграфом, являются, вообще говоря, обычными и включены для целей тренировки. Лишь небольшая

часть из них — задачи повышенной трудности. Они содержат результаты, которые используются в последующем тексте. Поскольку соответствующие результаты могут быть установлены без больших усилий, мы считаем, что было бы лучше включить их в виде упражнений, нежели чрезмерно увеличивать основной текст, приводя все доказательства. Во всяком случае, при серьезном чтении тренировка необходима. Задачи в конце каждой главы обладают, напротив, обычно повышенной трудностью. Хотя имеется соблазн как-то выделить эти задачи (например, отметить их звездочкой), как обладающие повышенной трудностью, небольшое раздумье показывает, что педагогически это было бы очень плохо. Целью такой книги, как эта, является, кроме всего прочего, подготовить студента к самостоятельному решению задач в том виде, как они появляются в исследованиях в области анализа, математической физики, в технике, экономике и т. д. Очень редко встречается случай, когда задача, появившаяся таким образом, четко поставлена так, что ее можно считать отмеченной звездочкой. Более того, для студента важно сделать вывод, что сложность задачи практически никогда не оценивается ее формулировкой. Очень быстро можно убедиться, что некоторые весьма простые утверждения могут скрывать значительные трудности. Приняв во внимание эти соображения, мы перемешали задачи всех уровней трудности совершенно случайным образом, без указаний на сложность их решения. Отсюда следует, что, штурмуя эти задачи, читатель должен сделать все, что может, проявляя иногда упорство и, быть может, возвращаясь назад снова и снова по мере роста аналитического умения и зрелости.

Несколько замечаний относительно общего плана книги. Мы хотели не только изложить ряд основополагающих результатов, но и указать на разнообразие фундаментальных методов. Чтобы это сделать, мы в различных местах представляли разные доказательства теорем или указывали на другие возможные подходы в упражнениях. Существенно отдавать себе отчет в том, что, как и во всех частях математики, в теории матриц имеется много разных подходов. Важность владения многими методами вытекает из того факта, что для получения различных обобщений теории требуются различные подходы. На самом деле, некоторые результаты могут быть почти очевидными при использовании одного метода доказательства и в то же время быть трудно достижимыми или вообще недостижимыми, если избрать другой путь рассуждений. Хотя многие результаты теории матриц могут быть успешно и весьма элегантно выведены из общей теории операторов, важно также выделить особые методы и приемы, относящиеся к специальной области, имеющей дело с преобразованиями в конечномерных пространствах. Это

и явилось причиной того, что мы попытались вытащить из забвения ряд простых и эффективных подходов, которые использовались в «старые добрые времена», пятьдесят—семьдесят пять лет назад.

Свойством человеческого мышления объясняется то, что повторение и перекрестные точки зрения являются мощным педагогическим приемом. В этой связи уместно привести слова из «Охоты на Снарка» Льюиса Кэррола: «Я сказал это трижды: то, что я говорю три раза,— правда».

Перечислим теперь кратко некоторые из многих фундаментальных аспектов теории матриц, которые по необходимости не были включены в книгу.

Во-первых, мы исключили обсуждение всех вычислительных аспектов матричной теории. Вычислительные методы получили широкое развитие в последние годы, что стимулировалось необычайными возможностями современных цифровых машин и их планируемым развитием в будущем и, кроме того, необычайно возросшими потребностями физики и экономики.

Необходимость развития новых методов чисто вычислительной природы вытекает из того факта, что задача решения системы линейных уравнений с очень большим числом переменных или задача нахождения собственных значений и собственных векторов матриц большой размерности не может быть решена классическим способом. Любое подходящее решение этой задачи требует использования новых и тонких методов. В этой области проделана чрезвычайно большая работа, и мы подумали, что мудрым решением будет выделить отдельный том этой серии для изложения соответствующих результатов. Этот том будет написан Дж. Форсайтом.

Другим направлением является комбинаторная теория матриц, которая в последние годы растет как снежный ком. Математическая теория игр Бореля и фон Неймана, теория линейного программирования, математическая теория расписаний— все эти области объединяются, что, возможно, приведет к новому разделу теории матриц. На этом пути не только появляются новые и значительные алгебраические и аналитические проблемы, но и как результат этих методов и концепций оказывается определенное давление на способы получения вычислительных процедур. Том, касающийся этой области, будет написан А. Гоффманом.

Топологические аспекты входят классически в теорию матриц при изучении электрических цепей. Здесь мы встречаемся с прекрасным сочетанием аналитической, алгебраической и геометрической теорий.

Том, посвященный этой важной области, должен быть написан Л. Вайнбергом.

В предшествующем перечислении мы, естественно, едва лишь затронули область, называемую теорией матриц. На отдаленном горизонте мы видим монографию, посвященную курсу теории матриц повышенного типа. Среди других результатов она должна содержать различные аспекты теории функций от матриц, теории Лёвнера, зигелевской теории модулярных функций от матриц и теории R -матриц Вигнера. В более общей теории функционалов от матриц теория Бейкера, Кемпбелла и Хаусдорфа приводит к изучению мультипликативных интегралов. Эти аспекты теории взяли на себя главную роль во многих областях современной математической физики.

Другая обширная часть матричного анализа развита в связи с изучением многомерного анализа в математической статистике. Снова мы чувствуем, что результаты в этой области могут быть наилучшим образом изложены совместно с основными концепциями статистики.

На общей базе анализа и алгебры мы встречаем теорию представления групп со многими ее прекрасными приложениями к алгебре, анализу и математической физике. Эта область также требует отдельного тома.

В чисто алгебраической области имеется теория идеалов, рассматриваемая с помощью матриц, введенных Пуанкаре. Мы не упоминали об этом, поскольку использование таких матриц требует введения чисто формальных концепций. Однако в упражнениях постоянно встречаются упоминания о связи между комплексными числами, кватернионами и матрицами наряду с намеками на более общие взаимосвязи.

Изучение матриц с целочисленными элементами тесно связано с числовой теорией матриц. Несмотря на привлекательность этой области, мы чувствовали, что было бы лучше детально обсудить эти вопросы в отдельном томе.

Важно отметить, что даже приведенный выше перечень не отражает той большой роли, которую играют матрицы в современной математике и ее приложениях.

В конец каждой главы и иногда в упражнения мы включили большое количество ссылок на оригинальные статьи, монографии и различные книги по теории матриц. Читатель, интересующийся деталями, может основательно с ними познакомиться, воспользовавшись этими ссылками. Мы ни в коем случае не претендуем на исчерпывающую библиографию. Многие из значительных работ не были упомянуты.

В заключение мне хотелось бы выразить сердечную благодарность моим друзьям, скрупулезно просмотревшим несколько набросков рукописи. Благодаря их замечаниям и критике были сделаны существенные улучшения в содержании, стиле изложения и форме многих интересных упражнений.

Благодарю Поля Брока, Фань Цзы и Ольгу Таусски.

Искренне благодарю также А. Маданского и И. Олкина, прочитавших некоторые главы и сделавших замечания по улучшению некоторых интересных задач.

Я особенно рад выразить благодарность RAND Corporation за ее политику в области исследований, которая привела к существенной поддержке моей работы по теории матриц. Это только один из аспектов свободы действий, предоставляемых RAND Corporation с тем, чтобы одновременно развивать научные исследования и служить интересу нации.

Наконец, благодарю моего секретаря Жанетт Гиберт, которая безропотно печатала сотни страниц уравнений, без всяких жалоб делала правку за правкой и преданно помогала мне в вычитке текста.

Ричард Беллман

МАКСИМИЗАЦИЯ И МИНИМИЗАЦИЯ. ОБОСНОВАНИЕ

§ 1. Введение. Цель этой вводной главы состоит в том, чтобы указать на естественную связь, существующую между проблемой определения области значений однородной квадратичной функции и нахождением максимума или минимума произвольной функции двух переменных.

Мы исследуем проблему экстремальных значений квадратичной функции двух переменных (квадратичной формы двух переменных) весьма подробно. Для этого имеется ряд важных причин. Прежде всего мы демонстрируем три различных метода исследования: алгебраический, аналитический и геометрический. Каждый из этих методов при соответствующем истолковании может быть обобщен на многомерный случай, который мы будем рассматривать впоследствии. С методической точки зрения еще более важным является то, что алгебраические и аналитические осложнения, возникающие при исследовании двумерного случая, становятся трудно преодолимыми в N -мерном варианте задачи. Это обстоятельство выявляет острую потребность в новых обозначениях.

Таким образом, подробное изучение двумерного случая доставляет превосходный материал, служащий оправданием при введении новых понятий.

§ 2. Максимизация функции одной переменной. Пусть на замкнутом интервале $[a, b]$ задана действительная функция вещественной переменной x , допускающая разложение в сходящийся ряд Тейлора

$$f(x) = f(c) + (x-c)f'(c) + \frac{(x-c)^2}{2!}f''(c) + \dots \quad (1)$$

в каждой точке открытого интервала (a, b) .

Пусть c — стационарная точка функции $f(x)$, т. е. точка, в которой $f'(x) = 0$, и пусть требуется определить, является ли c точкой относительного максимума, относительного минимума или стационарной точкой более сложной природы.

Если c — стационарная точка, то ряд (1) принимает более простой вид

$$f(x) = f(c) + \frac{(x-c)^2}{2!} f''(c) + \dots \quad (2)$$

В случае, когда $f''(c) = 0$, мы должны рассмотреть последующие члены разложения. Однако если $f''(c) \neq 0$, то знак второй производной определяет характер стационарной точки: функция $f(x)$ имеет относительный минимум в точке $x=c$ при условии, что $f''(c) > 0$, и относительный максимум, если $f''(c) < 0$.

Упражнение

1. Указать достаточное условие относительного максимума функции, если $f''(c) = 0$.

§ 3. Максимизация функции двух переменных. Перейдем к рассмотрению тех же вопросов для функции двух переменных $f(x, y)$, определенной на прямоугольнике $a_1 \leq x \leq b_1$, $a_2 \leq y \leq b_2$ и допускающей разложение в сходящийся ряд Тейлора в каждой точке (c_1, c_2) внутри этой области.

Для достаточно малых $|x-c_1|$ и $|y-c_2|$ можно записать

$$f(x, y) = f(c_1, c_2) + (x-c_1) \frac{\partial f}{\partial c_1} + (y-c_2) \frac{\partial f}{\partial c_2} + \\ + \frac{(x-c_1)^2}{2} \frac{\partial^2 f}{\partial c_1^2} + (x-c_1)(y-c_2) \frac{\partial^2 f}{\partial c_1 \partial c_2} + \frac{(y-c_2)^2}{2} \frac{\partial^2 f}{\partial c_2^2} + \dots \quad (1)$$

где

$$\begin{aligned} \frac{\partial f}{\partial c_1} &= \frac{\partial f}{\partial x} & \text{при } x=c_1, \quad y=c_2, \\ \frac{\partial f}{\partial c_2} &= \frac{\partial f}{\partial y} & \text{при } x=c_1, \quad y=c_2 \end{aligned} \quad (2)$$

и т. д.

Пусть (c_1, c_2) — стационарная точка функции $f(x, y)$, что означает выполнение следующих равенств:

$$\frac{\partial f}{\partial c_1} = 0, \quad \frac{\partial f}{\partial c_2} = 0. \quad (3)$$

Тогда, как и в предшествующем параграфе, поведение функции $f(x, y)$ в окрестности точки (c_1, c_2) определяется характером квадратичных членов разложения (1), а именно:

$$Q_2(x, y) = a(x-c_1)^2 + 2b(x-c_1)(y-c_2) + c(y-c_2)^2, \quad (4)$$

где для упрощения обозначений принято

$$a = \frac{1}{2} \frac{\partial^2 f}{\partial c_1^2}; \quad 2b = \frac{\partial^2 f}{\partial c_1 \partial c_2}; \quad c = \frac{1}{2} \frac{\partial^2 f}{\partial c_2^2}. \quad (5)$$

Положим еще для удобства

$$x - c_1 = u, \quad y - c_2 = v \quad (6)$$

и перейдем к рассмотрению однородной квадратичной функции

$$Q(u, v) = au^2 + 2buv + cv^2. \quad (7)$$

Выражение такого вида будем называть *квадратичной формой*, в данном случае квадратичной формой двух переменных u и v .

Хотя мы интересуемся поведением $Q(u, v)$ лишь в окрестности нуля, $u = v = 0$, тот факт, что квадратичная форма $Q(u, v)$ является однородной, позволяет нам при желании рассмотреть ее значения при любых вещественных u и v .

Поскольку для любого k имеем $Q(ku, kv) = k^2 Q(u, v)$, значения квадратичной формы $Q(u, v)$ на окружности $u^2 + v^2 = k^2$ весьма просто связаны со значениями $Q(u, v)$ на окружности единичного радиуса $u^2 + v^2 = 1$. Поэтому удобно ограничиться значениями формы на единичной окружности.

Если $Q(u, v) > 0$ для всех отличных от нуля значений u и v , то $f(x, y)$ имеет в точке $x = c_1, y = c_2$ относительный минимум. Если же $Q(u, v) < 0$, то в этой точке имеется относительный максимум. В случае, когда $Q(u, v)$ может принимать как положительные, так и отрицательные значения, мы встречаемся со стационарной точкой более сложного типа, так называемой *седловой точкой*.

Хотя в связи с рассмотрением седловой точки возникает большое число очень интересных алгебраических и геометрических вопросов, в настоящей книге мы не будем их касаться.

Если $Q(u, v)$ — тождественный нуль, то мы приходим к еще более сложной проблеме, к счастью, не имеющей большого значения для приложений.

Упражнение

1. Можно ли свести исследование положительности или отрицательности однородной формы четвертой степени

$$Q(u, v) = a_0 u^4 + a_1 u^3 v + a_2 u^2 v^2 + a_3 u v^3 + a_4 v^4$$

к рассмотрению соответствующей задачи для квадратичных форм?

§ 4. Алгебраический подход. Рассмотрим теперь несколько простых соотношений, связывающих коэффициенты вещественной квадратичной формы a, b и c , которые позволяют ответить на вопрос, с каким из трех описанных выше случаев мы имеем дело.

Для того чтобы определить знак квадратичной формы, дополним выражение $au^2 + 2buv$ до полного квадрата и, предполагая, что $a \neq 0$, запишем $Q(u, v)$ в следующем виде:

$$Q(u, v) = a \left(u + \frac{bv}{a} \right)^2 + \left(c - \frac{b^2}{a} \right) v^2. \quad (1)$$

В случае, если $a = 0$, $c \neq 0$, мы можем выполнить аналогичное преобразование, поменяв местами аргументы u и v . Если же $a = c = 0$, то $Q(u, v)$ становится равной $2buv$. При этом очевидно, что квадратичная форма может принимать как отрицательные, так и положительные значения при $b \neq 0$ и тождественно равна нулю в случае, когда $b = 0$.

Будем в дальнейшем предполагать, что $a \neq 0$, так как в противном случае ход решения задачи ясен из приведенных рассуждений.

Из формулы (1) следует, что $Q(u, v) > 0$ для всех *нетривиальных* значений u и v (т. е. для всех u и v , за исключением нулевой точки $(0, 0)$), если только

$$a > 0, \quad c - \frac{b^2}{a} > 0. \quad (2)$$

Аналогично этому $Q(u, v) < 0$ для всех нетривиальных u и v при условии выполнения неравенств

$$a < 0, \quad c - \frac{b^2}{a} < 0. \quad (3)$$

Справедливо и обратное утверждение. Если квадратичная форма Q положительна для всех нетривиальных значений u и v , то коэффициенты формы удовлетворяют неравенствам (2).

Аналогичный результат справедлив также в случае, когда форма Q отрицательна для всех нетривиальных значений u и v .

Таким образом, мы доказали следующую теорему.

Теорема 1. *Выполнение неравенств*

$$a > 0, \quad \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0 \quad (4)$$

является необходимым и достаточным условием того, чтобы форма $Q(u, v)$ была положительна для всех нетривиальных значений u и v .

Упражнение

1. Показать, что необходимые и достаточные условия положительности формы $Q(u, v)$ для всех нетривиальных значений u и v могут быть записаны в виде $c > 0$, $ac - b^2 > 0$.

Отметим то обстоятельство, что мы снова говорим о необходимых и достаточных условиях. Разумеется, возможно существование ряда различных, но, конечно, эквивалентных, необходимых и достаточных условий. Как правило мы стремимся получить возможно большее число таких условий, так как в различных ситуациях одни из них более удобны, чем другие.

§ 5. Аналитический подход — I. Как отмечалось выше, для того чтобы определить область изменения квадратичной формы $Q(u, v)$, достаточно исследовать значения, принимаемые функцией $Q(u, v)$ на окружности единичного радиуса $u^2 + v^2 = 1$. Если Q положительна для всех нетривиальных значений u и v , то мы должны иметь

$$\min_{u^2+v^2=1} Q(u, v) > 0, \quad (1)$$

в то время как неравенство

$$\max_{u^2+v^2=1} Q(u, v) < 0 \quad (2)$$

есть требуемое условие того, что квадратичная форма $Q(u, v)$ отрицательна для всех значений u и v на единичной окружности.

Для того чтобы исследовать эти вариационные проблемы в симметричной форме, мы воспользуемся методом множителей Лагранжа. Рассмотрим задачу определения стационарных точек новой квадратичной функции

$$R(u, v) = au^2 + 2buv + cv^2 - \lambda(u^2 + v^2). \quad (3)$$

Условия $\frac{\partial R}{\partial u} = \frac{\partial R}{\partial v} = 0$ дают два линейных соотношения:

$$au + bv - \lambda u = 0, \quad bu + cv - \lambda v = 0. \quad (4)$$

Для того чтобы эти уравнения имели нетривиальное решение, множитель Лагранжа λ должен удовлетворять характеристическому уравнению

$$\begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} = 0 \quad (5)$$

или

$$\lambda^2 - (a+c)\lambda + ac - b^2 = 0. \quad (6)$$

Поскольку дискриминант

$$(a+c)^2 - 4(ac - b^2) = (a-c)^2 + 4b^2 \quad (7)$$

неотрицателен, корни уравнения (6) всегда вещественны. Обозначим их через λ_1 и λ_2 . Эти корни различны, если только не выполняются равенства $a=c$, $b=0$. Рассмотрим случай различных корней более подробно.

Если $b=0$, то для корней квадратного уравнения (6) получаем $\lambda_1 = a$, $\lambda_2 = c$.

При $\lambda_1 = a$ из (4) получаем уравнения

$$(a - \lambda_1)u = 0, \quad (c - \lambda_1)v = 0, \quad (8)$$

которые выполняются при произвольном u и $v=0$, если $a \neq c$. Так как мы рассматриваем случай различных корней то последнее

условие выполнено. В случае, если $b \neq 0$, мы получаем нетривиальные решения системы (4), используя лишь одно из уравнений и отбрасывая второе. При этом переменные u и v связаны соотношением

$$(a - \lambda_1)u = -bv. \quad (9)$$

Для того чтобы говорить о частном решении, добавим требование $u^2 + v^2 = 1$. Введение дополнительного условия такого вида называется *нормированием решения*.

Значения u и v , определенные с учетом нормирующего условия, равны

$$\begin{aligned} u_1 &= -\frac{b}{[b^2 + (a - \lambda_1)^2]^{1/2}}, \\ v_1 &= \frac{a - \lambda_1}{[b^2 + (a - \lambda_1)^2]^{1/2}}. \end{aligned} \quad (10)$$

Другая пара решений (u_2, v_2) определяется аналогичным образом при замене λ_1 на λ_2 .

§ 6. Аналитический подход — II. После того как из уравнения (5.6)¹⁾ найдены λ_1 и λ_2 , значения u_i и v_i определяются по формулам (5.10). Эти величины после подстановки дают разыскиваемые минимальное и максимальное значения квадратичной формы $au^2 + 2buv + cv^2$ на окружности $u^2 + v^2 = 1$.

Однако оказывается, что можно действовать значительно более изобретательно. Возвращаясь к линейным уравнениям (5.4) и умножая первое на u и второе на v , мы получаем

$$au^2 + 2buv + cv^2 - \lambda(u^2 + v^2) = 0. \quad (1)$$

Этот результат не является чем-то неожиданным; он представляет собой частный случай теоремы Эйлера, касающейся однородных функций; дело в том, что если $Q(u, v)$ — однородная функция второго порядка, то

$$u \frac{\partial Q}{\partial u} + v \frac{\partial Q}{\partial v} = 2Q. \quad (2)$$

Так как $u_i^2 + v_i^2 = 1$ для $i = 1, 2$, то в силу (2)

$$\begin{aligned} \lambda_1 &= au_1^2 + 2bu_1v_1 + cv_1^2, \\ \lambda_2 &= au_2^2 + 2bu_2v_2 + cv_2^2. \end{aligned} \quad (3)$$

Следовательно, одно решение квадратного уравнения (5.6) есть искомое максимальное значение, а другое — минимальное значение квадратичной формы.

¹⁾ Двойной номер в скобках означает ссылку на формулу другого параграфа той же главы.

В связи с этим отметим то замечательное обстоятельство, что максимальное и минимальное значения $Q(u, v)$ могут быть получены без вычисления координат точек, в которых они достигаются.

Однако, как мы увидим далее, эти точки обладают существенными особенностями.

Выведем теперь важное свойство точек (u_i, v_i) , не пользуясь, однако, выражениями для значений их координат.

Как мы видели в предыдущем параграфе, эти точки определяются системами уравнений

$$\begin{aligned} au_1 + bv_1 - \lambda_1 u_1 &= 0, & au_2 + bv_2 - \lambda_2 u_2 &= 0, \\ bu_1 + cv_1 - \lambda_1 v_1 &= 0, & bu_2 + cv_2 - \lambda_2 v_2 &= 0, \\ u_1^2 + v_1^2 &= 1; & u_2^2 + v_2^2 &= 1. \end{aligned} \quad (4)$$

Рассмотрим первую систему. После умножения первого уравнения на u_2 и второго — на v_2 , складывая, получаем

$$\begin{aligned} u_2 (au_1 + bv_1 - \lambda_1 u_1) + v_2 (bu_1 + cv_1 - \lambda_1 v_1) = \\ = au_1 u_2 + b(u_2 v_1 + u_1 v_2) + cv_1 v_2 - \lambda_1 (u_1 u_2 + v_1 v_2) = 0. \end{aligned} \quad (5)$$

Вторая система аналогично этому дает

$$au_1 u_2 + b(u_2 v_1 + u_1 v_2) + cv_1 v_2 - \lambda_2 (u_1 u_2 + v_1 v_2) = 0. \quad (6)$$

Вычитая из первого уравнения второе, найдем

$$(\lambda_1 - \lambda_2) (u_1 u_2 + v_1 v_2) = 0. \quad (7)$$

Так как по предположению $\lambda_1 \neq \lambda_2$, то

$$u_1 u_2 + v_1 v_2 = 0. \quad (8)$$

Геометрический смысл этого соотношения будет обсуждаться ниже.

Отметим также, что величина $u_1 v_2 - u_2 v_1$ не равна нулю. Действительно, предположим противное.

Тогда, учитывая (8), получаем два уравнения, линейных относительно u_1 и v_1 :

$$\begin{aligned} u_1 u_2 + v_1 v_2 &= 0, \\ u_1 v_2 - v_1 u_2 &= 0. \end{aligned} \quad (9)$$

Так как обе переменные u_1 и v_1 не равны нулю одновременно (вследствие нормирующего условия), должно выполняться равенство

$$\begin{vmatrix} u_2 & v_2 \\ v_2 & -u_2 \end{vmatrix} = 0. \quad (10)$$

Последнее условие противоречит нормирующему условию в системе (4).

Упражнение

1. Показать, что для любой пары точек (u_1, u_2) и (v_1, v_2) имеет место соотношение $(u_1^2 + v_1^2)(u_2^2 + v_2^2) = (u_1u_2 + v_1v_2)^2 + (u_1v_2 - u_2v_1)^2$, и, следовательно, $u_1v_2 - u_2v_1 \neq 0$, если u_i и v_i удовлетворяют приведенным выше условиям.

§ 7. Упрощающее преобразование. Воспользовавшись свойствами (u_i, v_i) (см. равенства (6.4) и (6.8)), посмотрим, что произойдет с квадратичной формой (6.1), если сделать замену переменных

$$u = u_1u' + u_2v', \quad v = v_1u' + v_2v'. \quad (1)$$

Это преобразование является взаимно однозначным, так как определитель $u_1v_2 - u_2v_1$ отличен от нуля.

Прежде всего мы видим, что

$$u^2 + v^2 = (u_1^2 + v_1^2)u'^2 + (u_2^2 + v_2^2)v'^2 + 2(u_1u_2 + v_1v_2)u'v' = u'^2 + v'^2. \quad (2)$$

Из этого следует, что значения, принимаемые квадратичной формой $Q(u, v)$ на окружности $u^2 + v^2 = 1$, совпадают со значениями, принимаемыми следующей квадратичной формой $Q(u_1u' + u_2v', v_1u' + v_2v')$ на окружности $u'^2 + v'^2 = 1$.

Рассмотрим теперь выражение для квадратичной формы Q в новых переменных. Производя преобразования, получаем

$$Q(u_1u' + u_2v', v_1u' + v_2v') = (au_1^2 + 2bu_1v_1 + cv_1^2)u'^2 + \\ + (au_2^2 + 2bu_2v_2 + cv_2^2)v'^2 + 2(au_1u_2 + b(u_1v_2 + u_2v_1) + cv_1v_2)u'v'. \quad (3)$$

С учетом соотношений (6.3), (6.6) и (6.8) окончательно приводим Q к виду

$$\lambda_1 u'^2 + \lambda_2 v'^2. \quad (4)$$

Замена переменных привела к исчезновению перекрестного произведения $2buv$. Подобное упрощение очень удобно для исследования алгебраических, аналитических и геометрических свойств квадратичных форм. Мы будем иметь возможность убедиться в этом в последующих главах, где аналогичное преобразование будет применено для многомерного случая.

Принципиальная часть теории квадратичных форм состоит в том, что подобные результаты имеют место для квадратичных форм любого числа переменных.

Упражнение

1. Основываясь на формуле (4), определить вид конического сечения, описываемого уравнением $Q(u, v) = 1$, для следующих случаев *):

- (a) $\lambda_1 > 0, \lambda_2 > 0$; (c) $\lambda_1 = \lambda_2 > 0$;
 (b) $\lambda_1 > 0, \lambda_2 < 0$; (d) $\lambda_1 = 0, \lambda_2 > 0$.

*) Возможность приведения квадратичной формы к виду (4) была пока установлена лишь в предположении $\lambda_1 \neq \lambda_2$. Однако при $\lambda_1 = \lambda_2$ обязательно будет $a = c, b = 0$ (см. (5.7)), и представление (4) имеет место и в этом случае. (Прим. ред.)

§ 8. Другое необходимое и достаточное условие. Используя полученное представление (7.4), мы можем установить следующее предложение.

Теорема 2. *Необходимое и достаточное условие положительности $Q(u, v)$ для всех нетривиальных значений u и v состоит в положительности корней уравнения*

$$\begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} = 0. \quad (1)$$

Упражнение

1. Показать непосредственно, что условие, установленное выше, эквивалентно утверждению, содержащемуся в теореме 1.

§ 9. Определенные и неопределенные формы. Введем теперь некоторые новые термины. Если квадратичная функция $Q(u, v) = au^2 + 2buv + cv^2$ для всех нетривиальных значений u и v больше нуля, $Q(u, v) > 0$, то мы будем говорить, что $Q(u, v)$ — *положительно определенная форма*. Если $Q(u, v) < 0$ для тех же значений u и v , то мы будем называть $Q(u, v)$ *отрицательно определенной* формой. Если $Q(u, v)$ может принимать как положительные, так и отрицательные значения, то мы будем говорить, что это *неопределенная* форма. Если для всех нетривиальных u и v выполняется неравенство $Q(u, v) \geq 0$, то $Q(u, v)$ называется *неотрицательной* (или неотрицательно определенной) формой. Аналогично определяется *неположительная* форма.

Упражнения

1. Показать, что если $a_1u^2 + 2b_1uv + c_1v^2$ и $a_2u^2 + 2b_2uv + c_2v^2$ — положительно определенные формы, то форма $a_1a_2u^2 + 2b_1b_2uv + c_1c_2v^2$ также положительно определенная.

2. При каких условиях форма $(a_1u_1 + a_2u_2)^2 + (b_1u_1 + b_2u_2)^2$ будет положительно определенной?

3. Как во введенных терминах сформулировать условие того, что уравнение $au^2 + 2buv + cv^2 = 1$ соответствует эллипсу, гиперболе или параболе?

§ 10. Геометрический подход. Приведем теперь иной вариант подхода, рассмотренного выше. Этот последний является чрезвычайно важным для понимания геометрического смысла корней λ_1 и λ_2 и величин (u_i, v_i) . Предположим, что уравнение

$$au^2 + 2buv + cv^2 = 1 \quad (1)$$

соответствует эллипсу, изображенному на рис. 1.

Величина $r = (u^2 + v^2)^{1/2}$ обозначает длину радиуса-вектора от начала координат до точки (u, v) на эллипсе.

Воспользуемся тем фактом, что проблема определения максимума квадратичной формы $Q(u, v)$ на окружности $u^2 + v^2 = 1$

эквивалентна задаче определения минимума $u^2 + v^2$ на кривой $Q(u, v) = 1$.

Формализм, связанный с введением множителя Лагранжа, как и ранее, приводит к уравнениям

$$\begin{aligned} u - \lambda(au + bv) &= 0, \\ v - \lambda(bu + cv) &= 0. \end{aligned} \quad (2)$$

Это дает

$$\begin{vmatrix} 1 - a\lambda & -b\lambda \\ -b\lambda & 1 - c\lambda \end{vmatrix} = 0 \quad (3)$$

или в эквивалентной форме

$$\begin{vmatrix} a - 1/\lambda & b \\ b & c - 1/\lambda \end{vmatrix} = 0. \quad (4)$$

Если λ_i — корень уравнения (4) и (u_i, v_i) — соответствующая точка экстремума, то мы, как и ранее, можем заключить, что

$$u_i^2 + v_i^2 = \lambda_i. \quad (5)$$

Из этого равенства следует вывод, что один из корней уравнения (4) равен квадрату минимального расстояния от начала

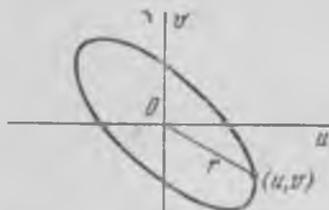


Рис. 1.

координат до эллипса, а другой — квадрату максимального расстояния. Отметим также, что сформулированная нами вариационная задача дает в ходе решения длины большой и малой полуосей эллипса. В условии (6.8) мы узнаем теперь хорошо известное свойство перпендикулярности или ортогональности главных осей эллипса.

Линейное преобразование (7.1), очевидно, есть вращение, так как сохраняет как положение начала координат, так и расстояние до него. Мы видим, что это как раз то вращение, которое приводит к совмещению осей эллипса с осями координат.

Упражнения

1. Используя вышеприведенные факты, показать, что площадь эллипса дается формулой

$$\frac{\pi}{(ac - b^2)^{1/2}}$$

2. Используя алгебраический подход (см. § 4), показать, что необходимые и достаточные условия положительной определенности формы $Q = au^2 + 2buv + cv^2 + 2duw + ew^2 + 2fuv$ имеют вид

$$a > 0, \quad \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0, \quad \begin{vmatrix} a & b & d \\ b & c & f \\ d & f & e \end{vmatrix} > 0.$$

3. Аналогично, используя аналитический подход (см. § 5), показать, что необходимым и достаточным условием положительной определенности Q является положительность всех корней уравнения

$$\begin{vmatrix} a-\lambda & b & d \\ b & c-\lambda & f \\ d & f & e-\lambda \end{vmatrix} = 0.$$

4. Показать, что если Q — положительно определенная форма, то уравнению $Q(u, v, w) = 1$ соответствует эллипсоид. Определить объем этого эллипсоида.

§ 11. Обсуждение. Мы исследовали в деталях двумерный случай. Соответствующие вопросы элементарны, однако истоки некоторых из них не вполне очевидны. Все рассмотрение проведено для того, чтобы читателю было легче понять необходимость перехода к новым обозначениям и он смог оценить их преимущества.

Результаты, которые для двумерного случая легко можно было предвидеть и которые были проверены непосредственно выкладками, будут естественным образом получены в общем случае. Основные идеи и основные приемы, с которыми мы будем иметь дело, содержатся в уже проведенных рассуждениях.

Упражнения к гл. I

1. При каких значениях x_1 и x_2 квадратичная форма

$$(a_{11}x_1 + a_{12}x_2 - b_1)^2 + (a_{21}x_1 + a_{22}x_2 - b_2)^2$$

принимает минимальное значение и чему оно равно?

2. Показать, что выражение $(x_1^2 + y_1^2)(x_2^2 + y_2^2)$ может быть записано в форме

$$(a_1x_1x_2 + a_2x_1y_2 + a_3x_2y_1 + a_4y_1y_2)^2 + (b_1x_1x_2 + b_2x_1y_2 + b_3x_2y_1 + b_4y_1y_2)^2,$$

где коэффициенты a_i и b_i не зависят от x_i и y_i , и определить все такие коэффициенты.

3. Показать, что аналогичный результат не имеет места для выражения

$$(x_1^2 + y_1^2 + z_1^2)(x_2^2 + y_2^2 + z_2^2).$$

4. Если $x_1u^2 + 2x_2uv + x_3v^2$ и $y_1u^2 + 2y_2uv + y_3v^2$ — положительно определенные формы, то

$$\begin{vmatrix} x_1y_1 & x_2y_2 \\ x_2y_2 & x_3y_3 \end{vmatrix} \geq \begin{vmatrix} x_1 & x_2 \\ x_2 & x_3 \end{vmatrix} \begin{vmatrix} y_1 & y_2 \\ y_2 & y_3 \end{vmatrix}.$$

5. Использовать этот результат для трактовки упражнения 1 § 9.

6. Обосновать справедливость формального введения множителя Лагранжа при рассмотрении максимальных и минимальных значений функции

$$(au^2 + 2buv + cv^2)/(u^2 + v^2).$$

7. Какое линейное преобразование оставляет квадратичную форму

$$Q(x_1, x_2) = \lambda(x_1^2 + x_2^2) + (1 - \lambda)(x_1 - x_2)^2$$

инвариантной? Здесь $0 < \lambda < 1$.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 1. Общее рассмотрение проблемы максимума и минимума функции нескольких переменных можно найти в любой из книг по математическому анализу. Более полное обсуждение может быть найдено в книге

Хенкок (H. Hancock), *The Theory of Maxima and Minima*, Ginn & Company, Boston, 1907.

В последней главе мы будем изучать более сложную задачу определения максимума и минимума при ограничениях. Проблема определения связи между числом относительных максимумов, относительных минимумов и стационарных точек других типов принадлежит области топологии и здесь обсуждаться не будет; см.

Морс (M. Morse), *The Calculus of Variations in the Large*, Amer. Math. Soc. 18 (1934).

§ 2. Часто пользуются другой терминологией: Швердфегер и Мирский используют термин *положительно полуопределенная*; Халмош — *неотрицательно полуопределенная*.

ВЕКТОРЫ И МАТРИЦЫ

§ 1. Введение. В первой главе мы исследовали вопрос определения локального максимума и минимума функции двух переменных. Если перейти к рассмотрению соответствующей проблемы для функции N переменных и поступать так же, как и ранее, то мы сразу же придем к задаче определения простых необходимых и достаточных условий положительности квадратичной формы N переменных

$$Q(x_1, x_2, \dots, x_N) = \sum_{i,j=1}^N a_{ij}x_ix_j. \quad (1)$$

Как станет ясно позднее, в гл. 5, алгебраический метод, изложенный в предыдущей главе, дает простое и изящное решение этой проблемы. Однако, так как мы заинтересованы в более глубоком понимании квадратичных форм, чем это требуется для данной частной задачи, будем предполагать, что такого решения не существует.

В настоящей главе нашей целью будет введение такой системы обозначений, которая даст нам возможность развить аналитический подход при минимуме аналитических трудностей. В соответствии с указанной целью будем стремиться ввести обозначения, не зависящие от размерности.

Несколько неожиданным оказывается то обстоятельство, что исследование квадратичных форм существенно упрощается при использовании обозначений, введенных первоначально для изучения линейных преобразований вида

$$y_i = \sum_{j=1}^N a_{ij}x_j, \quad i = 1, 2, \dots, N. \quad (2)$$

§ 2. Векторы. Определим *вектор* как набор N комплексных чисел. Вектор вида

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_N \end{pmatrix} \quad (1)$$

будем называть *вектором-столбцом*. Если N чисел расположены горизонтально:

$$\mathbf{x} = (x_1, x_2, \dots, x_N), \quad (2)$$

то \mathbf{x} называется *вектором-строкой*.

Так как все операции, которые нам потребуются, можно проводить, пользуясь лишь векторами-столбцами, в дальнейшем будем применять термин «вектор» для величины, заданной формулой (1).

Строчные буквы, такие как \mathbf{x} , \mathbf{y} , \mathbf{z} и \mathbf{a} , \mathbf{b} , \mathbf{c} , повсюду будем применять для обозначения векторов. Когда рассматривается некоторое множество векторов, мы будем использовать верхние индексы: x^1 , x^2 и т. д. Величины x_i называются *компонентами* вектора \mathbf{x} , в то время как число N называется *размерностью* вектора \mathbf{x} . Одномерные векторы будем называть *скалярами*. Это обычные величины, встречающиеся в анализе.

Через $\bar{\mathbf{x}}$ обозначается вектор, компоненты которого комплексно сопряжены элементам вектора \mathbf{x} . Если все компоненты \mathbf{x} вещественны, то будем говорить, что \mathbf{x} — вещественный вектор.

§ 3. Сложение векторов. Приступим к изложению алгебры векторов, устанавливающей правила обращения с этими величинами. Может быть предложено сколь угодно много различных правил, определяющих операции с векторными величинами. Те из них, которые приняты нами, найдут оправдание в том, что они позволяют просто и изящно изложить некоторые важные вопросы.

Говорят, что два вектора \mathbf{x} и \mathbf{y} *равны*, в том и только в том случае, если равны их компоненты: $x_i = y_i$ для $i = 1, 2, \dots, N$. Простейшая операция для двух векторов есть *сложение*. Сумма двух векторов \mathbf{x} и \mathbf{y} записывается как $\mathbf{x} + \mathbf{y}$ и определяется вектором

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_N + y_N \end{pmatrix}. \quad (1)$$

Отметим, что знак плюс, связывающий векторы \mathbf{x} и \mathbf{y} , по смыслу отличен от знака плюс, соединяющего компоненты x_i и y_i . Вместе с тем он обладает теми же аналитическими свойствами, и потому нет ничего плохого в использовании одного и того же символа в обоих случаях.

Упражнения

1. Показать, что операция сложения векторов обладает свойством коммутативности: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ — и ассоциативности: $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$.

2. Показать, что вследствие этого вектор $\mathbf{x}^1 + \mathbf{x}^2 + \dots + \mathbf{x}^M$ определен однозначно.

3. Определить разность двух векторов непосредственно и в терминах операции сложения, т. е. $\mathbf{x} - \mathbf{y}$ есть вектор \mathbf{z} такой, что $\mathbf{y} + \mathbf{z} = \mathbf{x}$.

§ 4. Умножение вектора на скаляр. Умножение вектора \mathbf{x} на скаляр c_1 определяется соотношением

$$c_1 \mathbf{x} = \mathbf{x} c_1 = \begin{pmatrix} c_1 x_1 \\ c_1 x_2 \\ \vdots \\ c_1 x_N \end{pmatrix}. \quad (1)$$

Упражнения

1. Показать, что $(c_1 + c_2)(\mathbf{x} + \mathbf{y}) = c_1 \mathbf{x} + c_1 \mathbf{y} + c_2 \mathbf{x} + c_2 \mathbf{y}$.

2. Определим нулевой вектор (с обозначением $\mathbf{0}$) как вектор, все компоненты которого равны нулю. Показать, что это эквивалентно определению вектора $\mathbf{0}$ как вектора, для которого равенство $\mathbf{x} + \mathbf{0} = \mathbf{x}$ справедливо для всех \mathbf{x} .

3. Показать, что эквивалентное определение нулевого вектора дается условием $c_1 \mathbf{0} = \mathbf{0}$ для всех скаляров c_1 .

§ 5. Скалярное произведение двух векторов. Введем теперь важную скалярную функцию двух векторов \mathbf{x} и \mathbf{y} , которую будем называть *скалярным произведением*. Эта функция будет записываться как (\mathbf{x}, \mathbf{y}) и определяться соотношением

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^N x_i y_i. \quad (1)$$

Следующие свойства скалярного произведения выводятся непосредственно из определения:

$$(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x}), \quad (2a)$$

$$(\mathbf{x} + \mathbf{y}, \mathbf{z} + \mathbf{w}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{x}, \mathbf{w}) + (\mathbf{y}, \mathbf{z}) + (\mathbf{y}, \mathbf{w}), \quad (2b)$$

$$(c_1 \mathbf{x}, \mathbf{y}) = c_1 (\mathbf{x}, \mathbf{y}). \quad (2в)$$

Скалярное произведение — это лишь один из способов «умножения» двух векторов. Вместе с тем имеются и другие определения операции умножения, которые мы не будем использовать в настоящей книге.

Важность введенного понятия скалярного произведения связана с тем, что величину (\mathbf{x}, \mathbf{x}) можно рассматривать как квадрат «длины» вещественного вектора \mathbf{x} .

Упражнения

1. Показать, что если x — вещественный вектор, то $(x, x) > 0$ для всех x , кроме $x=0$.

2. Показать, что $(ux+vy, ux+vy) = u^2(x, x) + 2uv(x, y) + v^2(y, y)$ есть неотрицательная квадратичная форма скалярных переменных u и v , если x и y — вещественные векторы, и что, следовательно, для любых двух вещественных векторов x и y справедливо неравенство $(x, y)^2 \leq (x, x)(y, y)$ (неравенство Коши)

3. Дать геометрическую интерпретацию этого результата.

4. Используя этот результат, показать, что для любых двух вещественных векторов

$$(x+y, x+y)^{1/2} \leq (x, x)^{1/2} + (y, y)^{1/2}.$$

5. Почему последнее неравенство называется «неравенством треугольника»?

6. Показать, что для любых двух комплексных векторов x и y справедливо неравенство

$$|(x, y)|^2 \leq (x, \bar{x})(y, \bar{y}).$$

§ 6. Ортогональность. Два вещественных вектора называются *ортогональными*, если они удовлетворяют соотношению

$$(x, y) = 0. \quad (1)$$

Важность понятия ортогональности в существенной мере становится понятной после рассмотрения следующего легко устанавливаемого результата.

Теорема 1. Пусть $\{x^i\}$ — множество вещественных взаимно ортогональных векторов таких, что $(x^i, x^j) = 0$ при $i \neq j$, и пусть эти векторы нормированы условием

$$(x^i, x^i) = 1.$$

Тогда если данный вектор x представлен суммой

$$x = \sum_{i=1}^M c_i x^i, \quad (2)$$

то имеют место следующие выражения для коэффициентов c_i :

$$c_i = (x, x^i). \quad (3)$$

Кроме того, справедливо соотношение

$$(x, x) = \sum_{i=1}^M c_i^2. \quad (4)$$

Упражнения

1. В силу каких причин понятие ортогональности для комплексных векторов вводится соотношением

$$(x, \bar{y}) = 0$$

и соответствующее условие нормирования — соотношением $(x, \bar{x}) = 1$?

2. Показать, что $(x, \bar{y}) = \overline{(x, y)}$.

3. В N -мерном евклидовом пространстве заданы векторы

$$y^1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad y^2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad y^N = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Показать, что векторы y^i взаимно ортогональны и нормированы. Векторы такого типа называются *ортонормированными*. Пусть $x = \sum_{i=1}^N c_i y^i$, определить величины c_i и обсудить геометрический смысл результата.

§ 7. Матрицы. Введем понятие *матрицы*. Таблицу комплексных чисел, записанную в форме

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{pmatrix}, \quad (1)$$

будем называть *квадратной матрицей*. Это только один из возможных видов матриц. Мы будем рассматривать таблицы подобного типа не обязательно квадратные и применять для них тот же термин «матрица». Впоследствии, когда будут вводиться другие виды матриц, мы будем пользоваться термином «матрица» с добавлением соответствующего прилагательного.

Величины a_{ij} называются *элементами*, а целое число N — *размером (порядком)* матрицы A .

Величины $a_{i1}, a_{i2}, \dots, a_{iN}$ составляют i -ю строку матрицы A , а величины $a_{1j}, a_{2j}, \dots, a_{Nj}$ — ее j -й столбец.

Повсюду будем обозначать матрицы прописными буквами X, Y, Z и A, B, C . Иногда будем пользоваться сокращенным обозначением

$$A = \|a_{ij}\|. \quad (2)$$

Определитель, соответствующий таблице (1), будем записывать как $|A|$ или $|a_{ij}|$.

Простейшее соотношение между матрицами — это их равенство. Две матрицы называются *равными* тогда и только тогда, когда все их соответствующие элементы равны. Действуя так же, как и при рассмотрении векторов, определим операцию сложения. Сумма двух матриц A и B записывается как $A+B$ и определяется формулой

$$A+B = \|a_{ij} + b_{ij}\|. \quad (3)$$

Произведение матрицы A на скаляр c_1 определяется соотношением

$$c_1 A = A c_1 = \|c_1 a_{ij}\|. \quad (4)$$

Наконец, под \bar{A} понимается матрица, элементы которой комплексно сопряжены с элементами матрицы A . Будем называть матрицу A вещественной, если все ее элементы вещественные.

Упражнения

1. Показать, что нулевая матрица O (матрица, все элементы которой равны нулю) однозначно определена условием $A + O = A$ для всех A .
2. Показать, что сложение матриц ассоциативно и коммутативно.
3. Показать, что матрица $A_1 + A_2 + \dots + A_N$ определена однозначно.

§ 8. Умножение вектора на матрицу. Для того чтобы сделать нашу алгебру более содержательной, определим некоторые операции умножения.

Для обоснования новых понятий вернемся к линейному преобразованию

$$y_i = \sum_{j=1}^N a_{ij} x_j, \quad i = 1, 2, \dots, N, \quad (1)$$

где коэффициенты a_{ij} — комплексные величины. Целью векторно-матричных обозначений является построение аппарата для исследования линейных преобразований, поэтому вполне понятно, что при выводе фундаментальных свойств векторов и матриц мы периодически обращаемся к этому основному соотношению. Дело в том, что мы хотим, чтобы операции сложения и умножения матриц были введены не произвольно, а наоборот — существенно определялись аналитическими и геометрическими свойствами объектов нашего исследования.

Если два вектора x и y связаны соотношением (1), то можно записать

$$y = Ax. \quad (2)$$

Это соотношение определяет произведение вектора на матрицу. Обратим внимание на порядок, в котором записаны сомножители этого произведения.

Упражнения

1. Показать, что $(A + B)(x + y) = Ax + Ay + Bx + By$.
2. Рассмотреть единичную матрицу I , определяемую таблицей

$$I = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & & & & 0 \end{pmatrix}$$

Для элементов единичной матрицы имеем $I = \|\delta_{ij}\|$, где δ_{ij} — символ Кронекера:

$$\delta_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

Показать, что

$$\delta_{ij} = \sum_{k=1}^N \delta_{ik} \delta_{kj}.$$

3. Показать, что $Ix = x$ для всех x и что это соотношение единственным образом определяет матрицу I .

4. Показать, что

$$(Ax, Ax) = \sum_{i=1}^N \left(\sum_{j=1}^N a_{ij} x_j \right)^2.$$

§ 9. Умножение матрицы на матрицу. Теперь рассмотрим, каким образом можно определить умножение матрицы на матрицу. Рассмотрим второе линейное преобразование

$$z = By, \quad (1)$$

которое устанавливает соответствие между компонентами векторов y и z . Для того чтобы выразить компоненты z через компоненты вектора x , где, как указывалось выше, $y = Ax$, запишем

$$z_i = \sum_{k=1}^N b_{ik} y_k = \sum_{k=1}^N b_{ik} \left(\sum_{j=1}^N a_{kj} x_j \right) = \sum_{j=1}^N \left(\sum_{k=1}^N b_{ik} a_{kj} \right) x_j. \quad (2)$$

Если теперь ввести новую матрицу $C = \|c_{ij}\|$, определяемую соотношениями

$$c_{ij} = \sum_{k=1}^N b_{ik} a_{kj}, \quad i, j = 1, 2, \dots, N, \quad (3)$$

то можно записать

$$z = Cx. \quad (4)$$

Так как формально

$$z = By = B(Ax) = (BA)x, \quad (5)$$

мы приходим к определению произведения матриц A и B :

$$C = BA, \quad (6)$$

где матрица C дается формулами (3).

Снова обратим внимание на порядок, в котором записаны сомножители матричного произведения.

Упражнения

1. Показать, что $(A + B)(C + D) = AC + AD + BC + BD$.

2. Пусть

$$T(\theta) = \begin{vmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{vmatrix}.$$

Показать, что

$$T(\theta_1)T(\theta_2) = T(\theta_2)T(\theta_1) = T(\theta_1 + \theta_2).$$

8. Пусть A — диагональная матрица (т. е. такая, что $a_{ij} = 0$, если $i \neq j$), и пусть B — матрица подобного же типа. Показать, что AB — также диагональная матрица и что $AB = BA$.

4. Пусть матрица A — треугольная (т. е. такая, что $a_{ij} = 0$ для $j > i$), и пусть B — матрица того же типа. Показать, что AB — снова треугольная матрица, но что в общем случае $AB \neq BA$.

8. Пусть

$$A = \begin{vmatrix} a_1 & a_2 \\ -a_2 & a_1 \end{vmatrix}, \quad B = \begin{vmatrix} b_1 & b_2 \\ -b_2 & b_1 \end{vmatrix}.$$

Показать, что AB — матрица того же типа и что $AB = BA$. (Как мы увидим впоследствии, эти матрицы эквивалентны комплексным числам вида $a_1 + ia_2$, если a_1 и a_2 вещественны.)

6. Используя соотношение $|AB| = |A||B|$, показать, что

$$(a_1^2 + a_2^2)(b_1^2 + b_2^2) = (a_1b_1 - a_2b_2)^2 + (a_2b_1 + a_1b_2)^2.$$

7. Пусть

$$A = \begin{vmatrix} a_1 + ia_2 & a_3 + ia_4 \\ -a_2 + ia_1 & a_1 - ia_2 \end{vmatrix}$$

и B — матрица аналогичного вида. Показать, что AB — матрица того же типа, но что, вообще говоря, $AB \neq BA$.

8. Пользуясь соотношением $|AB| = |A||B|$, представить выражение $(a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2)$ как сумму четырех квадратов.

9. Пусть

$$A = \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ -a_2 & a_1 & -a_4 & a_3 \\ -a_3 & a_4 & a_1 & -a_2 \\ -a_4 & -a_3 & a_2 & a_1 \end{vmatrix}$$

и B — матрица подобного вида. Показать, что AB — матрица того же типа, но что в общем случае $AB \neq BA$. Вычислить $|A|$. (Эти матрицы эквивалентны кватернионам.)

10. Рассмотреть дробно-линейное преобразование

$$w = \frac{a_1z + b_1}{c_1z + d_1} = T_1(z).$$

Показать, что если $T_2(z)$ — такое же преобразование с коэффициентами a_1, b_1, c_1, d_1 , замененными на a_2, b_2, c_2, d_2 , то $T_1(T_2(z))$ и $T_2(T_1(z))$ — преобразования того же типа.

11. Показать, что если выражение $a_1d_1 - b_1c_1 \neq 0$, то $T_1^{-1}(z)$ — преобразование того же вида. Каков вид коэффициентов преобразования $T_1^{-1}(z)$, если $a_1d_1 - b_1c_1 = 0$?

12. Рассмотреть соответствие между преобразованием $T_1(z)$ и матрицей

$$A_1 = \begin{vmatrix} a_1 & b_1 \\ c_1 & d_1 \end{vmatrix}.$$

записываемое как $A_1 \sim T_1(z)$. Показать, что если $A_1 \sim T_1(z)$ и $A_2 \sim T_2(z)$, то $A_1A_2 \sim T_1(T_2(z))$. Найти условие, достаточное для выполнения равенства

$$T_1(T_2(z)) = T_2(T_1(z)) \text{ для всех } z.$$

13. Как, пользуясь приведенными результатами, получить представление для итераций преобразования $T_1(z)$?

14. Показать, что

$$\begin{vmatrix} -1 & -1 \\ 1 & 0 \end{vmatrix}^3 = I.$$

15. Если $X = \|x_{ij}\|$, где $x_{ij} = (-1)^{N-j} \binom{N-j}{i-1}$, то $X^2 = I$.

16. Из того обстоятельства, что мы можем установить соответствия

$$I \sim \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \quad \text{и} \quad -I \sim \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix},$$

вывести, что может быть построена арифметика, не использующая понятия отрицательного числа.

§ 10. Некоммутативность. Одним из интереснейших свойств матриц, делающим их изучение таким увлекательным, является то, что произведение матриц некоммутативно. Иными словами, в общем случае

$$AB \neq BA. \quad (1)$$

Простой пример дается матрицами второго порядка

$$A = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}, \quad B = \begin{vmatrix} 2 & 1 \\ 4 & 3 \end{vmatrix}. \quad (2)$$

Более интересные примеры появляются в упражнениях 7 и 9 § 9. Для матриц (2) имеем

$$AB = \begin{vmatrix} 10 & 7 \\ 22 & 15 \end{vmatrix}, \quad BA = \begin{vmatrix} 5 & 8 \\ 13 & 20 \end{vmatrix}. \quad (3)$$

Если $AB = BA$, то мы будем говорить, что матрицы A и B коммутативны.

Теория матриц дает естественный переход от обычной области скаляров и их доступной алгебры к более интересному реальному миру объектов, в котором имеется множество различных видов алгебр, обладающих своими специфическими свойствами.

Упражнения

1. Показать, что $AI = IA = A$ для всех матриц A и что это соотношение единственным образом определяет единичную матрицу I .

2. Показать, что $A0 = 0A = 0$ для всех матриц A и что это соотношение единственным образом определяет нулевую матрицу 0 .

3. Пусть строки матрицы A состоят из компонент векторов a^1, a^2, \dots, a^N , а столбцы матрицы B состоят из компонент векторов b^1, b^2, \dots, b^N . Тогда

$$*) \binom{\alpha}{k} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}, \quad \text{причем} \quad \binom{\alpha}{0} = \binom{0}{0} = 1. \quad (\text{Прим. ред.})$$

для произведения матриц имеем

$$AB = \|(a', b')\|.$$

4. Если $AX = XA$ для всех X , то матрица A представляет собой произведение скаляра и единичной матрицы I .

§ 11. Ассоциативность. Хотя в новой алгебре коммутативность не имеет места, свойство ассоциативности, к счастью, сохраняется. Иными словами, для любых матриц A , B и C мы имеем

$$(AB)C = A(BC). \quad (1)$$

Это означает, что произведение ABC вполне определено и без помощи круглых скобок. Для доказательства этого результата воспользуемся правилом «немного индекса», которое заключается в том, что любой повторяющийся индекс должен суммироваться по всем своим возможным значениям. Приняв это условие, i -й элемент произведения AB можно записать как

$$a_{ik}b_{kj}. \quad (2)$$

Применяя этот прием и определяя произведение, приведенное выше, мы получаем соотношения

$$\begin{aligned} (AB)C &= \|(a_{ik}b_{ki})c_{ij}\|, \\ A(BC) &= \|a_{ik}(b_{kj}c_{ij})\|, \end{aligned} \quad (3)$$

которые устанавливают равенство выражений $(AB)C$ и $A(BC)$.

Упражнения

1. Показать, что $ABCD = A(BCD) = (AB)(CD) = (ABC)D$ и что в общем случае произведение $A_1A_2 \dots A_n$ определяется однозначно.

2. Показать, что функция A^n определяется однозначно и что $A^{m+n} = A^m A^n$; $m, n = 1, 2, \dots$

3. Показать, что A^m и B^n коммутативны, если коммутативны A и B .

4. Получить выражение для

$$X^n = \begin{vmatrix} x_1(n) & x_2(n) \\ x_3(n) & x_4(n) \end{vmatrix}, \quad \text{где } X = \begin{vmatrix} x_1 & x_2 \\ x_3 & x_4 \end{vmatrix}$$

— заданная матрица. Используя соотношение $X^{n+1} = XX^n$, вывести рекуррентную формулу для $x_i(n)$ и таким образом получить $x_i(n)$ в аналитической форме.

5. Использовать эти соотношения для случая, когда

$$X = \begin{vmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{vmatrix}, \quad x_1, x_2 \text{ вещественны,}$$

$$X = \begin{vmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{vmatrix}, \quad x_1, x_2 \text{ комплексные.}$$

6. Используя эти соотношения, определить явные выражения для элементов X^n , когда

$$X = \begin{vmatrix} \lambda & 1 \\ 0 & \lambda \end{vmatrix}, \quad X = \begin{vmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{vmatrix}.$$

7. Найти все матрицы размеров 2×2 , удовлетворяющие уравнению $X^2 = X$.

8. Найти все матрицы размеров 2×2 , удовлетворяющие уравнению $X^n = X$, где n — целое положительное число.

§ 12. **Инвариантные векторы.** Если действовать так же, как в § 5 гл. 1, то нетрудно убедиться, что проблема определения максимального и минимального значений квадратичной формы

$Q = \sum_{i,j=1}^N a_{ij}x_i x_j$ для значений x_i , удовлетворяющих условию

$\sum_{i=1}^N x_i^2 = 1$, может быть сведена к проблеме определения величин λ , для которых система линейных однородных уравнений

$$\sum_{j=1}^N a_{ij}x_j = \lambda x_i, \quad i = 1, 2, \dots, N, \quad (1)$$

имеет нетривиальные решения.

В векторно-матричных обозначениях соотношения (1) можно представить в форме одного уравнения

$$Ax = \lambda x. \quad (2)$$

Записанное таким образом уравнение имеет очень простой смысл. Мы ищем такие векторы x , которые преобразуются матрицей A в векторы, отличающиеся от исходных лишь скалярным множителем λ .

Если считать, что компоненты вектора x определяют некоторое направление в N -мерном пространстве, то нашей целью является определение инвариантного направления.

Мы продолжим исследование возникающих при этом вопросов в последующих главах. Тем временем введем некоторые новые обозначения, которые будут полезны нам в дальнейшем.

§ 13. **Квадратичная форма как скалярное произведение.** Приведем еще один аргумент, обосновывающий введенные нами обозначения. Квадратичная форма $Q(u, v) = au^2 + 2buv + cv^2$ может быть записана в форме

$$u(au + bv) + v(bu + cv). \quad (1)$$

Следовательно, если вектор x и матрицу A определить формулами

$$x = \begin{vmatrix} u \\ v \end{vmatrix}, \quad A = \begin{vmatrix} a & b \\ b & c \end{vmatrix}, \quad (2)$$

то

$$Q(u, v) = (x, Ax). \quad (8)$$

Аналогично для N -мерной квадратичной формы

$$Q(x) = \sum_{i,j=1}^N a_{ij} x_i x_j, \quad (4)$$

где без потери общности можно положить $a_{ij} = a_{ji}^*$), запишем

$$Q(x) = x_1 \left[\sum_{j=1}^N a_{1j} x_j \right] + x_2 \left[\sum_{j=1}^N a_{2j} x_j \right] + \dots + x_N \left[\sum_{j=1}^N a_{Nj} x_j \right] = (x, Ax). \quad (5)$$

Здесь вектор x имеет компоненты x_i и $A = \|a_{ij}\|$.

Упражнения

1. Означает ли выполнение равенства $(x, Ax) = (x, Bx)$ для всех x , что $A = B$?

2. При каких условиях $(x, Ax) = 0$ для всех x ?

§ 14. Транспонированная матрица. Соотношение

$$A' = \|a_{ji}\| \quad (1)$$

является определением весьма важной функции матрицы A — *транспонированной матрицы* A' . Строки матрицы A' являются столбцами матрицы A и строки A — столбцами A' .

К рассмотрению этой новой матрицы приводит следующее обстоятельство. Пусть заданы два множества векторов $\{x\}$ и $\{y\}$. образуем всевозможные скалярные произведения (x, y) , где вектор x берется из множества $\{x\}$, а вектор y — из множества $\{y\}$.

Предположим, что множество $\{x\}$ преобразуется посредством матричного умножения на матрицу A и дает новое множество векторов $\{Ax\}$. образуя скалярные произведения с векторами y , мы получаем множество значений (Ax, y) .

Заметим, что

$$(Ax, y) = y_1 \left[\sum_{j=1}^N a_{1j} x_j \right] + y_2 \left[\sum_{j=1}^N a_{2j} x_j \right] + \dots + y_N \left[\sum_{j=1}^N a_{Nj} x_j \right]. \quad (2)$$

После перегруппировки получаем

$$(Ax, y) = x_1 \left[\sum_{i=1}^N a_{i1} y_i \right] + x_2 \left[\sum_{i=1}^N a_{i2} y_i \right] + \dots + x_N \left[\sum_{i=1}^N a_{iN} y_i \right] = (x, A'y). \quad (3)$$

* См. упражнение 7 к § 15 настоящей главы. (Прим. ред.)

Другими словами, когда речь идет о скалярном произведении, действие преобразования A на множество $\{x\}$ эквивалентно действию преобразования A' на множество $\{y\}$. В таком случае мы можем рассматривать A' как *индуцированное* или *сопряженное преобразование*. Эта простая, но плодотворная идея глубоко проникла в классический и современный анализ.

§ 15. Симметрические матрицы. Проведенное рассмотрение позволяет предположить, что матрицы, удовлетворяющие условию

$$A = A', \quad (1)$$

обладают особыми свойствами и могут играть важную роль в изучении квадратичных форм. На самом деле это действительно так. Матрицы такого типа называются *симметрическими* и характеризуются условием

$$a_{ij} = a_{ji}. \quad (2)$$

Первая часть настоящей книги будет посвящена исследованию основных свойств матриц этого класса для случая, когда элементы a_{ij} вещественны.

В дальнейшем термин «симметрическая» будет использоваться для обозначения вещественной симметрической матрицы. Если при этом возникает опасность путаницы, то мы будем добавлять слова «вещественная» или «комплексная симметрическая» в зависимости от типа рассматриваемой матрицы.

Упражнения

1. Показать, что $(A')' = A$.
2. Показать, что $(A + B)' = A' + B'$, $(AB)' = B'A'$,

$$(A_1 A_2 \dots A_n)' = A_n' \dots A_2' A_1', \quad (A^n)' = (A')^n.$$

3. Показать, что если матрицы A и B симметрические, то матрица AB не обязательно симметрическая.

4. Показать, что если B — симметрическая матрица, то матрица $A'BA$ симметрическая.

5. Показать, что $(Ax, By) = (x, A'By)$.

6. Показать, что $|A| = |A'|$.

7. Показать, что предположение $a_{ij} = a_{ji}$ не приводит к потере общности при рассмотрении квадратичной формы $Q(x) = \sum_{i,j=1}^n a_{ij}x_i x_j$.

§ 16. Эрмитовы матрицы. Как было отмечено выше, важной скалярной функцией комплексных векторов оказывается не обычное скалярное произведение, а выражение вида (x, \bar{y}) .

Если заметить, что

$$(Ax, \bar{y}) = (x, \bar{z}). \quad (1)$$

где $z = \overline{A'}y$, то мы видим, что сопряженное преобразование определяется теперь матрицей $\overline{A'}$, комплексно сопряженной с матрицей A' . Матрицы, для которых

$$A = \overline{A'}, \quad (2)$$

называются *эрмитовыми* по имени великого французского математика Шарля Эрмита.

Для упрощения обозначений вместо $\overline{A'}$ будем писать A^* .

Как мы увидим, все важнейшие свойства симметрических матриц имеют прямые аналогии для эрмитовых матриц. Более того, при желании можно ввести обозначение

$$[x, y] = (x, \overline{y}), \quad (3)$$

в терминах которого свойства обоих типов матриц могут быть получены одновременно. Каждый из этих путей имеет свои преимущества, и читатель сможет сделать свой выбор, как только он освоит основную технику.

Упражнения

1. Вещественная эрмитова матрица — симметрическая.
2. $(A^*)^* = A$, $(AB)^* = B^*A^*$, $(A_1A_2 \dots A_n)^* = A_n^* \dots A_2^*A_1^*$.
3. Если матрица $A + iB$ эрмитова (A и B вещественны), то $A' = A$, $B' = -B$.

§ 17. Ортогональные матрицы. Инвариантность расстояний. Отправляясь от измерения длин в геометрии Евклида, примем величину (x, x) за длину вещественного вектора x .

Чрезвычайно любопытным и в то же время чрезвычайно важным, как мы увидим далее, является вопрос определения такого линейного преобразования $y = Tx$, которое оставляет неизменной величину (x, x) . Иными словами, мы хотим определить матрицу T так, чтобы уравнение

$$(x, x) = (Tx, Tx) \quad (1)$$

удовлетворялось для всех x . Так как

$$(Tx, Tx) = (x, T'Tx) \quad (2)$$

и $T'T$ — симметрическая матрица, то легко прийти к заключению, что условие (1) приводит к равенству

$$T'T = I. \quad (3)$$

Вещественная матрица T , обладающая таким свойством, называется *ортогональной*.

Упражнения

1. Показать, что матрица T' ортогональна тогда и только тогда, когда ортогональна матрица T .

2. Показать, что любая ортогональная матрица размеров 2×2 с определителем, равным ± 1 , может быть записана в форме

$$\begin{vmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{vmatrix}.$$

Каков геометрический смысл этого результата?

3. Показать, что столбцы матрицы T представляют собой ортогональные векторы.

4. Показать, что произведение двух ортогональных матриц есть снова ортогональная матрица.

5. Показать, что определитель ортогональной матрицы равен ± 1 .

6. Пусть T_N — ортогональная матрица размеров $N \times N$, а матрица размеров $(N+1) \times (N+1)$ имеет вид

$$T_{N+1} = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & T_N & \\ \vdots & & & \\ 0 & & & \end{vmatrix}.$$

Показать, что матрица T_{N+1} — ортогональная.

7. Показать, что если матрица T — ортогональная, то преобразование $x = Ty$ означает, что $y = T'x$.

8. Пусть $AB = BA$ и T — ортогональная матрица, тогда матрицы TAT' и TBT' коммутативны.

§ 18. Унитарные матрицы. Так как соответствующей мерой комплексных векторов является величина (x, \bar{x}) , то легко заключить, что аналогом условия инвариантности (17.3) является соотношение

$$T^*T = I. \quad (1)$$

Матрицы, обладающие этим свойством, называются *унитарными* и играют при изучении эрмитовых матриц ту же роль, что и ортогональные в теории симметрических матриц.

Упражнения

1. Показать, что матрица T^* унитарна тогда и только тогда, когда унитарна матрица T .

2. Показать, что произведение двух унитарных матриц есть унитарная матрица.

3. Показать, что определитель унитарной матрицы по модулю равен единице.

4. Показать, что если матрица T унитарна, то равенство $x = Ty$ означает, что $y = T^*x$.

5. Получить результат, соответствующий приведенному в упражнении 2 § 17, для элементов унитарной матрицы размеров 2×2 . (Аналогичный, но более сложный результат в эллиптических функциях имеет место для представления элементов ортогональной матрицы размеров 3×3 . См. К а с п а р и (F. Caspary), Zur Theorie der Thetafunktionen mit zwei argumenten, Kroecker J. 44, 74—86; К а с п а р и (F. Caspary), Sur les systèmes orthogonaux, formés par les fonctions theta, Comptes Rendus de Paris 54, 490—493.

6. Будет ли вещественная унитарная матрица ортогональной?

7. Будет ли комплексная ортогональная матрица унитарной?

8. Любая ортогональная матрица размеров 3×3 может быть представлена как произведение

$$\begin{pmatrix} \cos a & \sin a & 0 \\ -\sin a & \cos a & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos b & 0 & -\sin b \\ 0 & 1 & 0 \\ \sin b & 0 & \cos b \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos c & \sin c \\ 0 & -\sin c & \cos c \end{pmatrix}$$

Каков геометрический смысл этого результата?

Упражнения к гл. 2

1. Доказать, что

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1N} & x_1 \\ a_{21} & a_{22} & \dots & a_{2N} & x_2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} & x_N \\ x_1 & x_2 & \dots & x_N & 0 \end{vmatrix} = -(a^{11}x_1^2 + \dots + (a^{jj} + a^{jj})x_jx_j + \dots),$$

где a^{jj} — алгебраическое дополнение элемента a_{jj} в определителе $|a_{ij}|$.

2. Пусть l_{ij} обозначает матрицу, полученную перестановкой i -й и j -й строк в единичной матрице I . Доказать, что

$$l_{ij}^2 = I, \quad l_{ik}l_{kj}l_{ji} = l_{kj}.$$

3. Показать, что матрица $l_{ij}A$ совпадает с A , с тем лишь исключением, что i -я и j -я ее строки переставлены, в то время как произведение Al_{ij} означает перестановку i -го и j -го столбцов матрицы A .

4. Пусть H_{ij} — матрица, ij -й элемент которой есть h , а остальные элементы равны нулю. Показать, что произведение $(I + H_{ij})A$ есть матрица, совпадающая с A , с тем лишь исключением, что ее i -я строка заменена i -й строкой, к которой прибавлена j -я, предварительно умноженная на h . В то же время матрица $A(I + H_{ij})$ отличается от матрицы A аналогичным преобразованием столбцов.

5. Пусть матрица H_r равна единичной I с тем лишь исключением, что ее rr -й элемент равен k . Чему равны произведения H_rA и AH_r ?

6. Если матрица A вещественна и $AA^t = 0$, то $A = 0$.

7. Если $AA^b = 0$, то $A = 0$.

8. Показать, что если T — ортогональная матрица, то ее элементы ограничены общей константой. Аналогично, если U — унитарная матрица, то ее элементы также ограничены.

9. Пусть d_1 — определитель системы линейных однородных уравнений

$$\left[\sum_{j=1}^N a_{ij}x_j \right] \left[\sum_{j=1}^N a_{kj}x_j \right] = 0, \quad i, k = 1, 2, \dots, N,$$

где неизвестными считаются $N(N+1)/2$ величин x_jx_j , $i, j = 1, 2, \dots, N$. Тогда

$$d_1 = |a_{ij}|^{N+1}. \quad (\text{Шефли.})$$

10. Показать, что пределы $\lim_{n \rightarrow \infty} A^n$, $\lim_{n \rightarrow \infty} B^n$ могут существовать, в то время как предел $\lim_{n \rightarrow \infty} (AB)^n$ не существует. Это достаточно показать для матриц A и B второго порядка.

11. Предположим, что мы имеем систему линейных уравнений $\sum_{j=1}^N a_{ij}x_j = b_i$, $i=1, 2, \dots, N$. Если, образуя линейные комбинации данных уравнений, можно свести систему к виду $x_1=c_1; x_2=c_2; \dots; x_N=c_N$, то последние равенства представляют собой решение исходной системы и притом единственное. (Р. Робинсон.)

12. Введем скобки Якоби $[A, B]=AB-BA$, коммутатор матриц A и B . Показать непосредственным вычислением, что

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

13. Пусть $r_1 = e^{2\pi i/n}$ — неприводимый корень из единицы и пусть $r_k = e^{2\pi i k/n}$, $k=1, 2, \dots, n-1$. Рассмотрим матрицу

$$T = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & r_1 & \dots & r_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & r_{n-1} & \dots & r_{n-1}^{n-1} \end{pmatrix}.$$

Показать, что $n^{-1/2}T$ — унитарная матрица. Следовательно, если $x_k = \sum_{j=0}^{n-1} e^{2\pi i k j/n} y_j$, то

$$y_k = \frac{1}{n} \sum_{j=0}^{n-1} e^{-2\pi i k j/n} x_j$$

$$\sum_{k=0}^{n-1} |x_k|^2 = n \sum_{k=0}^{n-1} |y_k|^2.$$

Это преобразование называется *конечным преобразованием Фурье*.

14. Предположим, что

$$\sum_{i,j=1}^N a_{ij}x_i y_j = \sum_{k=1}^N c_k \left\{ \sum_{s=k}^N b_{ks}x_s \right\} \left\{ \sum_{l=k}^N d_{kl}y_l \right\}$$

при $b_{kk}=d_{kk}=1$ для всех k . Тогда $|a_{ij}| = \prod_{k=1}^N c_k^{-1}$.

15. Рассмотрим преобразование Гаусса $B = \|b_{ij}\|$ матрицы $A = \|a_{ij}\|$:

$$b_{ik} = \delta_{ik} a_{ik}, \dots, b_{ik} = a_{ik}^{-1} (a_{11} a_{ik} - a_{1k} a_{11}), \quad k > 1.$$

Пусть $A_{11} = \|a_{ij}\|$, $i, j=2, \dots, N$. Показать, что

$$|\lambda I - B| = (\lambda a_{11}^{-1} - 1) |\lambda |\lambda I - A_{11}| - |\lambda I - A|.$$

(Д. М. Котелянский.)

¹⁾ Берчнелл (J. L. Burchnell), Proc. Edinburgh Math. Soc. (2), 9 (1954), 100—104.

8. Любая ортогональная матрица размеров 3×3 может быть представлена как произведение

$$\begin{pmatrix} \cos a & \sin a & 0 \\ -\sin a & \cos a & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos b & 0 & -\sin b \\ 0 & 1 & 0 \\ \sin b & 0 & \cos b \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos c & \sin c \\ 0 & -\sin c & \cos c \end{pmatrix}.$$

Каков геометрический смысл этого результата?

Упражнения к гл. 2

1. Доказать, что

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1N} & x_1 \\ a_{21} & a_{22} & \dots & a_{2N} & x_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} & x_N \\ x_1 & x_2 & \dots & x_N & 0 \end{vmatrix} = -(a^{11}x_1^2 + \dots + (a^{jj} + a^{jj})x_j^2 + \dots),$$

где a^{jj} — алгебраическое дополнение элемента a_{ij} в определителе $|a_{ij}|$.

2. Пусть l_{ij} обозначает матрицу, полученную перестановкой i -й и j -й строк в единичной матрице I . Доказать, что

$$l_{ij}^2 = I, \quad l_{ik}l_{kj}l_{ji} = l_{kj}.$$

3. Показать, что матрица $l_{ij}A$ совпадает с A , с тем лишь исключением, что i -я и j -я ее строки переставлены, в то время как произведение Al_{ij} означает перестановку i -го и j -го столбцов матрицы A .

4. Пусть H_{ij} — матрица, ij -й элемент которой есть h , а остальные элементы равны нулю. Показать, что произведение $(I + H_{ij})A$ есть матрица, совпадающая с A , с тем лишь исключением, что ее i -я строка заменена i -й строкой, к которой прибавлена j -я, предварительно умноженная на h . В то же время матрица $A(I + H_{ij})$ отличается от матрицы A аналогичным преобразованием столбцов.

5. Пусть матрица H_r равна единичной I с тем лишь исключением, что ее rr -й элемент равен k . Чему равны произведения H_rA и AH_r ?

6. Если матрица A вещественна и $AA' = 0$, то $A = 0$.

7. Если $AA^0 = 0$, то $A = 0$.

8. Показать, что если T — ортогональная матрица, то ее элементы ограничены общей константой. Аналогично, если U — унитарная матрица, то ее элементы также ограничены.

9. Пусть d_1 — определитель системы линейных однородных уравнений

$$\left[\sum_{j=1}^N a_{ij}x_j \right] \left[\sum_{j=1}^N a_{kj}x_j \right] = 0, \quad i, k = 1, 2, \dots, N,$$

где неизвестными считаются $N(N+1)/2$ величины x_ix_j , $i, j = 1, 2, \dots, N$. Тогда

$$d_1 = |a_{ij}|^{N+1}. \quad (\text{Шефли.})$$

10. Показать, что пределы $\lim_{n \rightarrow \infty} A^n$, $\lim_{n \rightarrow \infty} B^n$ могут существовать, в то время как предел $\lim_{n \rightarrow \infty} (AB)^n$ не существует. Это достаточно показать для матриц A и B второго порядка.

11. Предположим, что мы имеем систему линейных уравнений $\sum_{j=1}^N a_{ij}x_j = b_i$, $i=1, 2, \dots, N$. Если, образуя линейные комбинации данных уравнений, можно свести систему к виду $x_1=c_1; x_2=c_2; \dots; x_N=c_N$, то последние равенства представляют собой решение исходной системы и притом единственное. (Р. Робинсон.)

12. Введем скобки Якоби $[A, B]=AB-BA$, коммутатор матриц A и B . Показать непосредственным вычислением, что

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

13. Пусть $r_1 = e^{2\pi i/n}$ — неприводимый корень из единицы и пусть $r_k = e^{2\pi i k/n}$, $k=1, 2, \dots, n-1$. Рассмотрим матрицу

$$T = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & r_1 & \dots & r_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & r_{n-1} & \dots & r_{n-1}^{n-1} \end{pmatrix}.$$

Показать, что $n^{-1/2}T$ — унитарная матрица. Следовательно, если $x_k = \sum_{j=0}^{n-1} e^{2\pi i k j/n} y_j$, то

$$y_k = \frac{1}{n} \sum_{j=0}^{n-1} e^{-2\pi i k j/n} x_j$$

и

$$\sum_{k=0}^{n-1} |x_k|^2 = n \sum_{k=0}^{n-1} |y_k|^2.$$

Это преобразование называется *конечным преобразованием Фурье*.

14. Предположим, что

$$\sum_{i,j=1}^N a_{ij}x_i y_j = \sum_{k=1}^N c_k \left\{ \sum_{s=k}^N b_{ks}x_s \right\} \left\{ \sum_{t=k}^N d_{kt}y_t \right\}$$

при $b_{kk}=d_{kk}=1$ для всех k . Тогда $|a_{ij}| = \prod_{k=1}^N e_k^2$.

15. Рассмотрим преобразование Гаусса $B = [b_{ij}]$ матрицы $A = [a_{ij}]$:

$$b_{i1} = \delta_{i1} a_{i1}, \dots, b_{ik} = a_{ik}^{-1} (a_{i1} a_{ik} - a_{i1} a_{1k}), \quad k > 1.$$

Пусть $A_{11} = |a_{ij}|$, $i, j=2, \dots, N$. Показать, что

$$|\lambda I - B| = (\lambda a_{11}^{-1} - 1) |\lambda | \lambda I - A_{11} | - |\lambda I - A|.$$

(Д. М. Котелянский.)

¹⁾ Берчнелл (J. L. Burchnell), Proc. Edinburgh Math. Soc. (2), 9 (1954), 100—104.

16. Показать, что матрицы

$$A = \begin{vmatrix} 1 & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \quad B = \begin{vmatrix} 1 & -a & -b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

удовлетворяют соотношению $AB = I$.

17. Показать, что

$$\begin{vmatrix} x & y & z \\ z & x & y \\ y & z & x \end{vmatrix} = (x+y+z)(x+\omega y+\omega^2 z)(x+\omega^2 y+\omega z) = x^3 + y^3 + z^3 - 3xyz,$$

где ω — кубический корень из единицы.

18. Пользуясь этим результатом, показать, что если $Q(x) = x_1^3 + x_2^3 + x_3^3 - 3x_1x_2x_3$, то $Q(x)Q(y) = Q(z)$, где z — билинейная форма x_i и y_i ($z_i = \sum a_{ijk}x_jy_k$) с коэффициентами a_{ijk} , не зависящими от x_i и y_i .

19. Доказать, что

$$\begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2 & x_1 & -x_4 & x_3 \\ -x_3 & x_4 & x_1 & -x_2 \\ -x_4 & -x_3 & x_2 & x_1 \end{vmatrix} = (x_1^2 + x_2^2 + x_3^2 + x_4^2)^2$$

и вследствие этого

$$(x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) = z_1^2 + z_2^2 + z_3^2 + z_4^2,$$

где z_i — билинейные формы x_i и y_i . Гурвицем было показано, что произведение суммы N квадратов, умноженное на произведение суммы N квадратов, есть сумма N квадратов в указанном выше смысле только тогда, когда $N = 1, 2, 4, 8^1$).

Изложение теории воспроизводящихся форм см. у Мак-Даффи ²⁾.

20. Доказать, что матрица A , элементы которой определяются формулами

$$a_{ij} = \begin{cases} (-1)^{j-1} \binom{j-1}{i-1}, & i < j; \\ (-1)^{j-1}, & i = j; \\ 0, & i > j, \end{cases}$$

удовлетворяет соотношению $A^2 = I$. Здесь $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

21. Пусть $y_i = y_i(x_1, x_2, \dots, x_N)$ — система N функций независимых переменных x_i , $i = 1, 2, \dots, N$. Матрица $J = J(y, x) = \|\partial y_i / \partial x_j\|$ называется матрицей Якоби, а ее определитель — якобианом преобразования. Показать, что

$$J(z, y)J(y, x) = J(z, x).$$

22. Рассмотреть связь между N^2 переменными y_{ij} и N^2 переменными x_{ij} , определяемую соотношением $Y = AXB$, где A и B — постоянные матрицы. Показать, что $|J(y, x)| = |A|^N |B|^N$.

¹⁾ Гурвиц (A. Hurwitz), Über die Komposition der quadratischen Formen von beliebig vielen Variablen, Math. Werke, bd II, Basel, 1933, 565—571.

²⁾ Мак-Даффи (C. S. MacDuffee), On the Composition of Algebraic Forms of Higher Degree, Bull. Amer. Math. Soc. 51 (1945), 198—211; Радон (J. Radon), Lineare Scharen orthogonaler Matrizen, Abh. Math. Sem. Hamb. I (1921), 1—14.

23. Если $Y = XX'$, где X — нижняя треугольная матрица, то $|J(X, Y)| = 2^N \prod_{i=1}^N x_{ii}^{N-i+1}$.

24. Показать, что проблема определения максимума функции $(x, Ax) - 2(x, b)$ приводит к векторному уравнению $Ax = b$.

25. Аналогично показать, что проблема минимизации функции пары векторов x и y : $(x, Bx) + 2(x, Ay) + (y, By) - 2(a, x) - 2(b, y)$ приводит к системе уравнений $Bx + Ay = a$, $A'x + By = b$ (предполагается, что $B = B'$).
Прим. ред.)

26. Пусть $f(x)$ — функция переменной x , принимающей только два значения: $x = 0$, $x = 1$. Показать, что $f(x)$ может быть записана в форме $a + bx$, где $a = f(0)$, $b = f(1) - f(0)$.

27. Пусть $g(x)$ — функция того же типа, которая сама принимает значения 0 или 1. Тогда $f(g(x)) = a_1 + b_1x$. Показать, что коэффициенты a_1 и b_1 представляют собой линейные комбинации коэффициентов a и b и что, таким образом, результат замены x на $g(x)$ эквивалентен матричному преобразованию вектора с компонентами a и b .

28. Пусть $f(x_1, x_2)$ — функция двух переменных, каждая из которых принимает только два значения: 0 или 1. Показать, что можно записать

$$f(x_1, x_2) = a_1 + a_2x_1 + a_3x_2 + a_4x_1x_2.$$

29. Пусть $g_1(x_1, x_2)$, $g_2(x_1, x_2)$ — функции того же типа, которые сами могут принимать только значения 0 или 1. Тогда, записав

$$f(g_1, g_2) = a'_1 + a'_2x_1 + a'_3x_2 + a'_4x_1x_2,$$

мы приходим к матричному преобразованию:

$$\begin{pmatrix} a'_1 \\ a'_2 \\ a'_3 \\ a'_4 \end{pmatrix} = M \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}.$$

Вычислить матрицу M для случая, когда $g_1(x_1, x_2) = x_1x_2$, $g_2(x_1, x_2) = x_1(1 - x_2)$.

30. Обобщить предыдущий результат на случай, когда мы имеем функцию N переменных $f(x_1, x_2, \dots, x_N)$ и x_i может принимать любое конечное множество значений. (Приведенные результаты используются при изучении различных типов логических сетей. См., например, Беллман, Холланд и Калаба (R. Bellman, J. Holland and R. Kalaba), *Dynamic Programming and the Synthesis of Logical Nets*, J. Assoc. Comp. Mach., 1959.)

31. Пусть A — заданная матрица размеров 2×2 и X — неизвестная матрица размеров 2×2 . Показать, что уравнение $AX - XA = I$ не имеет решения.

32. Показать, что тот же результат имеет место для случая, когда A и X — матрицы размеров $N \times N$.

33. Рассмотреть связь между $N(N+1)/2$ переменными y_{ij} и $N(N+1)/2$ переменными x_{ij} , определенную уравнением $Y = AXA'$, где X и Y — симметрические матрицы. Показать, что $|J(Y, X)| = |A|^{N+1}$. Две статьи полуописательного характера, обсуждающие вопросы такого типа, принадлежат Димеру и Олкину¹⁾.

¹⁾ Димер и Олкин (W. L. Deemer and I. Olkin), *Jacobian of Matrix Transformations Useful in Multivariate Analysis*, *Biometrika* 38 (1951), 345—367.

84. Построить симметрическую ортогональную матрицу размеров 4×4 , элементы которой равны ± 1 . Этот вопрос и его возможные развития интересны в связи с неравенством Адамара, см. § 7 гл. 8 и (несколько неожиданно) в связи с задачей планирования экспериментов¹⁾.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

Читателю, который интересуется истоками теории матриц, необходимо обратиться к монографии

Мак-Даффи (C. C. MacDuffee), *The Theory of Matrices*, Chelsea Publishing Co., New York, 1946^{*)}.

Хотя мы иногда будем ссылаться на различные теоремы, пользуясь именами их авторов, мы не будем серьезно пытаться установить автора каждого результата. Как и во многих других областях математики, имеется определенное количество ошибочных приписок. В тех случаях, когда мы знали об этом, мы пытались исправить положение вещей.

§ 8. Следует отметить, что при иной физической ситуации могут быть введены другие виды алгебр с совсем различными определениями операций «сложения» и «умножения». Как показал Олкин, при рассмотрении задач социометрики весьма полезно определение умножения в форме $A \cdot B = \|a_i b_{ij}\|$. Впоследствии будет показано, что несколько неожиданным образом указанное здесь произведение Шура возникает в теории дифференциальных уравнений в частных производных.

§ 10. Подробное обсуждение свойства коммутативности имеется в статье

Таусски (O. Tausky), *Commutativity in Finite Matrices*, *Amer. Math. Monthly* 64 (1957), 229—235.

Интересное развитие коммутативности можно найти в работе

Фридман (B. Friedman), *n-commutative Matrices*, *Math. Ann.* 136 (1958), 343—347.

Решение уравнений $XA = AX$, $XA = A'X$ содержится в статье

Фаукс (H. O. Foukes), *J. London Math. Soc.* 17 (1942), 70—80.

§ 11. Может быть построена другая алгебра, в которой нарушена как коммутативность, так и ассоциативность произведения.

¹⁾ Пэли (R. E. A. C. Paley), *On Orthogonal Matrices*, *J. Math. and Physics* 12 (1933), 311—320; Плаккетт и Бермен (R. L. Plackett and J. P. Burman), *The Design of Optimum Multifactorial Experiments*, *Biometrika* 88 (1946), 305—325.

^{*)} См. также Ф. Р. Гантмахер, *Теория матриц*, «Наука», 1967. (Прим. ред.)

Особенно интересный пример дает алгебра Кэли. Для обсуждения этих вопросов см.

Алберт (A. A. Albert), *Modern Higher Algebra*, University of Chicago Press, Chicago, 1937; Биркгоф и Маклейн (G. A. Birkhoff and S. MacLane), *Survey of Modern Algebra*, The Macmillan Company, New York, 1958.

§ 14. В случае более общих преобразований или операторов A' часто называют *сопряженным* преобразованием или сопряженным оператором. Важность этого понятия связана с тем, что иногда изучение сопряженного преобразования может оказаться проще рассмотрения исходного. Более того, во многих случаях векторные свойства преобразования A выражаются более просто, когда они формулируются в терминах A' . Пример этого мы получим в гл. 14 при рассмотрении марковских матриц.

§ 16. Обозначение H^* для матрицы H' предложено, по-видимому, Островским; см. статью

Островский (A. Ostrowski), Über die Existenz einer endlichen Basis bei gewissen Funktionen Systemen, *Math. Ann.* 78 (1917), 94—119.

ДИАГОНАЛИЗАЦИЯ И КАНОНИЧЕСКИЕ ФОРМЫ СИММЕТРИЧЕСКИХ МАТРИЦ

§ 1. Резюме. Рассмотрение проблемы стационарных значений квадратичной формы $Q(\mathbf{x}) = \sum_{i,j=1}^N a_{ij}x_i x_j$ на сфере $x_1^2 + x_2^2 + \dots + x_N^2 = 1$ привело нас к задаче нахождения нетривиальных решений системы линейных однородных уравнений

$$\sum_{j=1}^N a_{ij}x_j = \lambda x_i, \quad i = 1, 2, \dots, N. \quad (1)$$

Мы прервали наше изложение на этом месте, чтобы ввести векторно-матричные обозначения. В заключение экскурса указывалось, что новый аппарат позволит в простой и изящной форме трактовать этот и связанные с ним вопросы.

Отметим, что условия симметрии $a_{ij} = a_{ji}$ в уравнениях (1) выполнены автоматически в силу происхождения уравнений. Это простое, но существенное свойство позволит нам получить большое количество информации, касающейся природы решений. Используя свойство симметрии, мы преобразуем $Q(\mathbf{x})$ к более простой форме, которая играет важнейшую роль в теории матриц и квадратичных форм.

Продолжим теперь наше изложение.

§ 2. Решение системы линейных однородных уравнений. Нам необходим следующий фундаментальный результат.

Лемма. Для того чтобы линейная система

$$\sum_{j=1}^N b_{ij}x_j = 0, \quad i = 1, 2, \dots, N, \quad (1)$$

имела нетривиальное решение, необходимо и достаточно равенство нулю ее определителя

$$|b_{ij}| = 0. \quad (2)$$

Как обычно, под «нетривиальным» мы понимаем решение, в котором по крайней мере одно из неизвестных отлично от нуля. Этот результат фактически является частным случаем более

общего, относящегося к линейной системе, в которой число уравнений не равно числу неизвестных. Обсуждение этих вопросов имеется в приложении А. Однако здесь мы дадим простое индуктивное доказательство, устанавливающее все необходимые нам факты.

Доказательство леммы. Необходимость очевидна. Если $|b_{ij}| \neq 0$, систему можно решить по правилу Крамера и получить тем самым единственное тривиальное решение $x_1 = x_2 = \dots = x_N = 0$. Докажем достаточность. Очевидно, что результат справедлив для $N = 1$. Тогда посмотрим, можно ли установить истинность результата для N , предполагая его справедливым для $N - 1$. По крайней мере один из элементов b_{ij} не нуль, так как в противном случае результат тривиально верен. Предположим, что не равен нулю один из элементов первой строки, при этом, не теряя общности, можно принять, что этот элемент есть b_{11} .

Обращаясь к линейным уравнениям (1), исключим x_1 из первого и второго уравнений, из первого и третьего и т. д. Полученная система имеет вид

$$\begin{aligned} (b_{22} - \frac{b_{21}b_{12}}{b_{11}})x_2 + \dots + (b_{2N} - \frac{b_{21}b_{1N}}{b_{11}})x_N &= 0, \\ \vdots & \\ (b_{N2} - \frac{b_{N1}b_{12}}{b_{11}})x_2 + \dots + (b_{NN} - \frac{b_{N1}b_{1N}}{b_{11}})x_N &= 0. \end{aligned} \quad (3)$$

Установим соответствие между определителем этой системы и исходным определителем порядка N . Вычитая первую строку определителя, умноженную предварительно на b_{21}/b_{11} , из второй; первую строку, умноженную на b_{31}/b_{11} , из третьей и т. д., мы приходим к соотношению

$$\begin{vmatrix} b_{11} & b_{12} & \dots & b_{1N} \\ b_{21} & b_{22} & \dots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \dots & b_{NN} \end{vmatrix} = \begin{vmatrix} b_{11} & \dots & \dots & \dots \\ 0 & (b_{22} - \frac{b_{21}b_{12}}{b_{11}}) & \dots & (b_{2N} - \frac{b_{21}b_{1N}}{b_{11}}) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & (b_{N2} - \frac{b_{N1}b_{12}}{b_{11}}) & \dots & (b_{NN} - \frac{b_{N1}b_{1N}}{b_{11}}) \end{vmatrix}. \quad (4)$$

Отсюда следует, что определитель $(N - 1)$ -мерной системы (3) есть нуль, так как по предположению $b_{11} \neq 0$, а $|b_{ij}| = 0$. Из нашего индуктивного предположения следует, что существует нетривиальное решение системы (3): x_2, x_3, \dots, x_N .

Полагая

$$x_1 = - \sum_{j=2}^N b_{1j} \frac{x_j}{b_{11}}, \quad (5)$$

получаем тем самым искомое нетривиальное решение системы (1): x_1, x_2, \dots, x_n .

Упражнение

1. Показать, что если A — вещественная матрица и уравнение $Ax = 0$ имеет нетривиальное решение, то всегда существует нетривиальное вещественное решение.

§ 3. Собственные векторы и собственные значения. Используя векторно-матричные обозначения, запишем систему (1.1) в виде

$$Ax = \lambda x. \quad (1)$$

Из леммы § 2 следует, что нетривиальное решение этого уравнения существует в том и только в том случае, если λ является корнем уравнения

$$|a_{ij} - \lambda \delta_{ij}| = 0, \quad (2)$$

которое обычно будем записывать в виде

$$|A - \lambda I| = 0. \quad (3)$$

Это уравнение называется *характеристическим уравнением* матрицы A . Как полиномиальное уравнение относительно λ , оно имеет N корней, которые называются *характеристическими числами*, или *собственными значениями* матрицы A . Если все корни различны, то мы будем иногда пользоваться термином *простые*, противопоставляя его термину *кратные*.

Каждому собственному значению ставится в соответствие *собственный вектор*, определенный с точностью до скалярного множителя. Этот собственный вектор может быть найден индуктивным путем, намеченным в § 2, или с помощью метода, приведенного в приложении А. Ни один из этих способов не является привлекательным, особенно для больших значений N , так как их применение связано с большим объемом вычислений. В настоящее время не имеется простых методов нахождения собственных значений и собственных векторов матриц большого размера.

Как было отмечено в предисловии, в настоящей книге мы умышленно не приводим библиографии по вычислительным методам, применяемым для определения численных значений собственных значений и собственных векторов*).

Если λ — кратный корень характеристического уравнения, то для произвольной квадратной матрицы может не быть равного кратности корня числа соответствующих собственных векторов.

*) По этому поводу см. Д. К. Фаддеев и В. Н. Фаддеева. Вычислительные методы в линейной алгебре, Физматгиз, 1963; Дж. Х. Уилкинсон, Алгебраическая проблема собственных значений, «Наука», 1970. (Прим. ред.)

Эти вопросы обсуждаются во второй части настоящей книги, посвященной изучению общих, не обязательно симметрических матриц. В случае симметрических матриц кратные корни приводят к определенным неудобствам, однако они легко преодолеваются. Мы покажем, что вещественная симметрическая матрица порядка N имеет N различных собственных векторов.

Упражнения

1. Матрицы A и A' имеют одинаковые собственные значения.
2. $T'AT - \lambda I = T'(A - \lambda I)T$, если матрица T — ортогональная. Показать, что матрицы A и $T'AT$ имеют одинаковые собственные значения, если T — ортогональная матрица.
3. Матрицы A и T^*AT имеют одинаковые собственные значения, если матрица T — унитарная.
4. Матрицы SAT и A имеют одинаковые собственные значения, если $ST = I$.
5. Показать непосредственным вычислением для матриц A и B размеров 2×2 , что матрицы AB и BA имеют тождественные характеристические уравнения.
6. Справедлив ли этот результат в общем случае?
7. Показать, что произведение собственного вектора на любой скалярный множитель, кроме нуля, также есть собственный вектор. Используя этот факт, показать, что мы всегда можем выбрать собственный вектор x так, что $(x, x) = 1$.
8. Показать, рассматривая матрицы размеров 2×2 , что собственные значения матрицы $A + B$ не могут быть получены в общем случае как сумма собственных значений матриц A и B .
9. Показать, что аналогичное утверждение справедливо для собственных значений матрицы AB .
10. Для случая матрицы размеров 2×2 получить соотношение между собственными значениями матриц A и A^2 .
11. Существует ли соответствующая связь между собственными значениями матриц A и A^n для $n = 3, 4, \dots$?

§ 4. Два фундаментальных свойства симметрических матриц. Приведем простые доказательства двух фундаментальных результатов, которые по существу являются основой анализа вещественных симметрических матриц. Хотя мы интересуемся симметрическими матрицами, элементы которых вещественны, мы будем иногда добавлять слово «вещественный» для того, чтобы подчеркнуть этот факт и предупредить любую возможную ошибку.

Теорема 1. *Собственные значения вещественной симметрической матрицы вещественны.*

Доказательство. Предположим обратное. Так как A — вещественная матрица, то из характеристического уравнения $|A - \lambda I| = 0$ следует, что число, комплексно сопряженное каждому собственному значению, также есть собственное значение. Мы получаем этот результат и всю дальнейшую информацию из того факта, что если

$$Ax = \lambda x, \quad (1)$$

то справедливо также равенство

$$A\bar{x} = \bar{\lambda}x. \quad (2)$$

Из этих соотношений следует

$$\begin{aligned} (\bar{x}, Ax) &= \lambda(\bar{x}, x), \\ (x, A\bar{x}) &= \bar{\lambda}(x, x). \end{aligned} \quad (3)$$

Так как матрица A симметрическая, то $(\bar{x}, Ax) = (A\bar{x}, x) = (x, A\bar{x})$ и полученные соотношения дают

$$0 = (\lambda - \bar{\lambda})(x, \bar{x}), \quad (4)$$

откуда $\lambda = \bar{\lambda}$. Последнее противоречит принятому предположению. Это означает также, что собственные векторы вещественной симметрической матрицы A всегда могут быть выбраны вещественными. Мы будем поступать именно таким образом.

Второй результат содержится в следующей теореме.

Теорема 2. *Собственные векторы, соответствующие различным собственным значениям вещественной симметрической матрицы, ортогональны.*

Доказательство. Из равенств

$$\begin{aligned} Ax &= \lambda x, \\ Ay &= \mu y, \end{aligned} \quad (5)$$

$\lambda \neq \mu$, мы получаем

$$\begin{aligned} (y, Ax) &= \lambda(y, x), \\ (x, Ay) &= \mu(x, y). \end{aligned} \quad (6)$$

Так как $(x, Ay) = (Ax, y) = (y, Ax)$, то вычитание дает

$$0 = (\lambda - \mu)(x, y), \quad (7)$$

откуда следует $(x, y) = 0$.

Этот результат является основным по своей важности. Его обобщение на операторы более общего вида представляет собой один из краеугольных камней классического анализа.

Упражнения

1. Один собственный вектор не может соответствовать двум различным собственным значениям.

2. Показать для матрицы размеров 2×2 , что два различных вектора могут соответствовать одному собственному значению.

3. Показать на примере, что существуют симметрические матрицы A и B размеров 2×2 такие, что определитель $|A - \lambda B|$ тождественно равен нулю. При каком условии, налагаемом на элементы матрицы B , мы можем утверждать, что все корни уравнения $|A - \lambda B| = 0$ вещественны?

4. Показать, что если матрицы второго порядка A и B вещественные симметрические и, кроме того, B положительно определенная, то все корни уравнения $|A - \lambda B| = 0$ вещественны.

5. Показать, что собственные значения эрмитовой матрицы вещественны и собственные векторы, соответствующие различным собственным значениям, ортогональны. Для доказательства воспользоваться обобщенным скалярным произведением (x, y)

6. Пусть элементы a_{ij} матрицы A зависят от параметра t . Показать, что производная от определителя $|A|$ по t может быть записана как сумма N определителей, где k -й определитель получен дифференцированием элементов k -й строки при неизменных остальных строках.

7. Показать, что производная от $|A - \lambda I|$ по λ равна $-\sum_{k=1}^N |A_k - \lambda I|$,

где A_k — матрица размеров $(N-1) \times (N-1)$, полученная из A вычеркиванием k -й строки и k -го столбца.

8. Из результата предыдущего упражнения следует, что если λ — простое собственное значение матрицы A , то по крайней мере один из определителей $|A_k - \lambda I|$ отличен от нуля.

9. Используя этот результат, показать, что если λ — простое собственное значение матрицы A , то компоненты соответствующего собственного вектора всегда могут быть записаны в виде полиномов от λ и от элементов матрицы A .

§ 5. Приведение к диагональной форме. Различные собственные значения. Мы без особых усилий можем получить весьма важный результат, если примем упрощающее предположение, что матрица A имеет различные собственные значения $\lambda_1, \lambda_2, \dots, \lambda_N$. Пусть x^1, x^2, \dots, x^N — соответствующие собственные векторы, нормированные условием

$$(x^i, x^i) = 1, \quad i = 1, 2, \dots, N. \quad (1)$$

Рассмотрим матрицу T , столбцами которой являются векторы x^i :

$$T = \|x^1, x^2, \dots, x^N\|. \quad (2)$$

Тогда T' есть матрица, строками которой являются векторы x^i :

$$T' = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^N \end{bmatrix}. \quad (3)$$

Вследствие ортогональности x^i как собственных векторов, соответствующих различным собственным значениям симметрической матрицы, мы получаем, что

$$T'T = \|(x^i, x^j)\| = \|\delta_{ij}\|, \quad (4)$$

т. е. T — ортогональная матрица. Мы можем теперь утверждать, что произведение AT имеет простую форму

$$AT = \|\lambda_1 x^1, \lambda_2 x^2, \dots, \lambda_N x^N\|, \quad (5)$$

т. е. столбцами матрицы-произведения будут векторы $\lambda_i x^i$.

Еще раз учитывая ортогональность векторов x^i , получаем

$$T'AT = \|\lambda_i(x^i, x^j)\| = \|\lambda_i \delta_{ij}\| = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix}. \quad (6)$$

На главной диагонали матрицы в правой части последнего равенства находятся собственные значения $\lambda_1, \lambda_2, \dots, \lambda_N$, все остальные элементы матрицы равны нулю. Матрица такого типа, как отмечалось ранее, называется *диагональной*.

Умножая последнее выражение справа на T' и слева на T , в силу свойства ортогональности $TT' = I$ получаем важный результат

$$A = T \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix} T'. \quad (7)$$

Описанный процесс называется приведением матрицы к диагональной форме. Как мы увидим впоследствии, это представление играет фундаментальную роль в теории симметрических матриц. Введем обозначение

$$\Lambda = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix}. \quad (8)$$

Упражнения

1. Показать, что $\Lambda^k = \|\lambda_i^k \delta_{ij}\|$ и что $A^k = T\Lambda^k T'$ при $k=1, 2, \dots$
2. Показать, что матрица A , имеющая различные собственные значения, удовлетворяет своему характеристическому уравнению. Этот факт является частным случаем более общего результата, который мы установим позднее.
3. Найти собственные векторы, соответствующие собственным значениям матриц A^k , $k=2, 3, \dots$, если матрица A имеет различные собственные значения.

§ 6. Приведение квадратичной формы к каноническому виду. Покажем теперь, что рассмотренное матричное преобразование приводит к важному преобразованию квадратичной формы $Q(x)$. Полагая $x = Ty$, где T определяется формулой (5.2), мы получаем фундаментальное соотношение

$$(x, Ax) = (Ty, ATy) = (y, T'ATy) = (y, \Lambda y) \quad (1)$$

или в скалярной форме

$$\sum_{i,j=1}^N a_{ij}x_i x_j = \sum_{i=1}^N \lambda_i y_i^2. \quad (2)$$

Так как матрица T ортогональная, то мы видим, что $x = Ty$ означает

$$T'x = T'Ty = y. \quad (3)$$

Следовательно, каждому значению x соответствует одно определенное значение y , и наоборот. Таким образом, мы получили чрезвычайно полезный результат. Множество значений, принимаемых формой $Q(x)$ на сфере $(x, x) = 1$, совпадает со значениями, принимаемыми квадратичной формой $(y, \Lambda y)$ на сфере $(y, y) = 1$.

Пока мы установили этот факт только для случая, когда все λ_i различны. Как мы увидим в гл. 4, этот результат справедлив в общем случае и представляет собой одно из основных положений теории квадратичных форм.

Упражнения

1. Пусть матрица A имеет различные положительные собственные значения. Используя полученный результат, вычислить объем N -мерного эллипсоида $(x, Ax) = 1$.

2. Следуя намеченному выше пути, доказать, что собственные значения эрмитовой матрицы вещественны и что собственные векторы, соответствующие различным собственным значениям, ортогональны. Для доказательства воспользоваться обозначением $[x, y]$, введенным в § 16 гл. 2.

3. Показать, что если собственные значения эрмитовой матрицы A различны, то мы можем найти такую унитарную матрицу T , что $A = T\Lambda T^*$. Это частный случай более общего результата, который будет доказан в следующей главе.

4. Пусть A — вещественная косимметрическая матрица, $A' = -A$. Показать, что ее собственные значения или нули, или чисто мнимые.

5. Пусть T — ортогональная матрица. Показать, что все ее собственные значения имеют абсолютную величину, равную 1*).

6. Пусть T — унитарная матрица. Показать, что все ее собственные значения имеют абсолютную величину, равную 1.

7. Предположим, что мы пытаемся получить представление (5.7), не делая предположения о простоте собственных значений симметрической матрицы A . Начнем с утверждения, что всегда можно найти симметрическую матрицу B с произвольно малыми элементами, обладающую тем свойством, что $A + B$ имеет простые собственные значения.

Мы не останавливаемся на этом подробно, так как доказательство этого утверждения несколько сложнее, чем можно было бы предположить, см. § 16 гл. 11. Пусть $\{\mu_i\}$ — собственные значения матрицы $A + B$. Тогда, как мы знаем, существует ортогональная матрица S такая, что

$$A + B = S \begin{vmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ 0 & & & \mu_N \end{vmatrix} S'.$$

*) Имеется в виду вещественная ортогональная матрица. (Прим. ред.)

Так как матрица S ортогональная, то ее элементы равномерно ограничены. Пусть $\{B_n\}$ — такая последовательность матриц, стремящихся к нулю, что соответствующая последовательность ортогональных матриц $\{S_n\}$ сходится. Тогда предел матриц $\{S_n\}$ должен быть ортогональной матрицей, обозначим его через T . Так как $\lim_{n \rightarrow \infty} \mu_i = \lambda_i$, то мы имеем

$$A = \lim_{n \rightarrow \infty} (A + B_n) = \lim_{n \rightarrow \infty} S_n \begin{vmatrix} \lim_{n \rightarrow \infty} \mu_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lim_{n \rightarrow \infty} \mu_N \end{vmatrix} \lim_{n \rightarrow \infty} S_n^*$$

Ввиду того что это доказательство основывается на аналитических понятиях, которые будут введены лишь позднее, мы не излагаем его в основном тексте. Оно иллюстрирует очень полезный математический принцип, состоящий в том, что результаты, действительные для общих вещественных симметрических матриц, всегда могут быть первоначально установлены для матриц с различными собственными значениями с последующим переходом к пределу.

§ 7. Положительно определенные квадратичные формы и матрицы. В § 9 гл. I мы ввели понятие положительно определенной квадратичной формы двух переменных и родственные понятия неотрицательной и неположительно определенной форм. Распространим эти определения на N -мерные квадратичные формы. Если $A = \|a_{ij}\|$ — вещественная симметрическая матрица и

$$Q_N(\mathbf{x}) = \sum_{i,j=1}^N a_{ij}x_i x_j > 0$$

для всех нетривиальных x_i , то мы будем говорить, что $Q_N(\mathbf{x})$ — положительно определенная форма и A — положительно определенная матрица.

Аналогично, если H — эрмитова матрица и форма $P_N(\mathbf{x}) = \sum_{i,j=1}^N h_{ij}x_i \bar{x}_j > 0$ для всех комплексных нетривиальных x_i , то мы будем говорить, что форма $P_N(\mathbf{x})$ и матрица H положительно определенные.

Упражнения

1. Пусть A — симметрическая матрица с различными собственными значениями; получить необходимые и достаточные условия того, что A — положительно определенная матрица.

2. Для заданной произвольной симметрической матрицы A существует ли скаляр c_1 такой, что матрица $A + c_1 I$ положительно определенная?

3. Показать, что мы можем записать симметрическую матрицу A в форме

$$A = \sum_{i=1}^N \lambda_i E_i,$$

где F_i — неотрицательные симметрические матрицы. При этом

$$A^k = \sum_{i=1}^N \lambda_i^k E_i$$

для $k=1, 2, \dots$

4. Если $p(\lambda)$ — полином относительно λ со скалярными коэффициентами:

$p(\lambda) = \sum_{k=0}^m c_k \lambda^k$, то через $p(A)$ обозначают матрицу $p(A) = \sum_{k=0}^m c_k A^k$. Пока-

зать, что $p(A) = \sum_{i=1}^N p(\lambda_i) E_i$.

Упражнения к гл. 3

1. Пусть A и B — две симметрические матрицы. Тогда, если матрица B положительно определенная, то все корни уравнения $|A - \lambda B| = 0$ вещественны. Что можно сказать о векторах x , удовлетворяющих соотношению $Ax = \lambda Bx$?

2. Каждая матрица единственным образом представима в форме $A = H + S$, где H — эрмитова матрица, а S — *косоэрмитова* матрица, т. е. такая, что $S^* = -S$.

3. В качестве обобщения теоремы 1 показать, что если λ — собственное значение вещественной матрицы A , то $|\operatorname{Im} \lambda| < d(N(N-1)2)^{1/2}$, где

$$d = \max_{1 \leq i, j \leq N} |a_{ij} - a_{ji}|/2. \quad (\text{Бендиксон.})$$

4. Вообще пусть A — комплексная матрица; тогда если

$$d_1 = \max_{i, j} |a_{ij}|, \quad d_2 = \max_{i, j} |a_{ij} + \bar{a}_{ji}|/2, \\ d_3 = \max_{i, j} |a_{ij} - \bar{a}_{ji}|/2,$$

то мы имеем

$$|\lambda| < Nd_1, \quad |\operatorname{Re} \lambda| < Nd_2, \quad |\operatorname{Im} \lambda| < Nd_3.$$

5. Показать, что для любой комплексной матрицы справедливы неравенства *)

$$\sum_{i=1}^N |\lambda_i|^2 < \sum_{i, j=1}^N |a_{ij}|^2, \\ \sum_{i=1}^N |\operatorname{Re} \lambda_i|^2 < \sum_{i, j=1}^N |(a_{ij} + \bar{a}_{ji})/2|^2, \\ \sum_{i=1}^N |\operatorname{Im} \lambda_i|^2 < \sum_{i, j=1}^N |(a_{ij} - \bar{a}_{ji})/2|^2. \quad (\text{Шур.})$$

6. Мы не исключали возможности того, что одному собственному значению могут соответствовать несколько линейно независимых собственных векторов. Как мы увидим, это возможно, если матрица A имеет кратные собственные значения, и этого не может быть, если собственные значения просты. Наибо-

*) Писются гораздо более тонкие результаты. См. статьи Брауэра (A. Brauer) в Duke Math. J., 1946, 1947, 1948, где имеются также дальнейшие ссылки. См. также Паркер (W. V. Parker), Characteristic Roots and Fields of Value of a Matrix, Bull. Amer. Math. Soc. 57 (1951), 103—105.

*) См. теорему Шура ниже в § 13 гл. 11. (Прим. ред.)

ПРИВЕДЕНИЕ СИММЕТРИЧЕСКИХ МАТРИЦ К ДИАГОНАЛЬНОЙ ФОРМЕ В ОБЩЕМ СЛУЧАЕ

§ 1. Введение. В этой главе мы хотим показать, что результаты, полученные в гл. 3 в предположении простоты собственных значений, в действительности имеют место в случае симметрических матриц общего вида. Доказательство, которое мы приведем, дает превосходные возможности для обсуждения понятия линейной независимости и метода ортогонализации Грама—Шмидта. При этом будут рассмотрены некоторые другие интересные методы.

§ 2. Линейная зависимость. Пусть x^1, x^2, \dots, x^k — набор k N -мерных векторов. Мы говорим, что эти векторы *линейно зависимы*, если существуют скаляры c_1, c_2, \dots, c_k , из которых по крайней мере один отличен от нуля, такие, что

$$c_1 x^1 + c_2 x^2 + \dots + c_k x^k = 0, \quad (1)$$

где 0 означает нулевой вектор. Если не существует таких скаляров, то мы говорим, что векторы *линейно независимы*.

Ссылаясь на результаты, полученные в приложении А, можно утверждать, что понятие линейной независимости представляет интерес только для случаев $k \leq N$, так как любые k векторов, где $k > N$, связаны соотношением вида (1).

Упражнения

1. Показать, что произвольные взаимно ортогональные ненулевые векторы линейно независимы.
2. Пусть задан ненулевой вектор в произвольном N -мерном пространстве. Показать, что всегда можно найти набор из $N-1$ векторов, который вместе с данным вектором образует линейно независимую систему.

§ 3. Ортогонализация Грама — Шмидта. Пусть x^1, x^2, \dots, x^N — вещественные линейно независимые N -мерные векторы. Мы хотим показать, что можно образовать некоторые линейные комбинации этих базисных векторов, которые будут представлять собой систему взаимно ортогональных векторов.

Процедура, которой мы следуем, индуктивна. Начнем с определения двух новых векторов

$$\begin{aligned} y^1 &= x^1, \\ y^2 &= x^2 + a_{11}x^1, \end{aligned} \quad (1)$$

где a_{11} — скаляр, определенный из условия, что векторы y^1 и y^2 ортогональны. Соотношение

$$(y^1, y^2) = (x^1, x^2 + a_{11}x^1) = 0 \quad (2)$$

дает величину a_{11} :

$$a_{11} = -\frac{(x^1, x^2)}{(x^1, x^1)}. \quad (3)$$

Так как векторы x^1 по предположению линейно независимы, то вектор x^1 не может быть равен нулю, так что скалярное произведение (x^1, x^1) отлично от нуля.

Пусть, далее,

$$y^3 = x^3 + a_{21}x^1 + a_{22}x^2, \quad (4)$$

где теперь уже два коэффициента a_{21} и a_{22} определяются из условий ортогональности:

$$(y^3, y^1) = 0, \quad (y^3, y^2) = 0. \quad (5)$$

Эти условия упрощаются, если в силу (1) заменить уравнение (5) эквивалентными соотношениями

$$(y^3, x^1) = 0, \quad (y^3, x^2) = 0. \quad (6)$$

Последние приводят к системе уравнений, линейной относительно неизвестных коэффициентов a_{21} и a_{22} :

$$\begin{aligned} (x^3, x^1) + a_{21}(x^1, x^1) + a_{22}(x^2, x^1) &= 0, \\ (x^3, x^2) + a_{21}(x^1, x^2) + a_{22}(x^2, x^2) &= 0. \end{aligned} \quad (7)$$

Эти уравнения можно решить, используя правило Крамера, при условии, что определитель

$$D_2 = \begin{vmatrix} (x^1, x^1) & (x^1, x^2) \\ (x^1, x^2) & (x^2, x^2) \end{vmatrix} \quad (8)$$

не равен нулю.

Чтобы показать, что определитель D_2 не равен нулю, поступим следующим образом. Предположим противное, $D_2 = 0$. Тогда из леммы, установленной в § 2 гл. 3, следует, что существуют два скаляра r_1 и s_1 , из которых по крайней мере один отличен от нуля, удовлетворяющие линейным уравнениям

$$\begin{aligned} r_1(x^1, x^1) + s_1(x^1, x^2) &= 0, \\ r_1(x^1, x^2) + s_1(x^2, x^2) &= 0. \end{aligned} \quad (9)$$

Эти уравнения могут быть записаны в форме

$$\begin{aligned}(x^1, r_1 x^1 + s_1 x^2) &= 0, \\ (x^2, r_1 x^1 + s_1 x^2) &= 0.\end{aligned}\quad (10)$$

Умножая первое уравнение на r_1 , а второе на s_1 и складывая результаты, мы получаем

$$(r_1 x^1 + s_1 x^2, r_1 x^1 + s_1 x^2) = 0. \quad (11)$$

Однако это уравнение в силу вещественности коэффициентов может быть выполнено, только если $r_1 x^1 + s_1 x^2 = 0$, что противоречит принятому допущению о линейной независимости векторов x^i . Следовательно, определитель $D_2 \neq 0$, и система (7) имеет единственное решение.

На следующем шаге мы вводим вектор

$$y^1 = x^1 + a_{31} x^1 + a_{32} x^2 + a_{33} x^3. \quad (12)$$

Как и ранее, условия взаимной ортогональности

$$(y^1, x^1) = (y^1, x^2) = (y^1, x^3) = 0 \quad (13)$$

приводят к системе уравнений

$$(x^1, x^i) + a_{31}(x^1, x^i) + a_{32}(x^2, x^i) + a_{33}(x^3, x^i) = 0, \quad (14)$$

$$i = 1, 2, 3.$$

Мы можем решить эти уравнения относительно коэффициентов a_{31} , a_{32} и a_{33} , показав, что определитель

$$D_3 = |(x^i, x^j)|, \quad i, j = 1, 2, 3, \quad (15)$$

отличен от нуля. Доказательство того, что $D_3 \neq 0$, полностью аналогично проведенному для двумерного случая. Эту процедуру можно продолжить шаг за шагом до тех пор, пока не будет получен полный набор взаимно ортогональных векторов $\{y^i\}$.

Эти векторы могут быть нормированы условием $(y^i, y^i) = 1$. В этом случае мы говорим об ортонормированном наборе векторов. Определители D_k называются *определителями Грама*.

Упражнения

1. Рассмотрим отрезок $[-1, 1]$ и определим скалярное произведение двух вещественных функций $f(t)$ и $g(t)$ следующим образом:

$$(f, g) = \int_{-1}^1 f(t)g(t) dt.$$

Пусть $P_0(t) = \frac{1}{2}$. Другие элементы последовательности $\{P_n(t)\}$ вещественных полиномов степени n определены условиями $(P_n, P_m) = 0, m \neq n; (P_n, P_n) = 1$. Доказать, что $(P_n, t^m) = 0$ для $0 \leq m \leq n-1$, и построить несколько первых членов последовательности.

2. При том же определении скалярного произведения, что и в предыдущем упражнении, доказать, что

$$\left(\frac{d^n}{dt^n} (1-t^2)^n, t^m \right) = 0,$$

где $0 \leq m \leq n-1$, и, таким образом, выразить полиномы P_n через $\frac{d^n}{dt^n} (1-t^2)^n$.

Эти полиномы с точностью до постоянных множителей совпадают с классическими полиномами Лежандра.

3. Рассмотрим интервал $(0, \infty)$ и определим скалярное произведение двух вещественных полиномов $f(t)$ и $g(t)$ соотношением $(f, g) = \int_0^{\infty} e^{-t} f(t) g(t) dt$.

Пусть последовательность полиномов $\{L_n(t)\}$ определена условиями: $L_0(t) = 1$, $L_n(t)$ — полином степени n , и $(L_n, L_m) = 0$, $n \neq m$; $(L_n, L_n) = 1$. Доказать, что из этих условий следует равенство $(L_n, t^m) = 0$, $0 \leq m \leq n-1$, и сконструировать несколько первых членов последовательности.

4. Доказать, что $\left(e^t \frac{d^n}{dt^n} e^{-t} t^n, t^m \right) = 0$, $0 \leq m \leq n-1$, и, таким образом,

выразить полиномы L_n через $e^t \frac{d^n}{dt^n} e^{-t} t^n$. Эти полиномы с точностью до постоянных множителей совпадают с классическими полиномами Лагерра.

5. Рассмотрим интервал $(-\infty, \infty)$ и определим скалярное произведение двух вещественных полиномов $f(t)$ и $g(t)$:

$$(f, g) = \int_{-\infty}^{\infty} e^{-t^2} f(t) g(t) dt.$$

Пусть последовательность полиномов $\{H_n(t)\}$ определена следующим образом: $H_0(t) = 1$, $H_n(t)$ — полиномы степени n и $(H_m, H_n) = 0$, $m \neq n$; $(H_n, H_n) = 1$. Доказать, что $(H_n, t^m) = 0$, $0 \leq m \leq n-1$, и сконструировать несколько первых членов последовательности.

6. Доказать, что $\left(e^{t^2} \frac{d^n}{dt^n} e^{-t^2}, t^m \right) = 0$, $0 \leq m \leq n-1$, и выразить полиномы H_n через $e^{t^2} \frac{d^n}{dt^n} e^{-t^2}$. Эти полиномы с точностью до постоянных множителей совпадают с классическими полиномами Эрмита.

7. Показать, что если дан вещественный N -мерный вектор, нормированный условием $(x^1, x^1) = 1$, то мы можем найти $N-1$ дополнительных векторов x^2, x^3, \dots, x^N таких, что матрица $T = \|x^1, x^2, \dots, x^N\|$ ортогональная.

8. Получить аналог процедуры Грама—Шмидта для комплексных векторов.

§ 4. Положительность определителей Грама D_k . Все, что требовалось в предыдущем параграфе в той его части, где рассматривалась ортогонализация, — это неравенство нулю определителя D_k . Покажем теперь, что незначительное развитие приведенных ранее рассуждений позволит нам сделать вывод, что в действительности определители D_k положительны. Этот результат не очень важен на данном этапе, но сам метод доказательства интересен и иногда полезен.

Рассмотрим квадратичную форму

$$\begin{aligned} Q(u_1, u_2, \dots, u_k) &= \\ &= (u_1 x^1 + u_2 x^2 + \dots + u_k x^k, u_1 x^1 + u_2 x^2 + \dots + u_k x^k) = \\ &= \sum_{i, j=1}^k (x^i, x^j) u_i u_j, \quad (1) \end{aligned}$$

где u_i — вещественные величины, а x^i , как и ранее, образуют систему вещественных линейно независимых векторов. В силу линейной независимости векторов x^i квадратичная форма $Q > 0$ для всех нетривиальных значений u_i . Следовательно, свойство, которое мы хотим установить, является следствием общего результата, состоящего в том, что определитель

$$D = |a_{ij}| \quad (2)$$

любой положительно определенной формы

$$Q = \sum_{i, j=1}^N a_{ij} u_i u_j \quad (3)$$

положителен. Мы покажем, что свойство положительности просто следует из того факта, что рассматриваемый определитель всегда отличен от нуля.

Положительность определителя, как мы увидим, может быть легко установлена на основании результатов, которые мы получим в дальнейшем. Однако полезно проводить анализ различных результатов, выясняя, в какой мере они могут быть выведены из простых фактов.

Итак, начнем с замечания, что определитель D никогда не равен нулю. Действительно, если $D = 0$, то имеется нетривиальное решение системы однородных линейных уравнений

$$\sum_{j=1}^N a_{ij} u_j = 0, \quad i = 1, 2, \dots, N. \quad (4)$$

Эти равенства приводят нас к следующему:

$$Q = \sum_{i=1}^N u_i \left(\sum_{j=1}^N a_{ij} u_j \right) = 0, \quad (5)$$

что противоречит положительной определенности формы (3).

Это рассуждение в основном повторяет доказательство, которое мы привели в § 3.

Новое возникает при доказательстве положительности D . Рассмотрим семейство квадратичных форм, определяемое соотношением

$$P(\lambda) = \lambda Q + (1 - \lambda) \sum_{i=1}^N u_i^2, \quad (6)$$

где λ — скалярный параметр, изменяющийся на интервале $[0, 1]$.

Очевидно, что для всех λ в этом интервале квадратичная форма $P(\lambda)$ положительна при всех нетривиальных u_j . Следовательно, определитель квадратичной формы $P(\lambda)$ никогда не равен нулю.

Но для $\lambda=0$ определитель записывается в простой форме

$$\begin{vmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \\ 0 & & & & 1 \end{vmatrix} = 1 \quad (7)$$

и, очевидно, положителен. Поскольку, далее, определитель $P(\lambda)$ является непрерывной функцией параметра λ и не обращается в нуль при его изменении ($0 \leq \lambda \leq 1$), то положительность $P(\lambda)$ при $\lambda=0$ влечет положительность при $\lambda=1$. Это доказывает сформулированный результат.

§ 5. Одно тождество. Другой метод доказательства положительности определителя D_k для линейно независимых векторов x^i основывается на неравенстве, которое требуется нам снова в одной из следующих глав, посвященных неравенствам.

Теорема 1. Пусть $x^i, i=1, 2, \dots, k$, — набор N -мерных векторов, $N \geq k$; тогда

$$|(x^i, x^j)| = \frac{1}{k!} \sum_{\{i\}} \begin{vmatrix} x_{i_1}^1 & x_{i_2}^1 & \dots & x_{i_k}^1 \\ x_{i_1}^2 & x_{i_2}^2 & \dots & x_{i_k}^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_{i_1}^k & x_{i_2}^k & \dots & x_{i_k}^k \end{vmatrix}^2 \quad (1)$$

$(i, j=1, 2, \dots, k),$

где суммирование производится по всем индексам $\{i_k\}$, удовлетворяющим неравенствам $1 \leq i_1, i_2, \dots, i_k \leq N$.

Доказательство. Прежде чем дать доказательство, отметим, что в случае $k=2$ (1) превращается в тождество Лагранжа

$$\left(\sum_{i=1}^N x_i^2 \right) \left(\sum_{i=1}^N y_i^2 \right) - \left(\sum_{i=1}^N x_i y_i \right)^2 = \frac{1}{2} \sum_{i,j=1}^N (x_i y_j - x_j y_i)^2. \quad (2)$$

Для того чтобы используемый нами метод доказательства был ясен, достаточно рассмотреть случай $k=3$. Это позволит упростить

обозначения. Мы хотим показать, что

$$\begin{vmatrix} \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i y_i & \sum_{i=1}^N x_i z_i \\ \sum_{i=1}^N x_i y_i & \sum_{i=1}^N y_i^2 & \sum_{i=1}^N y_i z_i \\ \sum_{i=1}^N x_i z_i & \sum_{i=1}^N y_i z_i & \sum_{i=1}^N z_i^2 \end{vmatrix} = \frac{1}{3!} \sum_{1 < k, l, m < N} \begin{vmatrix} x_k & x_l & x_m \\ y_k & y_l & y_m \\ z_k & z_l & z_m \end{vmatrix}^2. \quad (3)$$

Начнем с равенства

$$\begin{vmatrix} x_k & x_l & x_m \\ y_k & y_l & y_m \\ z_k & z_l & z_m \end{vmatrix}^2 = \begin{vmatrix} x_k^2 + x_l^2 + x_m^2 & x_k y_k + x_l y_l + x_m y_m & x_k z_k + x_l z_l + x_m z_m \\ x_k y_k + x_l y_l + x_m y_m & y_k^2 + y_l^2 + y_m^2 & y_k z_k + y_l z_l + y_m z_m \\ x_k z_k + x_l z_l + x_m z_m & y_k z_k + y_l z_l + y_m z_m & z_k^2 + z_l^2 + z_m^2 \end{vmatrix}, \quad (4)$$

которое следует из правила перемножения определителей. Последний определитель может быть записан в виде суммы трех определителей, первый из которых является типичным:

$$\begin{vmatrix} x_k^2 & x_k y_k + x_l y_l + x_m y_m & x_k z_k + x_l z_l + x_m z_m \\ x_k y_k & y_k^2 + y_l^2 + y_m^2 & y_k z_k + y_l z_l + y_m z_m \\ x_k z_k & y_k z_k + y_l z_l + y_m z_m & z_k^2 + z_l^2 + z_m^2 \end{vmatrix}. \quad (5)$$

Вычитая из второго столбца первый, умноженный на y_k/x_k , и из третьего столбца первый, умноженный на z_k/x_k , мы получаем

$$\begin{vmatrix} x_k^2 & x_l y_l + x_m y_m & x_l z_l + x_m z_m \\ x_k y_k & y_l^2 + y_m^2 & y_l z_l + y_m z_m \\ x_k z_k & y_l z_l + y_m z_m & z_l^2 + z_m^2 \end{vmatrix}. \quad (6)$$

Записывая этот определитель как сумму двух определителей, каждый из которых имеет вид

$$\begin{vmatrix} x_k^2 & x_l y_l & x_l z_l + x_m z_m \\ x_k y_k & y_l^2 & y_l z_l + y_m z_m \\ x_k z_k & y_l z_l & z_l^2 + z_m^2 \end{vmatrix}, \quad (7)$$

и вычитая из третьего столбца второй, умноженный на z_l/y_l , окончательно получаем

$$\begin{vmatrix} x_k^2 & x_l y_l & x_m z_m \\ x_k y_k & y_l^2 & y_m z_m \\ x_k z_k & y_l z_l & z_m^2 \end{vmatrix}. \quad (8)$$

Суммируя по всем k, l и m , мы получаем левую часть равенства (3). Отметим, что описанная процедура привела к определителю вида (8), причем на каждом этапе преобразований мы

рассматривали один типичный определитель, общее число которых, как нетрудно видеть, равно 3×2 . Из этого следует, что дополнительно требуется множитель $1/3!$

§ 6. Диагонализация симметрической матрицы второго порядка. Для того чтобы показать, что произвольная вещественная симметрическая матрица может быть приведена к диагональной форме применением ортогонального преобразования, воспользуемся методом индукции, рассматривая первоначально случай 2×2 .

Пусть A — симметрическая матрица

$$A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a^1 & \\ & a^2 \end{vmatrix}, \quad (1)$$

и пусть λ_1 и x^1 — соответствующие собственное значение и собственный вектор. Это означает, что справедливы соотношения

$$Ax^1 = \lambda_1 x^1 \quad \text{или} \quad (a^1, x^1) = \lambda_1 x_{11}, \quad (a^2, x^1) = \lambda_1 x_{12}, \quad (2)$$

где x_{11} и x_{12} — компоненты вектора x^1 , которые будем считать нормированными условием $(x^1, x^1) = 1$.

Как мы знаем, можно построить ортогональную матрицу T_1 размеров 2×2 , одним из столбцов которой является вектор x^1 . Пусть другим столбцом будет вектор x^2 .

Мы хотим показать, что

$$T_1^t A T_1 = \begin{vmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{vmatrix}, \quad (3)$$

где λ_1 и λ_2 — не обязательно различные собственные значения матрицы A . При этом мы, конечно, сталкиваемся с трудностями, свойственными общему случаю. Для случая различных собственных значений мы уже привели простое доказательство.

Покажем первоначально, что

$$T_1^t A T_1 = \begin{vmatrix} \lambda_1 & b_{12} \\ 0 & b_{22} \end{vmatrix}, \quad (4)$$

где b_{12} и b_{22} — пока неизвестные параметры. Существенным обстоятельством является то, что элемент под главной диагональю равен нулю. Из уравнений (2) следует

$$T_1^t A T_1 = T_1^t \begin{vmatrix} \lambda_1 x_{11} & (a^1, x^2) \\ \lambda_1 x_{12} & (a^2, x^2) \end{vmatrix}. \quad (5)$$

Так как $T_1^t T_1 = I$, то, производя умножение, получаем *)

$$T_1^t \begin{vmatrix} \lambda_1 x_{11} & (a^1, x^2) \\ \lambda_1 x_{12} & (a^2, x^2) \end{vmatrix} = \begin{vmatrix} \lambda_1 & b_{12} \\ 0 & b_{22} \end{vmatrix}, \quad (6)$$

*) Здесь учтено, что $(x^1, x^1) = 1$ и $(x^2, x^2) = 0$. (Прим. ред.)

где b_{11} и b_{22} — параметры, значения которых будут определены ниже.

Начнем с параметра b_{11} и покажем, что он равен нулю. Это следует из того, что матрица $T_1'AT_1$ симметрическая:

$$(T_1'AT_1)' = T_1'A'(T_1)' = T_1'A'T_1 = T_1'AT_1. \quad (7)$$

Покажем, наконец, что величина b_{22} равна второму собственному значению матрицы A . Это следует из уже отмечавшегося факта, что собственные значения матрицы A совпадают с собственными значениями матрицы $T_1'AT_1$. Следовательно, $b_{22} = \lambda_2$.

Этим заканчивается доказательство двумерного случая. Этот случай существен не только для метода индукции. Он представляет для нас ценность и потому, что его изложение содержит все элементы общего доказательства.

§ 7. *N*-мерный случай. Применим метод индукции. Предположим, что для каждого $k, k = 1, 2, \dots, N$, мы можем определить ортогональную матрицу T_k , которая приводит вещественную симметрическую матрицу $A = \|a_{ij}\|, i, j = 1, 2, \dots, k$, к диагональной форме

$$T_k AT_k = \begin{vmatrix} \lambda_1 & & & & 0 \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & \lambda_k \end{vmatrix}. \quad (1)$$

Элементы главной диагонали λ_i должны быть тогда собственными значениями матрицы A . Сделав это предположение, которое, как мы знаем, справедливо для $N = 2$, покажем, что такое же приведение может быть выполнено для $k = N + 1$. Пусть

$$A_{N+1} = \|a_{ij}\| = \begin{vmatrix} a^1 \\ a^2 \\ \vdots \\ a^{N+1} \end{vmatrix}. \quad (2)$$

и пусть λ_1 и x^1 — соответствующие собственное значение и собственный вектор матрицы A_{N+1} .

Действуя так же, как и в двумерном случае, мы построим ортогональную матрицу T_1 , первый столбец которой есть x^1 . Пусть остальными столбцами будут x^2, x^3, \dots, x^{N+1} , так что матрица T_1 имеет вид

$$T_1 = \|x^1, x^2, \dots, x^{N+1}\|. \quad (3)$$

Тогда, как и ранее, первый шаг состоит в том, чтобы показать, что матрица $T_1' A_{N+1} T_1$ представима в виде

$$T_1' A_{N+1} T_1 = \begin{pmatrix} \lambda_1 & b_{12} & b_{13} & \dots & b_{1N} \\ 0 & & & & \\ 0 & & & & \\ \vdots & & & A_N & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix}. \quad (4)$$

при этом: а) все элементы первого столбца — нули, кроме первого элемента, равного λ_1 ;

б) величины $b_{12}, b_{13}, \dots, b_{1N}$ будут определены ниже;

в) матрица A_N имеет размеры $N \times N$.

Выполняя умножение, получаем

$$A_{N+1} T_1 = \begin{pmatrix} (a^1, x^1) & (a^1, x^2) & \dots & (a^1, x^{N+1}) \\ (a^2, x^1) & (a^2, x^2) & \dots & (a^2, x^{N+1}) \\ \vdots & \vdots & & \vdots \\ (a^{N+1}, x^1) & (a^{N+1}, x^2) & \dots & (a^{N+1}, x^{N+1}) \end{pmatrix} = \begin{pmatrix} \lambda_1 x_{11} & (a^1, x^2) & \dots & (a^1, x^{N+1}) \\ \lambda_1 x_{12} & (a^2, x^2) & \dots & (a^2, x^{N+1}) \\ \vdots & \vdots & & \vdots \\ \lambda_1 x_{1, N+1} & (a^{N+1}, x^2) & \dots & (a^{N+1}, x^{N+1}) \end{pmatrix}. \quad (5)$$

Так как T_1 — ортогональная матрица, то

$$T_1' A_{N+1} T_1 = \begin{pmatrix} \lambda_1 & b_{12} & \dots & b_{1, N+1} \\ 0 & & & \\ \vdots & & A_N & \\ \vdots & & & \\ 0 & & & \end{pmatrix}. \quad (6)$$

Для определения величин b_{1i} используем тот факт, что матрица $T_1' A_{N+1} T_1$ симметрическая.

Очевидно, что эти величины — нули и матрица A_N также симметрическая.

Тем самым доказано существование ортогональной матрицы T_1 такой, что

$$T_1' A_{N+1} T_1 = \begin{vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A_N & \\ \vdots & & & \\ 0 & & & \end{vmatrix}, \quad (7)$$

где A_N — симметрическая матрица.

Отметим, что собственными значениями матрицы A_N должны быть $\lambda_2, \lambda_3, \dots, \lambda_{N+1}$, остающиеся собственные значения матрицы A_{N+1} . Это следует из уже отмечавшейся тождественности собственных значений матриц A_{N+1} и $T_1' A_{N+1} T_1$ и из того обстоятельства, что характеристическое уравнение матрицы, стоящей в правой части равенства (6), имеет вид $(\lambda_1 - \lambda) |A_N - \lambda I| = 0$.

Применим теперь нашу индуктивную гипотезу. Пусть T_N — ортогональная матрица, приводящая A_N к диагональной форме. Построим матрицу $(N+1)$ -го порядка

$$S_{N+1} = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & T_N & \\ \vdots & & & \\ 0 & & & \end{vmatrix}, \quad (8)$$

которая, очевидно, также является ортогональной. Легко проверить, что

$$S_{N+1}' (T_1' A_{N+1} T_1) S_{N+1} = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_{N+1} \\ 0 & & & & \end{vmatrix}. \quad (9)$$

Так как можно записать

$$S_{N+1}' (T_1' A_{N+1} T_1) S_{N+1} = (T_1 S_{N+1})' A_{N+1} (T_1 S_{N+1}), \quad (10)$$

то $T_1 S_{N+1}$ — искомая ортогональная матрица, приводящая матрицу A_{N+1} к диагональной форме.

Таким образом, доказан следующий важный результат.

Теорема 2. *Вещественная симметрическая матрица A может быть приведена к диагональной форме с помощью некоторого ортогонального преобразования, или, что то же самое,*

существует ортогональная матрица T такая, что

$$T'AT = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix} \quad (11)$$

где λ_i —собственные значения матрицы A .

Это равносильно следующему результату:

$$(x, Ax) = \sum_{i=1}^N \lambda_i y_i^2, \quad (12)$$

где $y = T'x$.

Упражнение

1. Сохранятся ли результаты этого параграфа в случае, если A —комплексная симметрическая матрица?

§ 8. Необходимое и достаточное условие положительной определенности. Формула (7.12) сразу же приводит к следующей теореме:

Теорема 3. *Необходимым и достаточным условием положительной определенности матрицы A является положительность всех ее собственных значений.*

Аналогично мы видим, что необходимым и достаточным условием неотрицательной определенности матрицы A является неотрицательность всех ее собственных значений.

Упражнения

1. Матрица $A = BB'$ положительно определенная, если матрица B вещественна и $|B| \neq 0$.

2. Матрица $H = CC'$ положительно определенная, если $|C| \neq 0$.

3. Если A —вещественная симметрическая матрица, то матрица $I + \epsilon A$ положительно определенная при условии, что ϵ достаточно мало.

§ 9. Собственные векторы, соответствующие кратным собственным значениям. Если собственные значения матрицы A различны, то соответствующие собственные векторы линейно независимы, что следует из их ортогональности.

Рассмотрим теперь общий случай. Предположим, что λ_1 —корень кратности k . Верно ли, что существует k линейно независимых собственных векторов, соответствующих собственному значению λ_1 ? Если это так, то верно ли, что любой собственный вектор, соответствующий λ_1 , есть линейная комбинация этих k векторов?

Для того чтобы ответить на этот вопрос, воспользуемся представлением (7.11) и предположим, что $\lambda_1 = \lambda_2 = \dots = \lambda_k$, но что $\lambda_i \neq \lambda_1$ для $i = k+1, \dots, N$.

Умножая формулу (7.11) слева на матрицу T , запишем

$$AT = T \begin{vmatrix} \lambda_1 & & & & 0 \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & \lambda_N \end{vmatrix}. \quad (1)$$

откуда следует, что j -й столбец матрицы T есть собственный вектор матрицы A , соответствующий собственному значению λ_j . Так как матрица T ортогональная, то ее столбцы линейно независимы. Следовательно, если λ_1 — собственное значение кратности k , то ему соответствует k линейно независимых собственных векторов.

Остается показать, что любой другой собственный вектор y , соответствующий λ_1 , есть линейная комбинация этих k векторов. Запишем вектор y как линейную комбинацию N столбцов матрицы T :

$$y = \sum_{i=1}^N c_i x^i. \quad (2)$$

Коэффициенты c_i могут быть определены по правилу Крамера, так как определитель отличен от нуля вследствие линейной независимости векторов x^i .

Из ортогональности собственных векторов, соответствующих различным собственным значениям, следует, что $c_i = 0$, если только x^i — не собственный вектор, соответствующий λ_1 . Это доказывает, что вектор y есть линейная комбинация собственных векторов, соответствующих λ_1 .

§ 10. Теорема Гамильтона — Кэли для симметрических матриц. Воспользовавшись выражением (7.11), мы видим, что для любого полинома $p(\lambda)$ имеет место соотношение

$$p(A) = T \begin{vmatrix} p(\lambda_1) & & & & 0 \\ & p(\lambda_2) & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & p(\lambda_N) \end{vmatrix} T'. \quad (1)$$

Если, в частности, мы выберем $p(\lambda)$ в форме характеристического полинома матрицы A , $p(\lambda) = |A - \lambda I|$, то нетрудно видеть,

что $p(A) = 0$. Это дает доказательство знаменитого результата Кэли и Гамильтона для рассматриваемого частного случая.

Теорема 4. Любая симметрическая матрица удовлетворяет своему характеристическому уравнению.

Как мы увидим впоследствии, этот результат может быть распространен на произвольные квадратные матрицы.

Упражнение

1. Используя соображения непрерывности, получить теорему Гамильтона—Кэли для произвольной симметрической матрицы из результата, первоначально установленного для симметрической матрицы с простыми собственными значениями.

§ 11. Одновременное приведение к диагональной форме. Мы видели, что вещественная симметрическая матрица может быть приведена к диагональной форме посредством некоторого ортогонального преобразования. Естественно попытаться выяснить, возможно ли одновременное приведение к диагональной форме двух вещественных симметрических матриц. Ответ на этот вопрос дается следующей теоремой:

Теорема 5. Необходимым и достаточным условием существования ортогональной матрицы T такой, что

$$T'AT = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix}, \quad T'BT = \begin{vmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ 0 & & & \mu_N \end{vmatrix}, \quad (1)$$

является коммутативность матриц A и B .

Доказательство. Предположим, что матрица A имеет различные собственные значения. Тогда из того, что

$$Ax = \lambda x, \quad (2)$$

получаем

$$A(Bx) = B(Ax) = B(\lambda x) = \lambda(Bx). \quad (3)$$

Из этого соотношения следует, что Bx есть собственный вектор, соответствующий λ . Так как собственные значения по предположению различны, то любые два собственных вектора, соответствующих одному и тому же собственному значению, должны быть пропорциональны. Отсюда следует, что

$$Bx^i = \mu_i x^i, \quad i = 1, 2, \dots, N, \quad (4)$$

где μ_i —скаляры, которые должны быть собственными значениями матрицы B . Таким образом, мы видим, что матрицы A и B имеют одни и те же собственные векторы x^1, x^2, \dots, x^N .

Матрица T может быть взята в виде

$$T = \|x^1, x^2, \dots, x^N\|. \quad (5)$$

Рассмотрим теперь общий случай, когда λ_i — собственное значение кратности k с соответствующими собственными векторами x^1, x^2, \dots, x^k . Тогда из соотношения (3) мы получаем

$$Bx^i = \sum_{j=1}^k c_{ij}x^j, \quad i = 1, 2, \dots, k. \quad (6)$$

Посмотрим, возможно ли построить такую линейную комбинацию векторов x^i , которая была бы собственным вектором матрицы B .

Прежде всего отметим, что матрица $C = \|c_{ij}\|$ симметрическая, что следует из ортонормированности векторов x^i и симметричности B :

$$(x^j, Bx^i) = c_{ij} = (Bx^j, x^i) = c_{ji}. \quad (7)'$$

Рассмотрим линейную комбинацию $\sum_{i=1}^k a_i x^i$:

$$B\left(\sum_{i=1}^k a_i x^i\right) = \sum_{i=1}^k a_i \left(\sum_{j=1}^k c_{ij} x^j\right) = \sum_{j=1}^k \left(\sum_{i=1}^k c_{ij} a_i\right) x^j. \quad (8)$$

Следовательно, если a_i выбраны так, что

$$\sum_{i=1}^k c_{ij} a_i = r_1 a_j, \quad j = 1, 2, \dots, k, \quad (9)$$

то мы получаем

$$B\left(\sum_{i=1}^k a_i x^i\right) = r_1 \left(\sum_{i=1}^k a_i x^i\right), \quad (10)$$

откуда следует, что r_1 — собственное значение матрицы B и $\sum_{i=1}^k a_i x^i$ — соответствующий собственный вектор.

Соотношение (9) показывает, что r_1 — собственное значение матрицы C и a_i — компоненты соответствующего собственного вектора. Так что если T_k есть k -мерное ортогональное преобразование, приводящее матрицу C к диагональной форме, то векторы z^i , определяемые соотношением

$$\begin{pmatrix} z^1 \\ z^2 \\ \vdots \\ z^k \end{pmatrix} = T_k \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^k \end{pmatrix}, \quad (11)$$

образуют ортонормированный набор векторов, являющихся собственными векторами одновременно двух матриц A и B .

Выполняя аналогичные преобразования для собственных векторов, соответствующих каждому кратному собственному значению, мы получаем искомую матрицу T .

Необходимость условия следует из того факта, что две матрицы вида

$$A = T \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \\ 0 & & & & 0 \end{vmatrix} T', \quad B = T \begin{vmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_N \\ 0 & & & & 0 \end{vmatrix} T' \quad (12)$$

всегда коммутативны, если матрица T ортогональная.

§ 12. Одновременное приведение к сумме квадратов. Как отмечалось в предыдущем параграфе, одновременное приведение двух симметрических матриц A и B к диагональной форме посредством ортогонального преобразования возможно в том и только в том случае, когда матрицы коммутативны. Однако для многих целей достаточно привести матрицы A и B к диагональной форме посредством произвольного невырожденного преобразования. Мы хотим доказать следующую теорему:

Теорема 6. Если даны две вещественные симметрические матрицы A и B , причем матрица A положительно определена, то существует невырожденная матрица T такая, что

$$T'AT = I,$$

$$T'BT = \begin{vmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_N \\ 0 & & & & 0 \end{vmatrix} \quad (1)$$

Доказательство. Пусть S — ортогональная матрица такая, что

$$A = S \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \\ 0 & & & & 0 \end{vmatrix} S'. \quad (2)$$

Тогда если $x = Sy$, то мы получаем

$$(x, Ax) = \sum_{i=1}^N \lambda_i y_i^2, \quad (3)$$

$$(x, Bx) = (y, S'BSy).$$

Теперь выполним «растягивающее» преобразование

$$y_i = \frac{z_i}{\lambda_i^{1/2}}, \quad i = 1, 2, \dots, N, \quad (4)$$

так, что

$$y = S_1 z.$$

Тогда

$$\sum_{i=1}^N \lambda_i y_i^2 = (z, z) \quad (5)$$

и

$$(y, S'BSy) = (z, S_1' S' B S S_1 z).$$

Обозначим через C матрицу $S_1' S' B S S_1$, и пусть S_2 — ортогональная матрица, которая приводит матрицу C к диагональной форме. Если положить $z = S_2 w$, то в силу (5)

$$(z, Cz) = (w, S_2' C S_2 w) = \sum_{i=1}^N \mu_i w_i^2, \quad (6)$$

и в то же время

$$(z, z) = (w, w). \quad (6')$$

Сравнив (3), (5) и (6), мы заключаем, что

$$T = S S_2 S_1 \quad (7)$$

есть искомая невырожденная матрица.

Упражнения

1. Что можно сказать в связи с теоремой 6, если предположить, что матрица A является неотрицательно определенной?

2. Пусть обозначение $A > B$ для двух симметрических матриц означает, что матрица $A - B$ положительно определенная. Используя предыдущий результат, показать, что из неравенства $A > B > 0$ следует $B^{-1} > A^{-1}$.

§ 13. Эрмитовы матрицы. Очевидно, что тот же метод, который использовался для доказательства теоремы 2, дает возможность доказать теорему 7:

Теорема 7. Если H — эрмитова матрица, то существует унитарная матрица U такая, что

$$H = U \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix} U^*, \quad (1)$$

§ 14. Исходная проблема максимизации. Теперь мы в состоянии решить проблему, которая для нас послужила толчком к изучению положительности и отрицательности квадратичных форм, а именно, ответить на вопрос, когда стационарная точка функции $f(x_1, x_2, \dots, x_N)$ является точкой относительно минимума или относительно максимума.

Пусть $c = (c_1, c_2, \dots, c_N)$ — стационарная точка функции f , имеющей непрерывные смешанные производные второго порядка. Тогда достаточно рассмотреть квадратичную форму

$$Q = \left| \frac{\partial^2 f}{\partial x_i \partial x_j} \right| = |f_{c_i c_j}| \quad (1)$$

где, как и ранее, величины

$$\frac{\partial^2 f}{\partial c_i \partial c_j} = \frac{\partial^2 f}{\partial x_i \partial x_j} \quad (2)$$

вычислены в точке $x_1 = c_1, x_2 = c_2, \dots, x_N = c_N$.

Достаточным условием того, что c — точка относительного минимума, является положительная определенность формы Q , необходимым условием — неотрицательность формы Q . Критерий, приведенный в § 8, дает метод определения положительной и отрицательной определенности квадратичной формы Q . Однако если f — функция многих переменных и N велико, то этот метод практически неприменим. Далее, в гл. 6 мы получим наиболее удобный из критериев положительной определенности.

Упражнение

1. Показать, что условия

$$f_{c_1 c_1} < 0, \quad \begin{vmatrix} f_{c_1 c_1} & f_{c_1 c_2} \\ f_{c_2 c_1} & f_{c_2 c_2} \end{vmatrix} > 0$$

являются достаточными для того, чтобы функция $f(x_1, x_2)$ имела относительный максимум.

§ 15. Теория возмущений — I. Рассмотрим теперь проблему, представляющую большой теоретический и практический интерес. Пусть A — симметрическая матрица, имеющая собственные значения $\lambda_1, \lambda_2, \dots, \lambda_N$ и соответствующие собственные векторы x^1, x^2, \dots, x^N . Спрашивается, что можно сказать о собственных значениях и собственных векторах матрицы $A + \varepsilon B$, где B — симметрическая матрица, а ε — малый параметр? Так как мы интересуемся лишь формальной стороной теории, то мы не будем здесь рассматривать вопрос о том, каким должен быть параметр ε , чтобы название «малый» было действительно уместным.

Если матрицы A и B коммутативны, то они одновременно могут быть приведены к диагональной форме с помощью некоторого ортогонального преобразования. Следовательно, при со-

ответствующей перенумерации собственных значений матрицы B собственными значениями матрицы $A + \varepsilon B$ будут $\lambda_i + \varepsilon \mu_i$, $i = 1, 2, \dots, N$, в то время как собственные векторы останутся прежними.

Рассмотрим теперь общий случай, когда матрицы A и B некоммутативны, $AB \neq BA$. Для простоты остановимся только на случае различных собственных значений.

Вообще говоря, можно ожидать, что для достаточно малых ε собственные значения матрицы $A + \varepsilon B$ будут различными и близкими к собственным значениям матрицы A . Из этого следует, что собственные векторы матрицы $A + \varepsilon B$ также будут близки к собственным векторам матрицы A .

Один из способов доказательства этого состоит в следующем. Вместо характеристического уравнения $|A - \lambda I| = 0$ мы рассматриваем уравнение $|A + \varepsilon B - \lambda I| = 0$.

Проблема сводится к отысканию приближенного решения этого уравнения, причем необходимо учесть, что параметр ε мал и корни исходного уравнения известны. Пусть T — ортогональная матрица, приводящая матрицу A к диагональной форме. Тогда характеристическое уравнение может быть записано в виде

$$\begin{vmatrix} \lambda_1 + \varepsilon c_{11} - \lambda & \varepsilon c_{12} & \dots & \varepsilon c_{1N} \\ \varepsilon c_{21} & \lambda_2 + \varepsilon c_{22} - \lambda & \dots & \varepsilon c_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon c_{N1} & \varepsilon c_{N2} & \dots & \lambda_N + \varepsilon c_{NN} - \lambda \end{vmatrix} = 0,$$

где $T'BT = C = \|c_{ij}\|$.

Можно ожидать, что решение этого уравнения имеет вид

$$\lambda = \lambda_i + d_{1i}\varepsilon + d_{2i}\varepsilon^2 + \dots$$

Коэффициенты d_{1i} , d_{2i} и т. д. могут быть выражены через элементы c_{ij} ; мы оставляем задачу определения этих коэффициентов в качестве упражнения для читателя, имеющего опыт обращения с определителями.

Мы не останавливаемся подробно на этих вопросах, так как метод, приведенный в следующем параграфе, является более эффективным и, кроме того, может быть использован для изучения проблемы возмущений в случае операторов более общего вида.

§ 16. Теория возмущений — II. Предположим, что собственные значения и собственные векторы матриц $A + \varepsilon B$ и A не только близки, но и что эти величины могут быть представлены в виде рядов по степеням малого параметра ε .

Обозначим через μ_i и y^i собственные значения и собственные векторы матрицы $A + \varepsilon B$, а через λ_i и x^i — соответствующие величины для матрицы A .

Положим тогда

$$\begin{aligned} \mu_i &= \lambda_i + \varepsilon \lambda_{i1} + \varepsilon^2 \lambda_{i2} + \dots \\ y^i &= x^i + \varepsilon x^{i1} + \varepsilon^2 x^{i2} + \dots \end{aligned} \quad (1)$$

Чтобы определить неизвестные коэффициенты λ_{i1} , λ_{i2} , ... и неизвестные векторы x^{i1} , x^{i2} , ..., подставим эти выражения в уравнение

$$(A + \varepsilon B) y^i = \mu_i y^i \quad (2)$$

и приравняем коэффициенты при степенях ε .

Имеем

$$\begin{aligned} (A + \varepsilon B)(x^i + \varepsilon x^{i1} + \varepsilon^2 x^{i2} + \dots) &= \\ &= (\lambda_i I + \varepsilon \lambda_{i1} I + \varepsilon^2 \lambda_{i2} I + \dots)(x^i + \varepsilon x^{i1} + \varepsilon^2 x^{i2} + \dots). \end{aligned} \quad (3)$$

Отсюда следует, что

$$\begin{aligned} Ax^i &= \lambda_i x^i, \\ Ax^{i1} + Bx^i &= \lambda_i x^{i1} + \lambda_{i1} x^i, \\ Ax^{i2} + Bx^{i1} &= \lambda_i x^{i2} + \lambda_{i1} x^{i1} + \lambda_{i2} x^i, \\ &\dots \end{aligned} \quad (4)$$

Рассматривая эти уравнения, мы видим, что первое из них удовлетворяется тождественно, второе вводит две неизвестные величины: скаляр λ_{i1} и вектор x^{i1} . Аналогично, третье уравнение вводит еще два неизвестных: скаляр λ_{i2} и вектор x^{i2} .

На первый взгляд кажется, что рассматриваемый метод определения коэффициентов разложения (1) просто непригоден. Однако, исследуя второе уравнение, мы видим, что оно имеет форму

$$(A - \lambda_i I) x^{i1} = (\lambda_{i1} I - B) x^i, \quad (5)$$

где матрица коэффициентов $A - \lambda_i I$ вырожденная. Следовательно, это уравнение имеет решение, только если его правая часть $(\lambda_{i1} I - B) x^i$ обладает специальными свойствами. Только после того, как будет установлено, что $(\lambda_{i1} I - B) x^i$ обладает этими свойствами, мы сможем определить неизвестный скаляр λ_{i1} .

Чтобы упростить обозначения, запишем

$$\begin{aligned} x^{i1} &= y, \\ (\lambda_{i1} I - B) x^i &= z. \end{aligned} \quad (6)$$

Тогда задача формулируется следующим образом. Мы хотим определить, при каких ограничениях на вектор z уравнение

$$(A - \lambda_i I) y = z \quad (7)$$

имеет решение (λ_i — собственное значение матрицы A), и хотим найти это решение.

Пусть x^1, x^2, \dots, x^N — нормированные собственные векторы матрицы A . Представим y и z в виде линейных комбинаций этих векторов:

$$y = \sum_{j=1}^N b_j x^j, \quad (8)$$

$$z = \sum_{j=1}^N c_j x^j,$$

где

$$b_j = (y, x^j), \quad c_j = (z, x^j), \quad j = 1, 2, \dots, N. \quad (9)$$

Подставляя в (7), получаем уравнение

$$(A - \lambda_i I) \sum_{j=1}^N b_j x^j = \sum_{j=1}^N c_j x^j \quad (10)$$

или, так как $Ax^j = \lambda_j x^j$, $j = 1, 2, \dots, N$,

$$\sum_{j=1}^N b_j (\lambda_j - \lambda_i) x^j = \sum_{j=1}^N c_j x^j. \quad (11)$$

Приравнявая коэффициенты при x^j , получаем соотношение

$$b_j (\lambda_j - \lambda_i) = c_j, \quad j = 1, 2, \dots, N. \quad (12)$$

Эти уравнения совместны в том и только в том случае, когда

$$c_i = 0. \quad (13)$$

Тогда величина b_i произвольна, но остающиеся b_j определяются формулами

$$b_j = \frac{c_j}{\lambda_j - \lambda_i}, \quad j = 1, 2, \dots, N, \quad j \neq i. \quad (14)$$

Условие (13) означает, что уравнение (7) имеет решение в том и только в том случае, когда вектор z ортогонален собственному вектору x^i , соответствующему собственному значению λ_i .

Если это так, то существует однопараметрическое семейство решений уравнения (7), имеющее вид

$$y = b_i x^i + \sum_{j \neq i} \frac{c_j x^j}{\lambda_j - \lambda_i}, \quad (15)$$

где b_i произвольно.

Возвращаясь к (6), мы видим, что можно принять $b_i = 0$, так как различный выбор величин b_i влияет только на нормирование нового собственного вектора y^i .

Условие ортогональности дается соотношением

$$(x^i, (\lambda_{i1} - B)x^i) = 0. \quad (16)$$

Отсюда следует, что

$$\lambda_{i1} = (x^i, Bx^i). \quad (17)$$

Упражнения

1. Найти значение λ_{ij} .
2. Рассмотреть случай кратных корней, первоначально для матриц размеров 2×2 , а затем для матриц произвольных размеров.
3. Пусть λ_j — собственные значения матрицы A и $\lambda_j(z)$ — собственные значения матрицы $A + zB$. Показать, что

$$\lambda_j(A + zB) = \sum_{m=0}^{\infty} \lambda_j^{(m)} z^m, \quad \lambda_j^{(0)} = \lambda_j,$$

где

$$\lambda_j^{(m)} = (-1)^m m^{-1} \text{Sp} \left(\sum_{\substack{k_1 + \dots + k_m = m-1 \\ k_i \geq 0}} BS_j^{k_1} BS_j^{k_2} \dots BS_j^{k_m} \right), \quad S_j^{(0)} = -E_j.$$

Матрицы S_j определяются соотношением *) $A = \sum_j \lambda_j E_j$ и

$$S_j = \sum_{k \neq j} E_k / (\lambda_k - \lambda_j). \quad (\text{Като.})$$

4. Пусть A и B — матрицы размеров 10×10

$$A = \begin{vmatrix} 0 & 1 & \dots & 0 \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 1 \\ 0 & & & & 0 \end{vmatrix}, \quad B = \begin{vmatrix} 0 & 1 & \dots & 0 \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 1 \\ 10^{-10} & & & & 0 \end{vmatrix}.$$

Матрицы A и B отличаются лишь одним элементом $(10, 1)$, который для A равен 10^{-10} . Показать, что $|\lambda I - A| = \lambda^{10}$ и что $|\lambda I - B| = \lambda^{10} - 10^{-10}$. Следовательно, собственными значениями матрицы B являются 10^{-1} , $10^{-1}\omega$, ..., $10^{-1}\omega^9$, где ω — неприводимое значение корня десятой степени из единицы. Как можно объяснить это явление? (Дж. Форсайт.)

Упражнения к гл. 4

1. Показать, что вещественная симметрическая матрица может быть записана в форме

$$A = \sum_{i=1}^N \lambda_i E_i,$$

где E_i — неотрицательно определенные матрицы, удовлетворяющие условиям $E_i E_j = 0$, $i \neq j$; $E_i^2 = E_i$, и λ_i — собственные значения матрицы A . Такое представление называется спектральным разложением матрицы A .

*) По поводу обозначений см. упражнение 3 § 7 гл. 3, а также приводимое ниже упражнение 1 к настоящей главе. Предлагаемая в упражнении формула естественным образом выводится контурным интегрированием резольвенты матрицы $A + zB$ (по этому поводу см. упражнение 43 к гл. 6). (Прим. ред.)

2. Пусть A — вещественная кососимметрическая матрица. Как мы знаем, ее собственные значения или чисто мнимые или нули. Пусть $x + iy$ — собственный вектор, соответствующий $i\mu$, где μ — вещественная величина, отличная от нуля, а x и y вещественные. Показать, что векторы x и y ортогональны.

3. Пусть T_1 — некоторая ортогональная матрица, двумя столбцами которой являются x и y (см. предыдущее упражнение)

$$T_1 = (x, y, x^3, x^4, \dots, x^N).$$

Показать, что

$$T_1' A T_1 = \begin{vmatrix} \begin{pmatrix} 0 & \mu \\ -\mu & 0 \end{pmatrix} & 0 \\ 0 & A_{N-2} \end{vmatrix},$$

где матрица A_{N-2} снова кососимметрическая.

4. Доказать методом индукции, что если A — вещественная кососимметрическая матрица четного порядка, то можно найти вещественную ортогональную матрицу T такую, что

$$T' A T = \begin{vmatrix} \begin{pmatrix} 0 & \mu_1 \\ -\mu_1 & 0 \end{pmatrix} & & & & 0 \\ & \begin{pmatrix} 0 & \mu_2 \\ -\mu_2 & 0 \end{pmatrix} & & & \\ & & \ddots & & \\ 0 & & & & \begin{pmatrix} 0 & \mu_N \\ -\mu_N & 0 \end{pmatrix} \\ & & & & & 0 \end{vmatrix},$$

где некоторые μ_i могут быть нулями.

5. Показать, что если A — вещественная кососимметрическая матрица нечетного порядка, то ее каноническая форма имеет вид

$$T' A T = \begin{vmatrix} \begin{pmatrix} 0 & \mu_1 \\ -\mu_1 & 0 \end{pmatrix} & & & & 0 \\ & \ddots & & & \\ & & 0 & & \\ & & & & \begin{pmatrix} 0 & \mu_N \\ -\mu_N & 0 \end{pmatrix} \\ & & & & & 0 \end{vmatrix},$$

где некоторые μ_i могут быть нулями.

6. Показать, что определитель кососимметрической матрицы нечетного порядка равен нулю.

7. Определитель кососимметрической матрицы четного порядка есть квадрат полинома относительно элементов этой матрицы.

8. Пусть A — вещественная ортогональная матрица и λ — собственное значение, по абсолютной величине равное 1, но не равное ± 1 . Пусть соответствующий собственный вектор имеет вид $x + iy$, где x и y вещественные. Показать, что векторы x и y ортогональны.

где все неуказанные элементы — нули, показать, что

$$|\lambda I - H| = \prod_{k=1}^{N-1} \left(\lambda - 2 - 2 \cos \frac{2k\pi}{2N-1} \right),$$

и определить собственные векторы.

17. Используя последний результат, показать, что если $x_1 = 0$ и x_i вещественны, то

$$\sum_{i=1}^{N-1} (x_i - x_{i+1})^2 \geq 4 \sin^2 \frac{\pi}{2(2N-1)} \sum_{i=2}^N x_i^2.$$

Определить, когда имеет место равенство.

18. Показать, что если все x_i вещественны и $x_0 = x_{N+1} = 0$, то

$$\sum_{i=0}^N (x_i - x_{i+1})^2 \geq 4 \sin^2 \frac{\pi}{2(N+1)} \sum_{i=1}^N x_i^2,$$

и что если $x_1 = x_{N+1}$, $\sum_{i=1}^N x_i = 0$, то

$$\sum_{i=1}^N (x_i - x_{i+1})^2 \geq 4 \sin^2 \frac{\pi}{N} \sum_{i=1}^N x_i^2. \quad (\text{Фань Цзы, Таусски, Тодд.})$$

19. Найти собственные значения и собственные векторы матрицы, соответствующей квадратичной форме $\sum_{i,j=1}^N (x_i - x_j)^2$.

20. Пусть A и B — вещественные симметрические матрицы такие, что A неотрицательно определенная. Тогда, если $|A + iB| = 0$, существует нетривиальный вещественный вектор x такой, что $(A + iB)x = 0$. (Переманс — Дюпарк — Леккеркернер.)

21. Если A и B — симметрические матрицы и A неотрицательно определенная, то собственные значения матрицы AB вещественны.

22. Если A и B — симметрические матрицы, то собственные значения матрицы $AB - BA$ чисто мнимые.

23. Привести квадратичскую форму $Q_N(x) = x_1x_2 + x_2x_3 + \dots + x_{N-1}x_N$ к диагональной форме, определяя собственные векторы и собственные значения. Подобным же образом привести к диагональной форме $P_N(x) = x_1x_2 + x_2x_3 + \dots + x_{N-1}x_N + x_Nx_1$.

24. Если A — комплексная квадратная матрица, то A можно представить в виде HU , где H — неотрицательно определенная эрмитова матрица и U — унитарная матрица.

Дальнейшие исследования и ссылки на более ранние результаты Аутона, Уинтнера, Мурнагана можно найти у Вилльямсона¹⁾.

25. Если матрицы A , B и C симметрические и положительно определенные, то корни уравнения

$$|\lambda^2 A + \lambda B + C| = 0$$

имеют отрицательные вещественные части. (Пароди.)

26. Пусть A — квадратная матрица. Положим $|A| = |A|_+ - |A|_-$, где $|A|_+$ — сумма тех членов разложения определителя, которые снабжаются

¹⁾ D. Williamson, A. Generalization of the Polar Representation of Nonsingular Matrices, Bull. Amer. Math. Soc. 48 (1942), 856—863. [См. также Ф. Р. Гантмахер, Теория матриц, «Наука», 1966, стр. 249. (Прим. ред.)]

знаком плюс. Показать, что если A — положительно определенная матрица, то $|A| > 0$. (Шур — Ленг.)

27. Пусть A — вещественная матрица такая, что существует положительно определенная матрица M , связанная с матрицей A соотношением $A'M = MA$. Матрица A , обладающая таким свойством, называется *симметризуемой*. Показать, что матрица A обладает следующими свойствами:

а) все ее собственные значения вещественны,
 б) в терминах обобщенного скалярного произведения $[x, y] = (x, My)$ собственные векторы, соответствующие различным собственным значениям, ортогональны;

в) пусть матрица T называется «ортогональной», если она вещественна и ее столбцы ортогональны в обобщенном смысле. Тогда существует «ортогональная» матрица T такая, что матрица TAT' диагональна¹⁾.

28. Аналогичным образом обобщить понятие эрмитовой матрицы.

29. Назовем матрицу A *идемпотентной*, если $A^2 = A$. Показать, что необходимо и достаточное условие того, что симметрическая матрица A ранга k идемпотентна, состоит в том, что k собственных значений равны единице, а остальные $N - k$ равны нулю. (Понятие ранга определено в приложении А.)

30. Если A — симметрическая идемпотентная матрица, то ранг матрицы A равен ее следу*).

31. Невырожденная симметрическая идемпотентная матрица тождественна единичной матрице.

32. Если A — симметрическая идемпотентная матрица порядка N и ранга k , то матрица A будет положительно определенной, если $k = N$, и неотрицательно определенной, если $k < N$.

33. Если A — симметрическая идемпотентная матрица с элементом $a_{ii} = 0$ для некоторого значения i , то все элементы i -й строки и i -го столбца равны нулю.

34. Если каждая матрица A_i симметрическая и $\sum_{i=1}^m A_i = I$, то следующие

три условия эквивалентны:

а) каждая матрица A_i идемпотентная;

б) $A_i A_j = 0$, $i \neq j$;

в) $\sum_i n_i = N$, где n_i — ранг матрицы A_i и N — ее порядок.

35. Пусть A_i — набор симметрических матриц порядка N , и ранг матрицы A_i равен p_i . Пусть матрица $A = \sum_i A_i$ имеет ранг p .

Рассмотрим четыре условия:

C1. каждая матрица A_i идемпотентная,

C2. $A_i A_j = 0$, $i \neq j$,

C3. матрица A идемпотентная,

C4. $p = \sum_{i=1}^m p_i$.

Тогда

1) любые два из трех условий C1, C2, C3 означают выполнение всех условий C1, C2, C3, C4.

2) условия C3 и C4 означают выполнение C1 и C2.

¹⁾ А. Н. Колмогоров, Zur Theorie der Markoffschen Ketten, Math. Ann. 112 (1936), 155—160.

* По поводу следа матрицы см. § 8 гл. 6, стр. 118. (Прим. ред.)

Доказательство вышеприведенных фактов и некоторые их применения можно найти в работе Грейбила и Марсалья¹⁾.

36. Используя то обстоятельство, что для двух вещественных чисел x_1 и x_2 и любого $\varepsilon > 0$ имеет место неравенство $2x_1x_2 < \frac{x_1^2}{\varepsilon} + \varepsilon x_2^2$, показать, что квадратичная форма

$$\lambda \sum_{i=1}^k b_i x_i^2 + \sum_{i=k+1}^N b_i x_i^2 + \sum_{j=1}^k \sum_{i=k+1}^N a_{ij} x_i x_j + \sum_{i,j=1}^k a_{ij} x_i x_j$$

будет положительно определенной, если $b_i > 0$ и параметр λ достаточно велик.

37. Показать, что из положительной определенности формы $\sum_{i,j=1}^N a_{ij} x_i x_j$ не следует при $N > 3$, что форма $\sum_{i=1}^N a_{ii} x_i^2 + \sum_{i \neq j} |a_{ij}| x_i x_j$ положительно определена^{*)}.

38. Матрица $A = \|a_{ij}\|$, $i, j = 1, 2, 3, 4$, называется *матрицей Лоренца*, если преобразование $x = Ay$ оставляет квадратичную форму $Q(x) = x_1^2 - x_2^2 - x_3^2 - x_4^2$ неизменной, т. е. $Q(x) = Q(y)$. Показать, что произведение двух таких матриц снова матрица Лоренца.

39. Показать, что

$$y_1 = (x_1 + \beta x_2)/(1 - \beta^2)^{1/2}, \quad y_2 = (\beta x_1 + x_2)/(1 - \beta^2)^{1/2}, \quad y_3 = x_3, \quad y_4 = x_4,$$

где $0 < \beta^2 < 1$, есть преобразование Лоренца.

40. Показать, что любое преобразование Лоренца есть комбинация ортогонального преобразования переменных x_2, x_3, x_4 , оставляющего x_1 неизменным, преобразования типа, приведенного выше, и преобразования, состоящего в возможном изменении знака одной из переменных (отражения²⁾). Физическую интерпретацию преобразований Лоренца можно найти в книге Сайнджа³⁾.

41. Пусть H — неотрицательно определенная эрмитова матрица. Показать, что для любого заданного интервала $[a, b]$ можно найти последовательность комплексных функций $\{f_i(t)\}$, $i = 1, 2, \dots$, такую, что

$$h_{ij} = \int_a^b f_i(t) \overline{f_j(t)} dt, \quad i, j = 1, 2, \dots^4)$$

¹⁾ F. A. Graybill and G. Marsaglia, Idempotent Matrices and Quadratic Forms in the General Linear Hypothesis, Ann. Math. Stat. 28 (1957), 678—686.

^{*)} Если $a_{ij} < 0$ при $i \neq j$, то из положительной определенности формы $\sum_{i,j=1}^n a_{ij} x_i x_j$ следует положительная определенность формы $\sum_{i,j=1}^n |a_{ij}| x_i x_j$.
(Прим. ред.)

²⁾ Петровский И. Г., Лекции об уравнениях с частными производными, Гостехиздат, 1953.

³⁾ J. L. Synge, Relativity, the Special Theory, Interscience Publishers, Inc., New York, 1956.

⁴⁾ Шур (I. Schur), Math. Z. 1 (1918), 206.

типа или доказательство того, что она существует, является не-тривиальной задачей.

По этому поводу смотри

Харди, Литтлвуд, Поля (G. H. Hardy, J. E. Littlewood and G. Polya), *Inequalities*, Cambridge University Press, New York, 1934, 57—60. [Русский перевод: Неравенства, ИЛ, 1948.]

См. также соответствующие комментарии в конце гл. 5.

§ 8. Понятие о положительно определенной квадратичной форме как естественное обобщение понятия положительного скаляра — одно из наиболее полезных и плодотворных во всей математике. Статья

Фань Цзы (K у F а n), On Positive Definite Sequences, *Ann. Math.* 47 (1946), 593—607

указывает некоторые из многих возможных путей использования этого понятия в анализе.

Подробное и изящное изложение других применений имеется в монографии

Фань Цзы (K у F а n), *Les Fonctions definies-positives et les Fonctions completement monotones*, fascicule CXIV, *Mem. sci. math.*, 1950.

В приложениях в конце книги мы также укажем остроумные приемы использования квадратичных форм в различных разделах анализа.

Вопрос о диагонализации комплексных неэрмитовых симметрических матриц рассматривается в работе

Долф, Маклафлин и Маркс (C. L. Dolph, J. E. McLaughlin and I. Marx), *Symmetric Linear Transformations and Complex Quadratic Form*, *Comm. Pure and Appl. Math.* 7 (1954), 621—632.

Вопросы этого характера возникают как частные случаи более общей проблемы, касающейся собственных значений и собственных функций уравнения Штурма — Лиувилля с комплексными коэффициентами.

§ 15. Дальнейшее обсуждение этих проблем и ссылки можно найти в статье

Браун и Бассет (R. D. Brown, I. M. Bassett), *A Method for Calculating the First Order Perturbation of an Eigenvalue of a Finite Matrix*, *Proc. Phys. Soc.* 71 (1958), 724—732.

Некоторые интересные обобщения понятия положительной определенности имеются в работе

М. Г. Крейн, Об одном применении теоремы о неподвижной точке в теории линейных операторов в пространстве с индефинитной метрикой, *УМН* 5, № 2 (1950), 180—199,

где можно найти дальнейшие ссылки.

§ 16. В статье, посвященной физическим приложениям,

Лакс (M. Lax), Localized Perturbations, Phys. Rev. 94 (1954), 1391 рассматривается решение уравнения $(A + B)x = \lambda x$, где только несколько элементов матрицы B отличны от нуля.

Некоторые интересные результаты, касающиеся кососимметрических матриц, имеются в работах

Дрезин (M. P. Drazin), A Note on Skew-symmetric Matrices, Math. Gaz. 36 (1952), 253—255;

Джекобсон (N. Jacobson), Bull. Math. Soc. 45 (1939), 745—748.

В связи с упражнением 24 см. работу

Стенцель (H. Stenzel), Über die Darstellbarkeit einer Matrix als Product ..., Math. Z. 15, 1—25.

В заключение укажем, что некоторые интересные соотношения между ортогональностью и квадратичными формами можно найти в работах

Гребнер (W. Groebner), Über die Konstruktion von Systemen orthogonaler Polynome in ein- und zweidimensionalem Bereich. Monatsh. Math. 52 (1948), 38—54;

Ларчер (H. Larcher), Proc. Amer. Math. Soc. 10 (1959), 417—423.

УСЛОВНЫЕ ЭКСТРЕМУМЫ

§ 1. Введение. В предыдущих главах мы рассмотрели задачу об определении области значений, принимаемых квадратичной формой (x, Ax) на множестве векторов x таких, что $(x, x) = 1$. В частности, нас интересовало, может ли форма (x, Ax) принимать как положительные, так и отрицательные значения. В этой главе мы исследуем тот же вопрос при условии, что x удовлетворяет не только соотношению $(x, x) = 1$, но и некоторым дополнительным линейным ограничениям вида

$$(x, b^i) = c_i, \quad i = 1, 2, \dots, k. \quad (1)$$

В переводе на язык геометрии мы рассматривали множество значений, принимаемых квадратичной формой (x, Ax) на единичной сфере. В дополнение к этому условию мы теперь полагаем, что x одновременно принадлежит еще и нескольким плоскостям, или, иначе, является точкой заданной $(N - k)$ -мерной плоскости.

Ограничения такого рода самым естественным образом возникают в различных задачах алгебры, анализа и геометрии.

Прежде всего мы обобщим применявшийся в гл. 1 алгебраический метод, получив на его основе несколько полезных необходимых и достаточных условий положительной определенности. Эти результаты будут в свою очередь расширены в ходе изложения.

§ 2. Детерминантный критерий положительной определенности (критерий Сильвестра). В гл. 4 было показано, что для положительной определенности вещественной симметрической матрицы необходимо и достаточно, чтобы все ее собственные значения были положительны. Однако, хотя этот результат и имеет теоретическую ценность, его практическое применение довольно затруднительно. Поэтому необходимо получить условия, более пригодные для аналитических и вычислительных целей.

Критерий, сформулированный на языке собственных значений, не годится для приложений по той причине, что численное нахождение собственных значений матрицы высокого порядка

оказывается весьма трудным делом. Всякая попытка разложить определитель $|A - \lambda I|$ неизбежно обречена на неудачу вследствие чрезвычайно большого количества членов, присутствующих в таком разложении. Полное разложение определителя порядка N содержит $N!$ членов. Так как $10! = 3\,628\,800$, а $20! \cong 2433 \cdot 10^{18}$, то ясно, что прямые методы применить нельзя, даже имея в своем распоряжении самые совершенные вычислительные машины¹⁾.

Численные методы отыскания собственных значений и собственных векторов представляют собой один из наиболее важных разделов теории матриц. Как уже отмечалось ранее, мы не будем здесь затрагивать каких-либо сторон этой темы. Одна из последующих книг нашей серии будет посвящена исключительно этим вопросам.

Для квадратичной формы двух переменных

$$Q(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \quad (1)$$

мы получили представление

$$Q = a_{11} \left(x_1 + \frac{a_{12}}{a_{11}} x_2 \right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}} \right) x_2^2 \quad (2)$$

в предположении, что $a_{11} \neq 0$.

Отсюда мы заключили, что выполнение неравенств

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0 \quad (3)$$

является необходимым и достаточным условием положительной определенности матрицы A .

Рассмотрим теперь квадратичную форму трех переменных

$$Q(x_1, x_2, x_3) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_{23}x_2x_3 + 2a_{13}x_1x_3 + a_{33}x_3^2. \quad (4)$$

Так как условие $a_{11} > 0$ является, очевидно, необходимым для положительной определенности, в чем мы убеждаемся, положив $x_2 = x_3 = 0$, то можно написать

$$Q(x_1, x_2, x_3) = a_{11} \left(x_1 + \frac{a_{12}x_2}{a_{11}} + \frac{a_{13}x_3}{a_{11}} \right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}} \right) x_2^2 + 2 \left(a_{23} - \frac{a_{12}a_{13}}{a_{11}} \right) x_2x_3 + \left(a_{33} - \frac{a_{13}^2}{a_{11}} \right) x_3^2. \quad (5)$$

Если $Q(x_1, x_2, x_3)$ — положительно определенная квадратичная форма, то, положив

$$x_1 + \frac{a_{12}x_2 + a_{13}x_3}{a_{11}} = 0,$$

¹⁾ При скорости вычислений в одну операцию за микросекунду выполнение 20! операций потребовало бы 500 000 лет!

мы получим квадратичную форму переменных x_1 и x_2 , которая также должна быть положительно определенной. Отсюда, используя выведенные нами условия (3) для матриц второго порядка, мы получаем необходимые и достаточные условия положительной определенности квадратичной формы $Q(x_1, x_2, x_3)$:

$$a_{11} > 0, \quad a_{22} - \frac{a_{12}^2}{a_{11}} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0. \quad (6)$$

Первые два условия нам уже знакомы. Преобразуем теперь третье условие к более удобному виду. Рассмотрим определитель

$$D_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} - \frac{a_{12}^2}{a_{11}} & a_{23} - \frac{a_{12}a_{13}}{a_{11}} \\ 0 & a_{32} - \frac{a_{12}a_{13}}{a_{11}} & a_{33} - \frac{a_{13}^2}{a_{11}} \end{vmatrix}. \quad (7)$$

Чтобы получить последнее равенство, нужно вычесть из второй строки определителя D_3 первую, умноженную на $\frac{a_{12}}{a_{11}}$, а из третьей строки — первую, умноженную на $\frac{a_{13}}{a_{11}}$ (прием, ранее применявшийся в § 2 гл. 3).

Теперь условия (6) можно записать в следующей, наводящей на размышления форме:

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0. \quad (8)$$

Таким образом, у нас есть все необходимое для доказательства очень важного результата, содержащегося в теореме 1.

Теорема 1. *Необходимым и достаточным условием положительной определенности вещественной симметрической матрицы A является выполнение следующих условий:*

$$D_k > 0, \quad k = 1, 2, \dots, N, \quad (9)$$

где

$$D_k = |a_{ij}|, \quad i, j = 1, 2, \dots, k. \quad (10)$$

Детальное проведение доказательства мы оставляем читателю в качестве упражнения.

§ 3. Представление в виде суммы квадратов. Несколько продолжив наше рассмотрение, мы можем установить следующий результат.

Теорема 2. Если ни один из определителей D_k не равен нулю, то квадратичная форма N переменных x_1, x_2, \dots, x_N может быть представлена в виде

$$Q(x_1, x_2, \dots, x_N) = \sum_{k=1}^N \left(\frac{D_k}{D_{k-1}} \right) y_k^2, \quad D_0 = 1, \quad (1)$$

где

$$y_k = x_k + \sum_{j=k+1}^N c_{kj} x_j, \quad k=1, 2, \dots, N-1, \quad (2)$$

$$y_N = x_N,$$

а c_{ij} — рациональные функции элементов a_{ij} матрицы A .

Формула (1) представляет собой обобщение представления (2.5) на N -мерный случай.

Упражнение

1. К чему приведет равенство нулю одного или нескольких определителей D_k ?

§ 4. Связанные вариации и теоремы Финслера. Рассмотрим теперь вопрос о знакоопределенности квадратичной формы в случае, когда переменные x_i могут принимать лишь те значения, которые удовлетворяют соотношениям

$$\sum_{j=1}^N b_{ij} x_j = 0, \quad i=1, 2, \dots, k, \quad k < N. \quad (1)$$

Как обычно, все встречающиеся нам величины предполагаются вещественными. Не уменьшая общности, мы можем считать, что все k уравнений в (1) независимы. Тогда, используя эти уравнения, можно выразить k переменных x_i через остальные $N-k$ переменных в виде линейных функций, подставить затем полученные соотношения в выражение для Q и решать уже $(N-k)$ -мерную задачу, используя выведенные в предыдущих параграфах критерии.

Однако, хотя такой метод и может быть успешно применен, он связан с большим числом операций над определителями. Поэтому мы изберем другой путь. Все же мы рекомендуем читателю попытаться получить результат только что описанным способом, хотя бы для случая, когда число условий $k=1$, чтобы должным образом оценить другой подход, к изложению которого мы сейчас переходим.

Начнем с доказательства следующего результата Финслера.

Теорема 3. Если $(x, Ax) > 0$ для всех x , удовлетворяющих условию $(x, Bx) = 0$, где B — неотрицательно определенная мат-

рица, то существует такое число λ , что квадратичная форма $(x, Ax) + \lambda(x, Bx)$ положительно определена.

Доказательство. Сделаем линейную замену переменных $x = Ty$, где T — ортогональная матрица такая, что

$$(x, Bx) = \sum_{i=1}^k \mu_i y_i^2, \quad \mu_i > 0, \quad 1 \leq k < N. \quad (2)$$

Так как B по предположению — неотрицательно определенная матрица, то форма (x, Bx) , как мы знаем, может быть приведена к виду (2). Это же преобразование, примененное к форме (x, Ax) , дает

$$(x, Ax) = \sum_{i,j=1}^N c_{ij} y_i y_j. \quad (3)$$

Если квадратичная форма (x, Ax) положительно при всех x таких, что $(x, Bx) = \sum_{i=1}^k \mu_i y_i^2 = 0$, то неравенство

$$\sum_{i,j=k+1}^N c_{ij} y_i y_j > 0 \quad (4)$$

должно выполняться для всех нетривиальных наборов значений переменных $y_{k+1}, y_{k+2}, \dots, y_N$.

В $(N-k)$ -мерном пространстве векторов $(y_{k+1}, y_{k+2}, \dots, y_N)$ произведем ортогональное преобразование, которое приводит квадратичную форму (4) к сумме квадратов. Это преобразование не затрагивает переменных y_1, y_2, \dots, y_k . В N -мерном пространстве векторов y запишем это преобразование в виде $y = Sw$. Тогда

$$\begin{aligned} (x, Ax) = Q(w) &= \sum_{i,j=1}^N c_{ij} y_i y_j = \\ &= \sum_{i=k+1}^N \mu_i w_i^2 + \sum_{i=k+1}^N \sum_{j=1}^N d_{ij} w_i w_j + \sum_{i,j=1}^k d_{ij} w_i w_j, \end{aligned} \quad (5)$$

и теперь наша задача сводится к доказательству того, что квадратичная форма

$$\lambda \sum_{i=1}^k \mu_i w_i^2 + Q(w) \quad (6)$$

является положительно определенной, если $\mu_i > 0$, $i = 1, 2, \dots, N$, и λ достаточно велико. Для завершения доказательства остается только применить результат, сформулированный в упражнении 36 к гл. 4.

Возвращаясь к исходной вариационной задаче, заметим, что k уравнений (1) можно объединить в одно уравнение

$$\sum_{i=1}^k \left(\sum_{j=1}^N b_{ij} x_j \right)^2 = 0. \quad (7)$$

Рассматривая $\sum_{i=1}^k \left(\sum_{j=1}^N b_{ij} x_j \right)^2$ как неотрицательно определенную квадратичную форму, легко убедиться, что из теоремы 3 вытекает следующий результат.

Для того чтобы квадратичная форма $Q(x_1, x_2, \dots, x_N)$ была положительной при всех нетривиальных наборах переменных x_i , удовлетворяющих линейным уравнениям (1), необходимо и достаточно, чтобы квадратичная форма

$$P(x_1, x_2, \dots, x_N) = Q(x_1, x_2, \dots, x_N) + \lambda \sum_{i=1}^k \left(\sum_{j=1}^N b_{ij} x_j \right)^2 \quad (8)$$

была положительно определенной для всех достаточно больших положительных λ .

Упущение

1. Установить предыдущий результат, рассматривая максимум по всем переменным x_i функции

$$f(x) = - \frac{Q(x_1, x_2, \dots, x_N)}{\sum_{i=1}^k \left(\sum_{j=1}^N b_{ij} x_j \right)^2}. \quad (\text{Герштейн.})$$

§ 5. Случай $k = 1$. Попытаемся вывести из предыдущего результата какой-либо удобный критерий положительной определенности, причем начнем с часто встречающегося случая $k = 1$. Тогда квадратичная форма P имеет вид

$$P = \sum_{i,j=1}^N (a_{ij} + \lambda b_{1i} b_{1j}) x_i x_j. \quad (1)$$

В соответствии с теоремой 1 для того, чтобы P была положительно определенной при любом достаточно большом положительном λ , должны выполняться детерминантные неравенства

$$|a_{ij} + \lambda b_{1i} b_{1j}| > 0, \quad i, j = 1, 2, \dots, k, \quad (2)$$

при $k = 1, 2, \dots, N$.

Чтобы получить более простой эквивалент этих неравенств, рассмотрим следующее матричное тождество:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & \lambda b_1 & 1 & 0 \\ a_{21} & a_{22} & \dots & a_{2k} & \lambda b_2 & \cdot & 1 & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} & \lambda b_k & 0 & \cdot & 1 & 0 \\ b_1 & b_2 & \dots & b_k & -1 & b_1 & b_2 & \dots & b_k & 1 \end{vmatrix} = \begin{vmatrix} a_{11} + \lambda b_1^2 & a_{12} + \lambda b_1 b_2 & \dots & a_{1k} + \lambda b_1 b_k & \lambda b_1 \\ a_{21} + \lambda b_2 b_1 & a_{22} + \lambda b_2^2 & \dots & a_{2k} + \lambda b_2 b_k & \lambda b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} + \frac{\lambda b_k b_1}{0} & a_{k2} + \frac{\lambda b_k b_2}{0} & \dots & a_{kk} + \frac{\lambda b_k^2}{0} & \lambda b_k \\ \cdot & \cdot & \cdot & \cdot & -1 \end{vmatrix} \quad (3)$$

Так как равенство в (3) сохраняется при замене матриц их определителями, то положительность определителя $|a_{ij} + \lambda b_i b_{ij}|$, $i, j = 1, 2, \dots, k$, при всех больших λ эквивалентна отрицательности определителя

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & \lambda b_1 \\ a_{21} & a_{22} & \dots & a_{2k} & \lambda b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} & \lambda b_k \\ b_1 & b_2 & \dots & b_k & -1 \end{vmatrix} \quad (4)$$

при больших положительных λ . Следовательно, должны выполняться неравенства

$$\lambda \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & b_1 \\ a_{21} & a_{22} & \dots & a_{2k} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} & b_k \\ b_1 & b_2 & \dots & b_k & 0 \end{vmatrix} - \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix} < 0 \quad (5)$$

при всех больших положительных λ и $k = 1, 2, \dots, N$.

Таким образом, достаточное условие положительности Q при всех нетривиальных наборах значений переменных x_i , удовлетворяющих линейному уравнению

$$b_1 x_1 + b_2 x_2 + \dots + b_k x_k = 0,$$

заключается в том, что окаймленные определители удовлетворяют неравенствам

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & b_1 \\ a_{21} & a_{22} & \dots & a_{2k} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} & b_k \\ b_1 & b_2 & \dots & b_k & 0 \end{vmatrix} < 0, \quad (6)$$

$k = 1, 2, \dots, N$. Ясно также, что условия

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & b_1 \\ a_{21} & a_{22} & \dots & a_{2k} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} & b_k \\ b_1 & b_2 & \dots & b_k & 0 \end{vmatrix} \leq 0, \quad (7)$$

$k = 1, 2, \dots, N$, являются необходимыми, при этом некоторые из этих определителей могут быть равными нулю, не нарушая тем самым положительной определенности формы Q . В простейшем случае такая ситуация может иметь место, когда некоторые из величин b_i равны нулю.

Обращаясь к (5), мы видим, что если какой-либо из окаймленных определителей, скажем k -й, равен нулю, то необходимым условием положительной определенности является выполнение неравенства

$$\begin{vmatrix} a_{11} & \dots & a_{1k} \\ a_{21} & \dots & a_{2k} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_{k1} & \dots & a_{kk} \end{vmatrix} > 0. \quad (8)$$

Дальнейшее обсуждение некоторых частных случаев содержится в приведенных ниже упражнениях.

Упражнения

1. В предположении, что $b_N \neq 0$, получить предыдущий результат, выражая x_N через остальные переменные:

$$x_N = -(b_1 x_1 + b_2 x_2 + \dots + b_{N-1} x_{N-1})/b_N.$$

и рассматривая $Q(x)$ как квадратичную форму $N-1$ переменных x_1, x_2, \dots, x_{N-1} . Можно ли здесь применить индуктивный подход?

выполняется для любого вектора x . Это неравенство можно преобразовать к виду

$$(x, x) \geq - \frac{\begin{vmatrix} 0 & b_1 & b_2 & \dots & b_k \\ b_1 & & & & \\ \vdots & & & & \\ & & & (a^i, a^j) & \\ \vdots & & & & \\ b_k & & & & \end{vmatrix}}{|(a^i, a^j)|}, \quad (3)$$

так как определитель $|(a^i, a^j)|$ положителен вследствие линейной независимости векторов a^i .

Величина, стоящая в правой части неравенства (3), и является минимальным значением формы (x, x) при наложенных ограничениях, так как равенство в формуле (2) достигается, если выбрать x равным некоторой линейной комбинации векторов a^i :

$$x = \sum_{i=1}^k c_i a^i. \quad (4)$$

Чтобы удовлетворить уравнениям (1), коэффициенты c_i должны образовывать решение линейной системы уравнений

$$\sum_{i=1}^k c_i (a^i, a^j) = b_j. \quad (5)$$

Так как определитель $|(a^i, a^j)| \neq 0$, то эта система имеет единственное решение, которое и дает нам единственный минимизирующий вектор x .

Упражнения

1. Используя предыдущий результат, показать, что квадратичная форма

$$Q(x) = \begin{vmatrix} 0 & x_1 & x_2 & \dots & x_N \\ x_1 & & & & \\ \vdots & & & & \\ & & & A & \\ \vdots & & & & \\ x_N & & & & \end{vmatrix}$$

является отрицательно определенной, если матрица A положительно определена.

2. Дать независимое доказательство результата, предлагаемого в упражнении 1, показав, что соответствующая квадратичная форма

$$P(x) = 2z_1(x_1z_2 + x_2z_3 + \dots + x_Nz_{N+1}) + (x, Ax)$$

не может быть ни положительно определенной, ни даже неотрицательно определенной, если матрица A положительно определена.

3. Полагая, что квадратичная форма (x, Ax) обладает положительным минимумом на пересечении плоскостей $(x, a^i) = b_i$, $i = 1, 2, \dots, k$, найти этот минимум и получить таким образом другой вывод результата § 6.

§ 7. **Общий случай.** Вернемся теперь к общей задаче, сформулированной в § 4. Нетрудно использовать для решения этой общей задачи те же приемы, что и в случае $k=1$. Но так как обозначения теперь становятся слишком громоздкими, то стоит потратить немного времени на то, чтобы ввести понятие прямоугольной матрицы и показать, насколько это новое обозначение упрощает выкладки.

Как только мы вводим прямоугольные матрицы (которые неплохо было бы в отличие от квадратных называть «таблицами», как это хотел сделать Кэли), естественно возникают матрицы, элементы которых сами являются матрицами. В некоторых ранее введенных обозначениях уже содержался намек на такие матрицы, и мы снова встретимся с ними при изучении кронекеровских произведений.

§ 8. **Прямоугольные матрицы.** Совокупность комплексных величин a_{ij} , расположенных в виде таблицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1M} \\ a_{21} & a_{22} & \dots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NM} \end{pmatrix}, \quad (1)$$

будем называть *прямоугольной матрицей*. Здесь уже нет той оговорки, что $M=N$. Мы будем также называть такую таблицу матрицей размера $M \times N$, или $M \times N$ -матрицей. Существен порядок, в котором пишутся M и N .

При сложении двух $M \times N$ -матриц не возникает каких-либо затруднений, однако не совсем ясно, пожалуй, как их следует перемножать. Мы будем допускать умножение $M \times N$ -матрицы A на $K \times M$ -матрицу B . Это ограничение появляется вследствие того, что произведение прямоугольных матриц, как и произведение квадратных матриц, возникает при описании повторных линейных преобразований. Таким образом, произведение двух прямоугольных матриц

$$AB = C = \|c_{ij}\|, \quad (2)$$

или схематично

$$\begin{array}{c} M \\ \boxed{A} \end{array} \begin{array}{c} K \\ \boxed{B} \end{array} = \begin{array}{c} K \\ \boxed{C} \end{array} \begin{array}{c} N \\ \end{array},$$

представляет собой $K \times N$ -матрицу, где

$$c_{ij} = \sum_{k=1}^M a_{ik} b_{kj}. \quad (3)$$

Матрица A' , транспонированная к $M \times N$ -матрице A , является $N \times M$ -матрицей и получается из A заменой ее строк столбцами.

Во многих случаях прямоугольные матрицы можно использовать для унификации обозначений в том смысле, что с их помощью стирается различие между векторами и матрицами. Тем не менее, поскольку между понятиями вектора и квадратной матрицы все же имеется существенная разница, смешивать эти два понятия не всегда желательно, особенно в анализе.

Хотя в одном из ближайших параграфов мы и собираемся привести пример упрощения, которого иногда можно достигнуть в результате использования подобного рода унификации, мы все же призываем читателя к осторожности. Основную роль всегда играют кроющиеся за символами математические идеи, и поэтому не стоит рабски следовать каким-либо обозначениям.

Упражнения

1. Пусть x и y — N -мерные векторы-столбцы. Показать, что

$$xy' = \|x_i y_j\|, \quad i, j = 1, 2, \dots, N.$$

2. Если x — N -мерный вектор-столбец, то

$$|\lambda I - xx'| = \lambda^N - (x', x) \lambda^{N-1}.$$

§ 9. Клеточные матрицы. Ранее мы ввели матрицы, элементами которых служат комплексные числа. Рассмотрим теперь матрицы, элементы которых сами являются комплексными матрицами:

$$A = \|A_{ij}\|. \quad (1)$$

Такая запись особенно удобна тем, что с ее помощью, например, можно представить 4×4 -матрицу в виде

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}, \quad (2)$$

где A_{ij} заданы следующим образом:

$$\begin{aligned} A_{11} &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, & A_{12} &= \begin{vmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{vmatrix}, \\ A_{21} &= \begin{vmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{vmatrix}, & A_{22} &= \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix}. \end{aligned} \quad (3)$$

В то же время можно считать, что

$$\begin{aligned} A_{11} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, & A_{12} &= \begin{vmatrix} a_{14} \\ a_{24} \\ a_{34} \end{vmatrix}, \\ A_{21} &= \|a_{41} \ a_{42} \ a_{43}\|, & A_{22} &= \|a_{44}\|. \end{aligned} \quad (4)$$

Как мы увидим, в некоторых случаях эта новая запись обладает определенными преимуществами. Естественно, нам хотелось бы сохранить для этих матриц обычные правила сложения и умножения.

Легко проверить, что сложение выполняется именно так, как мы того ожидали. Кроме того, читателю в виде упражнения рекомендуется проверить, что произведение матриц $\|A_{ij}\|$ и $\|B_{ij}\|$ можно смело определить следующим образом:

$$\|A_{ij}\| \|B_{ij}\| = \|C_{ij}\|, \quad (5)$$

где

$$C_{ij} = \sum_{k=1}^N A_{ik} B_{kj}. \quad (6)$$

При этом все произведения $A_{ik} B_{kj}$ уже определены. Отметим, что при таком определении умножения ассоциативность сохраняется.

Упражнения

1. Из соотношения

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} \begin{vmatrix} D & -B \\ -C & A \end{vmatrix} = \begin{vmatrix} AD-BC & BA-AB \\ CD-DC & DA-CB \end{vmatrix}$$

вывести, что

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD-BC|,$$

если $BA=AB$ или $CD=DC$, и все необходимые произведения существуют.

2. Пусть $M=A+iB$, где A и B — вещественные матрицы, и пусть

$$M_1 = \begin{vmatrix} A & -B \\ B & A \end{vmatrix}.$$

Показать, что

$$|M_1 - \lambda I| = |M - \lambda I| |\bar{M} - \lambda I|. \quad (\text{Сараки и Важевски.})$$

3 (продолжение). Отсюда вывести, что $|M_1| = |M| |\bar{M}|$ и что собственными значениями матрицы M_1 являются собственные значения матриц M и \bar{M} .

4 (продолжение). Показать, что матрица M_1 симметрическая, если M — эрмитова матрица.

§ 10. Решение задачи в общем случае. Рассмотрим теперь общую задачу, сформулированную в § 4. Как утверждает теорема 3, квадратичная форма

$$P = \sum_{i,j=1}^n \left(a_{ij} + \lambda \sum_{r=1}^k b_r b_{rj} \right) x_i x_j \quad (1)$$

должна быть положительно определенной при больших λ . Следовательно, должны выполняться неравенства

$$\left| a_{ij} + \lambda \sum_{r=1}^k b_r b_{rj} \right| > 0, \quad i, j = 1, 2, \dots, n, \quad (2)$$

при $n = 1, 2, \dots, N$ и любом достаточно большом положительном λ .

Чтобы упростить эти условия, воспользуемся тем же самым приемом, который мы применили в § 5, а также нашими новыми обозначениями. Пусть

$$B_N = \begin{vmatrix} b_{11} & b_{12} & \dots & b_{1N} \\ b_{21} & b_{22} & \dots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \dots & b_{NN} \end{vmatrix}. \quad (3)$$

Рассмотрим следующее произведение квадратных матриц $(N+k)$ -го порядка:

$$\begin{vmatrix} A_N & \lambda B_N \\ B_N^* & -I_k \end{vmatrix} \begin{vmatrix} I_N & 0 \\ 0 & I_k \end{vmatrix} = \begin{vmatrix} A_N + \lambda B_N B_N^* & \lambda B_N \\ 0 & -I_k \end{vmatrix}, \quad (4)$$

где через I_k обозначена единичная матрица порядка k , а

$$A_N = \|a_{ij}\|, \quad i, j = 1, 2, \dots, N.$$

Заменяя в (4) матрицы их определителями, получаем

$$(-1)^k \begin{vmatrix} A_N & \lambda B_N \\ B_N^* & -I_k \end{vmatrix} = |A_N + \lambda B_N B_N^*|. \quad (5)$$

Следовательно, условия (2) заменяются неравенством

$$(-1)^k \begin{vmatrix} A_N & \lambda B_N \\ B_N^* & -I_k \end{vmatrix} > 0, \quad (6)$$

которое должно выполняться для всех достаточно больших λ .

Таким образом, как и прежде, достаточные условия положительной определенности получить легко. Необходимые условия требуют более детального рассмотрения вследствие того, что некоторые из элементов b_{ij} и окаймленных определителей могут оказаться равными нулю.

Упражнения к гл. 5

1. Пусть $\{A_i\}$ — произвольный набор квадратных матриц порядка r . Для того чтобы существовали унитарные $(r \times r)$ -матрицы U и V такие, что $UA_iV = D_i$, где матрицы D_i диагональные и вещественные, необходимо и достаточно, чтобы $A_i \bar{A}_i = A_j \bar{A}_i$ и $\bar{A}_j A_i = \bar{A}_i A_j$ для всех i и j . (Н Вигман.)

2. Пусть $\sum_{i,j=1}^N a_{ij} x_i x_j$ — положительно определенная квадратичная форма. Показать, что

$$\sum_{i,j=1}^N \begin{vmatrix} a_{ik} & a_{ij} \\ a_{jk} & a_{il} \end{vmatrix} x_i x_j$$

также является положительно определенной квадратичной формой $N - 1$ переменных $x_1, x_2, \dots, x_{k-1}, x_{k+1}, \dots, x_N$. Каковы собственные значения и собственные векторы матрицы этой квадратичной формы? Обобщить этот результат.

3. Введем следующее соотношение между комплексными числами и матрицами:

$$z = x + iy \sim Z = \begin{vmatrix} x & y \\ -y & x \end{vmatrix}.$$

Показать, что (а) $ZW = WZ$; б) $\bar{z} \sim Z'$; в) из $z \sim Z$ и $w \sim W$ следует, что $zw \sim ZW$.

4 (продолжение). С помощью введенного соответствия и теоремы Муавра найти вид Z^n при $n = 1, 2, \dots$

5 (продолжение). Очевидно, имеют место соответствия

$$1 \sim \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \quad i \sim \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}.$$

Рассматривая матрицу

$$Q = \begin{vmatrix} x_1 + ix_2 & x_3 + ix_4 \\ -x_3 + ix_4 & x_1 - ix_2 \end{vmatrix}$$

и учитывая эти соответствия, мы приходим к клеточной матрице

$$Q = \begin{pmatrix} \begin{pmatrix} x_1 & x_3 \\ -x_3 & x_1 \end{pmatrix} & \begin{pmatrix} x_3 & x_4 \\ -x_4 & x_3 \end{pmatrix} \\ \begin{pmatrix} -x_3 & x_4 \\ -x_4 & -x_3 \end{pmatrix} & \begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix} \end{pmatrix}$$

или, если опустить круглые скобки,

$$Q_s = \begin{vmatrix} x_1 & x_3 & x_3 & x_4 \\ -x_2 & x_1 & -x_4 & x_3 \\ -x_3 & x_4 & x_1 & -x_2 \\ -x_4 & -x_3 & x_2 & x_1 \end{vmatrix}.$$

Рассмотрим соответствие $Q \sim Q_s$. Показать, что из $Q \sim Q_s$ и $W \sim W_s$ следует соответствие $QW \sim Q_s W_s$. Отсюда найти $|Q_s|$ и собственные значения Q_s .

6. Как найти максимальное значение, принимаемое формой (x, Ax) , если на x наложены ограничения $(x, x) = 1$, $(x, c^i) = 0$, $i = 1, 2, \dots, k$?

7. Соотношение $A \geq B$, где A и B — симметрические матрицы, означает, что матрица $A - B$ неотрицательно определенная. Показать, что из $A \geq B$, вообще говоря, не следует, что $A^2 \geq B^2$.

8. Пусть $g(t) \geq 0$; рассмотрим квадратичную форму

$$Q_N(x) = \int_0^1 g(t) \left(\sum_{k=0}^N x_k t^k \right)^2 dt.$$

Показать, что $Q_N(x)$ неотрицательно определена и вывести отсюда детерминантные неравенства для величин $m_k = \int_0^1 g(t) t^k dt$. (Стилтьес.)

9. Пусть $g(t) \geq 0$; рассмотрим эрмитову форму

$$H_N(x) = \int_0^1 g(t) \left| \sum_{k=0}^N x_k e^{2\pi i k t} \right|^2 dt.$$

Показать, что $H_N(\mathbf{x})$ положительно определена и вывести отсюда детерминантные неравенства для величин $m_k = \int_0^1 g(t) e^{2\pi i k t} dt$. (О. Тёплиц.)

10. Показать, что $\sum_{m, n=0}^N \frac{x_m x_n}{m+n+1} < \pi \sum_{n=0}^N x_n^2$. (Неравенство Гильберта.)

11. Доказать теорему 1 с помощью индукции¹⁾.

12. Определить минимум выражения

$$Q_N(\mathbf{x}) = \int_0^\pi \left| e^{-i\theta} - \sum_{k=0}^N x_k e^{ik\theta} \right|^2 d\theta.$$

13 (продолжение). К какому пределу стремится этот минимум при $N \rightarrow \infty$? (Этот результат играет важную роль в теории прогнозирования. См. Гренандер и Сегё, Тёплицевы формы и их приложения, ИЛ, 1961.)

14. Определить минимум выражения

$$Q_N(\mathbf{x}) = \int_0^1 \left| t^k - \sum_{l=0}^N x_l t^{\lambda_l} \right|^2 dt,$$

где $\{\lambda_l\}$ — последовательность действительных чисел таких, что $\lambda_0 < \lambda_1 < \dots$

15. Найти минимальное и максимальное значения функции $(\mathbf{x}, A\mathbf{x}) + 2(\mathbf{b}, \mathbf{x})$ на сфере $(\mathbf{x}, \mathbf{x}) = 1$. Сколько здесь имеется стационарных точек и все ли они действительны?

16. Сколько нормалей, действительных и комплексных, можно провести из данной точки плоскости к эллипсу $x^2/a^2 + y^2/b^2 = 1$?

17. Если $A = \|a_{ij}\|$, $C = \|c_{ij}\|$ и $\sum_j a_{ij} = 0$, $\sum_j a_{ij} = 0$ и $c_{ij} = c_i + c_j$, то матрицы AB и $A(B+C)$ имеют одно и то же характеристическое уравнение. (А. Брауэр.)

18. Если A , C_1 и C_2 таковы, что $C_1 A = A C_2 = 0$, $C = C_1 + C_2$, то матрицы AB и $A(B+C)$ имеют одно и то же характеристическое уравнение. (Паркер.)

19. Если $ACA = 0$, то матрицы AB и $A(B+C)$ имеют одно и то же характеристическое уравнение. (Паркер.)

20. Пусть $\lambda_1 < \lambda_2 < \dots < \lambda_N$ и $\mu_1 < \mu_2 < \dots < \mu_N$ — характеристические числа соответственно матриц H и $S'HS$, причем матрица H эрмитова, а S — произвольная действительная невырожденная матрица. Тогда число положительных, отрицательных и нулевых характеристических чисел у этих матриц одинаково. Количественное уточнение этого результата (принадлежащее Сильвестру и Якоби), который часто называют законом инерции, можно найти в статье

Островский (А. М. Ostrowski), A Quantitative Formulation of Sylvester's Law of Inertia, Proc. Nat. Acad. Sci. U.S.A. 45 (1959), 740—743.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 2. Эти детерминантные критерии могут быть получены с помощью последовательности Штурма, получаемой из характеристического многочлена $|A - \lambda I|$. Этот же прием можно ис-

¹⁾ См. Селье (C. J. Seelye), Amer. Math. Monthly 65 (1958), 355—356.

пользовать для доказательства того, что собственные значения действительны. См.

Бернсайд и Пэнтон (W. S. Burnside and A. W. Panton), *Theory of Equations*, Longmans, Green & Co., Inc., New York, 1928, vol. II, p. 65, Exercise 39, etc.

Замечательное обобщение этих неравенств содержится в работе

Шур (I. Schur), *Über endliche Gruppen und hermitische Formen*, *Math. Z.* 1 (1918), 184—207.

Некоторые интересные результаты, связанные с положительностью, можно найти в работе

Гуд (I. J. Good), *A Note on Positive Determinants*, *J. London Math. Soc.* 22 (1947), 92—95.

§ 4. Затронутый вопрос охватывается исследованиями, изложенными в работах

Дайнс (L. L. Dines), *On the Mapping of Quadratic Forms*, *Bull. Amer. Math. Soc.* 47 (1941), 494—498;

Дайнс (L. L. Dines), *On Linear Combinations of Quadratic Forms*, *Bull. Amer. Math. Soc.* 49 (1943).

Мы также воспользовались результатом Финслера; см.

Финслер (P. Finsler), *Über das Vorkommen definiten und semidefiniten Formen in schären quadratischen Formen*, *Commentarii Mathematici Helvetici* 9 (1937), 188—192.

Этот результат независимо от Финслера был получен Герштейном. Данное им доказательство предлагается в качестве упражнения в конце главы. Ряд других доказательств имеется в недавно вышедших работах. См., например,

Афрайэт (S. N. Afriat), *The Quadratic Form Positive Definite on a Linear Manifold*, *Proc. Cambridge Phil. Soc.* 47 (1951), 1—6.

§ 5. Упражнения 3—6 в конце этого параграфа иллюстрируют важный эвристический принцип, который можно сделать строгим с помощью теории инвариантов. Этот принцип гласит: всякий раз, когда мы получаем условия положительности в форме неравенств, имеет место более точный результат в терминах расстояния, площади, объема и т. д., из которого данное неравенство следует с очевидностью.

§ 6. Задача определения условий положительности формы (x, Ax) при $x \geq 0$ рассматривается в статье

Гэддам (J. W. Gaddum), *Linear Inequalities and Quadratic Forms*, *Pacific J. Math.* 8 (1958), 411—414.

ФУНКЦИИ ОТ МАТРИЦЫ

§ 1. Введение. В этой главе мы сосредоточим свое внимание на понятии функции от симметрической матрицы. Сначала мы рассмотрим две важные матричные функции: обратную матрицу и квадратный корень из матрицы. Последняя функция представляет особый интерес в связи с вопросом о положительной определенности матриц. Затем мы займемся наиболее важными скалярными функциями от матриц — коэффициентами характеристического многочлена.

Определение матричных функций в случае матриц общего вида — задача несколько более сложная, чем это может показаться на первый взгляд. Поэтому всякое детальное обсуждение различных предложенных для этой цели способов мы откладываем до гл. 11. В случае симметрических матриц существование диагональной канонической формы устраняет большую часть трудностей. Эту диагональную форму, существование которой было установлено в гл. 4, мы используем также для того, чтобы получить параметрическое представление для элементов симметрической матрицы, которое нередко оказывается полезным при различных доказательствах. В качестве примера мы докажем принадлежащий Шуру интересный результат, касающийся произведения двух положительно определенных матриц.

В заключение мы выведем важное соотношение между определителем положительно определенной матрицы и соответствующей квадратичной формой, которое послужит основой для одной из последующих глав, посвященной неравенствам. Аналогичный результат мы получим и для эрмитовых матриц.

§ 2. Функции от симметрической матрицы. Как мы видели, всякую вещественную симметрическую матрицу можно представить в виде

$$A = T \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix} T', \quad (1)$$

где T — ортогональная матрица. Отсюда следует, что

$$A^k = T \begin{vmatrix} \lambda_1^k & & & 0 \\ & \lambda_2^k & & \\ & & \ddots & \\ 0 & & & \lambda_N^k \end{vmatrix} T' \quad (2)$$

для любого целого k . Следовательно, для всякой аналитической функции $f(z)$ мы можем определить матричную функцию $f(A)$ следующим образом:

$$f(A) = T \begin{vmatrix} f(\lambda_1) & & & 0 \\ & f(\lambda_2) & & \\ & & \ddots & \\ 0 & & & f(\lambda_N) \end{vmatrix} T', \quad (3)$$

если только скалярная функция $f(z)$ определена в точках $\lambda_1, \lambda_2, \dots, \lambda_N$.

§ 3. Обратная матрица. Следуя по тому же пути, мы приходим к следующему определению обратной матрицы:

$$A^{-1} = T \begin{vmatrix} \lambda_1^{-1} & & & 0 \\ & \lambda_2^{-1} & & \\ & & \ddots & \\ 0 & & & \lambda_N^{-1} \end{vmatrix} T', \quad (1)$$

причем все λ_i предполагаются отличными от нуля. Ясно, что A^{-1} обладает обычными свойствами обратной матрицы, именно:

$$AA^{-1} = A^{-1}A = I. \quad (2)$$

§ 4. Единственность обратной матрицы. Возникает естественный вопрос, является ли обратная матрица единственной, если она существует. Прежде всего заметим, что все λ_i отличны от нуля тогда и только тогда, когда $|A| \neq 0$. Действительно, если $\lambda = 0$ — корень уравнения $|A - \lambda I| = 0$, то мы видим, что и $|A| = 0$; обратно, если $|A| = 0$, то $\lambda = 0$ является корнем этого уравнения.

Матрица A называется *вырожденной (особенной)*, если $|A| = 0$. Если $|A| \neq 0$, то A называется *невырожденной (неособенной)*.

Тотчас же видно, что вырожденная матрица не может иметь обратной, так как из $AB = I$ следует, что $|AB| = 1$, или $|A| \cdot |B| = 1$. Если же $|A| = 0$, то последнее равенство невозможно.

С другой стороны, покажем теперь, что всякая невырожденная матрица, симметрическая или нет, имеет единственную об-

ратную. В случае симметрической матрицы A обратная матрица, определяемая новым способом, должна, очевидно, совпадать с (3.1).

Уравнение $AB = I$ эквивалентно N^2 уравнениям

$$\sum_{k=1}^N a_{ik}b_{kj} = \delta_{ij}, \quad i, j = 1, 2, \dots, N. \quad (1)$$

Фиксируя j , мы получаем систему линейных уравнений относительно b_{kj} , $k = 1, 2, \dots, N$. Эта система имеет своим определителем $|A| \neq 0$ при всех индексах j .

Таким образом, если A — невырожденная матрица, то существует единственное решение системы (1). Из (1) также легко увидеть, что это решение дается формулой

$$b_{ij} = \frac{\alpha_{ji}}{|A|}, \quad (2)$$

где α_{ji} — алгебраическое дополнение элемента a_{ji} матрицы A .

Упражнения

1. Показать тремя способами:

(а) непосредственно из (3.1),

(б) используя формулу (4.2),

(в) используя соотношение $AA^{-1} = A^{-1}A = I$,

что невырожденная симметрическая матрица имеет симметрическую обратную, причем обратная матрица единственна.

2. Пусть A — вырожденная матрица. Показать, что можно найти такую матрицу B с произвольно малыми элементами, что матрица $A+B$ будет невырожденной.

3. Пусть $|\lambda I - A| = \lambda^N + c_1\lambda^{N-1} + \dots + c_N$ — характеристический многочлен матрицы. Показать, что $A^{-1} = -(A^{N-1} + c_1A^{N-2} + \dots + c_{N-1}I)/c_N$, если только $c_N \neq 0$.

4. Если A_1, A_2, B_1 и B_2 — невырожденные матрицы одного и того же порядка, то

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}^{-1} = \begin{vmatrix} (A_1 - B_1B_2^{-1}A_2)^{-1} & (A_2 - B_2B_1^{-1}A_1)^{-1} \\ (B_1 - A_1A_2^{-1}B_2)^{-1} & (B_2 - A_2A_1^{-1}B_1)^{-1} \end{vmatrix}$$

5. Пусть A — прямоугольная матрица размеров $N \times M$. Рассмотрим систему несовместных линейных уравнений $x = Ay$, где x — N -мерный вектор, а y — M -мерный вектор. Чтобы получить наилучшее приближение для этой системы, мы определим y из условия минимума квадратичной функции $(x - Ay, x - Ay)$. Показать, что при условии несовместности система имеет единственное наилучшее приближение, которое задается формулой $y = (A'A)^{-1}A'x$. Определить после этого минимум функции $(x - Ay, x - Ay)$.

6. Пусть A и B — симметрические матрицы, причем $A \geq B$. Как и ранее, соотношение $A \geq B$ означает, что матрица $A - B$ неотрицательно определенная. Показать, что из $A \geq B > 0$ следует, что $B^{-1} \geq A^{-1}$.*

7. Если S — вещественная кососимметрическая матрица, то матрица $I + S$ невырожденная

* См. упражнение 2 к § 12 гл. 4. (Прим ред.)

8. Если матрица S вещественная и кососимметрическая, то матрица $T = (I - S)(I + S)^{-1}$ ортогональная (преобразование Кэли).

9. Если A — ортогональная матрица, причем $A + I$ — невырожденная матрица, то A можно представить в виде

$$A = (I - S)(I + S)^{-1},$$

где S — действительная кососимметрическая матрица.

10. Для любой матрицы A существует матрица J такая, что ее диагональные элементы равны ± 1 , а все остальные элементы — нули, причем матрица $JA + I$ будет невырожденной.

11. Используя предыдущий результат, показать, что всякую ортогональную матрицу A можно представить в виде

$$A = J(I - S)(I + S)^{-1},$$

где J — матрица, описанная в упражнении 10.

12. Показать, что

(а) если $A_1 \geq B_1$, $A_2 \geq B_2$, то $A_1 + A_2 \geq B_1 + B_2$;

(б) если $A_1 \geq B_1$, $B_1 \geq C_1$, то $A_1 \geq C_1$;

(в) если $A_1 \geq B_1$, $A_2 \geq B_2$, то не обязательно $A_1 A_2 \geq B_1 B_2$, даже если $A_1 A_2$ и $B_1 B_2$ — симметрические матрицы (таким образом, из $A \geq B$ еще не следует, что $A^2 \geq B^2$);

(г) если $A_1 \geq B_1$, то $T' A_1 T \geq T' B_1 T$.

13. Показать, что если A — положительно определенная матрица, то A^{-1} можно определить с помощью соотношения

$$(x, A^{-1}x) = \max_y [2(x, y) - (y, Ay)].$$

14 (продолжение). Используя это соотношение, показать, что $B^{-1} \geq A^{-1}$, если A и B — симметрические матрицы и $A \geq B > 0$ *).

15. Если A — матрица, все элементы которой — действительные целые числа, и $|A| = \pm 1$, то A^{-1} представляет собой матрицу этого же типа.

16. Если A — матрица, все элементы которой — комплексные целые числа, т. е. числа вида $x + iy$, где x и y — действительные целые числа, и если $|A| = \pm 1$ или $\pm i$, то A^{-1} представляет собой матрицу этого же типа.

17. Каковы необходимые и достаточные условия того, чтобы система $Ax = y$ имела целочисленные решения, если все компоненты вектора y и элементы матрицы A — целые числа.

18. Показать, что $|A + iB|^2 = |A|^2 |I + A^{-1}BA^{-1}B|$, если $H = A + iB$ — эрмитова матрица, а A невырождена (A и B вещественны).

19. Показать, что $(z, Hz) = (x, Ax) + (y, Ay) + 2(Bx, y)$, если $H = A + iB$ — эрмитова матрица, а $z = x + iy$.

20. Показать, что $(x, Ax) = (x, A_R x)$, где $A_R = (A + A')/2$. (Предполагается, что x — вещественный вектор. — Прим. ред.)

21. Как получить элементы матрицы A^{-1} , зная элементы матрицы B^{-1} , если A и B совпадают, за исключением одного столбца.

22. Если матрица A порядка $2N$ имеет вид

$$A = \begin{vmatrix} 0 & x_1 & 0 & \dots & \\ y_1 & 0 & x_2 & 0 & \dots \\ 0 & y_2 & 0 & x_3 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{vmatrix}$$

* См. упражнение 6 к настоящей главе. (Прим. ред.)

при условии, что $x_i, y_i \neq 0$ при нечетных номерах i , то элементы b_{ik} матрицы A^{-1} определяются по формуле

$$b_{ik} = \begin{cases} 0, & k \text{ нечетное;} \\ (-1)^{k/2-1} \prod_{j=0}^{k/2-1} \frac{x_{2j}}{y_{2j+1}}, & k \text{ четное,} \end{cases}$$

где $x_0 = 1$ (Клем-ит.)

23. Пользуясь предыдущим результатом, найти общий вид матрицы A^{-1} .

§ 5. Квадратные корни. Поскольку положительно определенная матрица является естественным обобщением положительного числа, возникает вопрос, имеет ли положительно определенная матрица положительно определенный квадратный корень.

Подобно тому, как мы поступали в § 2, можно определить $A^{1/2}$ соотношением

$$A^{1/2} = T \begin{vmatrix} \lambda_1^{1/2} & & & 0 \\ & \lambda_2^{1/2} & & \\ & & \ddots & \\ & & & \lambda_N^{1/2} \end{vmatrix} T' \quad (1)$$

и получить таким образом матрицу, которая удовлетворяет соотношению $B^2 = A$ и является положительно определенной, если A положительно определена.

Выясним теперь вопрос о единственности квадратного корня. Так как матрица B симметрическая, то она обладает представлением

$$B = S \begin{vmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_N \end{vmatrix} S', \quad (2)$$

где S — ортогональная матрица. Матрицы B и $B^2 = A$ перестановочны, следовательно, они могут быть приведены к диагональной форме с помощью одной и той же ортогональной матрицы T . Соотношение $B^2 = A$ показывает, что B имеет вид (1), где квадратные корни берутся со знаком плюс.

Упражнения

1. Сколько симметрических квадратных корней имеет положительно определенная $N \times N$ -матрица?
2. Может ли симметрическая матрица иметь несимметрические квадратные корни?

3. Если матрица B эрмитова и $B \geq 0$, то A совпадает с единственным неотрицательным эрмитовым квадратным корнем матрицы B , если $A^2 = B$ и $(A + A^*) \geq 0$. [Путнам (Putnam), On Square Roots and Logarithms of Operators, Purdue University, PRF-1421, 1958.]

§ 6. Параметрическое представление. Формула

$$A = T \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{pmatrix} T' \quad (1)$$

дает нам параметрическое представление элементов симметрической матрицы через элементы ортогональной матрицы T , именно,

$$a_{ij} = \sum_{k=1}^N \lambda_k t_{ik} t_{jk}. \quad (2)$$

Это представление, как мы скоро увидим, можно использовать для получения некоторых свойств элементов симметрической матрицы A .

Упражнение

1. Получить параметрическое представление элементов симметрической 2×2 -матрицы через $\cos \theta$ и $\sin \theta$.

§ 7. Результат Шура. Используя представление (6.2), легко доказать следующую теорему:

Теорема 1. Если $A = \|a_{ij}\|$ и $B = \|b_{ij}\|$ — положительно определенные матрицы, то матрица

$$C = \|a_i b_{ij}\|$$

также является положительно определенной.

Доказательство. Имеем

$$\begin{aligned} \sum_{i,j=1}^N a_i b_{ij} x_i x_j &= \sum_{i,j=1}^N b_{ij} x_i x_j \left(\sum_{k=1}^N \lambda_k t_{ik} t_{jk} \right) = \\ &= \sum_{k=1}^N \lambda_k \left(\sum_{i,j=1}^N b_{ij} x_i t_{ik} x_j t_{jk} \right), \quad \lambda_k > 0. \quad (1) \end{aligned}$$

Как квадратичная форма переменных $x_i t_{ik}$, выражение $\sum_{i,j=1}^N b_{ij} x_i t_{ik} x_j t_{jk}$ положительно, если хотя бы одна из величин $x_i t_{ik}$ отлична от нуля. Но так как

$$\sum_{i,k} x_i^2 t_{ik}^2 = \sum_i x_i^2 \sum_k t_{ik}^2 = \sum_i x_i^2, \quad (2)$$

то, очевидно, все $x_i t_{ik}$ равны нулю тогда и только тогда, когда равны нулю все x_i . Это и устанавливает положительную определенность матрицы $C = \|a_{ij} b_{ij}\|$.

Упражнение

1. Используя упражнение 2 к гл. 5, дать индуктивное доказательство этого результата, отталкиваясь от легко доказываемого случая $N=2$.

§ 8. Основные скалярные функции. Рассмотрим теперь некоторые свойства скалярных функций матрицы A , определяемых посредством характеристического полинома

$$|\lambda I - A| = \lambda^N - \varphi_1(A) \lambda^{N-1} + \varphi_2(A) \lambda^{N-2} + \dots + (-1)^N \varphi_N(A). \quad (1)$$

Матрицы, фигурирующие в этом параграфе, являются квадратными матрицами общего вида, не обязательно симметрическими.

Из соотношения между коэффициентами и корнями полиномиального уравнения следует, что

$$\varphi_1(A) = \lambda_1 + \lambda_2 + \dots + \lambda_N,$$

$$\varphi_2(A) = \sum_{i>j} \lambda_i \lambda_j,$$

⋮
⋮
⋮

$$\varphi_N(A) = \lambda_1 \lambda_2 \dots \lambda_N. \quad (2)$$

С другой стороны, выписывая коэффициенты при λ^{N-1} в том виде, в каком они получаются при раскрытии определителя $|A - \lambda I|$, мы видим, что

$$\varphi_1(A) = a_{11} + a_{22} + \dots + a_{NN}. \quad (3)$$

Полагая же $\lambda = 0$, убеждаемся, что

$$\varphi_N(A) = |A|. \quad (4)$$

Линейная функция $\varphi_1(A)$ играет первостепенную роль в теории матриц. Она носит название *следа матрицы* A и обозначается $\text{Sp } A$.

Из (3) следует, что

$$\text{Sp}(A + B) = \text{Sp}(A) + \text{Sp}(B), \quad (5)$$

а несложным подсчетом нетрудно убедиться в том, что

$$\text{Sp}(AB) = \text{Sp}(BA) \quad (6)$$

для любых двух матриц A и B . Эти соотношения справедливы, несмотря на тот факт, что никаких простых соотношений, связывающих характеристические числа матриц A , B и $A + B$ или

AB , не существует. Соотношения (5) и (6) являются частным случаем следующей теоремы:

Теорема 2. Для любых двух матриц A и B имеют место равенства

$$\varphi_k(AB) = \varphi_k(BA), \quad k = 1, 2, \dots, N. \quad (7)$$

Доказательство будет опираться на факт, уже известный нам, а именно, на то, что при $k = N$ равенство (7) имеет место, т. е. $|AB| = |BA|$ для любых двух матриц A и B .

Если A — невырожденная матрица, то

$$|\lambda I - AB| = |A(\lambda A^{-1} - B)| = |(\lambda A^{-1} - B)A| = |\lambda I - BA|, \quad (8)$$

откуда и следует утверждение теоремы.

Если матрица A вырожденная, то соотношение (8) можно получить предельным переходом, отправляясь от матрицы $A + \varepsilon I$, где ε мало, а матрица $A + \varepsilon I$ невырожденная.

В нижеследующих упражнениях намечен другой подход к доказательству этой теоремы.

Упражнения

1. Показать, что $|I + tA| = 1 + t \operatorname{Sp}(A) + \dots$
2. Показать, что $\varphi_k(TAT^{-1}) = \varphi_k(A)$; $\varphi_k(TAT') = \varphi_k(A)$, если T — ортогональная матрица.
3. Показать, что любой полином $p(a_{11}, a_{12}, \dots, a_{NN})$ от элементов матрицы A , удовлетворяющий условию $p(AB) = p(BA)$, где A и B — произвольные матрицы, должен быть полиномом от функций $\varphi_k(A)$.
4. Показать, что $\operatorname{Sp}((AB)^k) = \operatorname{Sp}((BA)^k)$ при $k = 1, 2, \dots$ и, следовательно, $\varphi_k(AB) = \varphi_k(BA)$ при $k = 1, 2, \dots, N$.
5. Показать, что если $\operatorname{Sp}(AX) = 0$ при всех матрицах X , то $A = 0$.
6. Показать, что результаты упражнения 2 и теоремы 2 этого параграфа являются частным случаем следующего результата: для любого набора из m матриц A_1, A_2, \dots, A_m порядка N величина $\varphi_k(A_1 A_2 \dots A_m)$ остается неизменной после циклической перестановки сомножителей A_1, A_2, \dots, A_m .

§ 9. Несобственный интеграл $\int_{-\infty}^{\infty} e^{-(x, Ax)} dx$. Важную роль в анализе играет интеграл

$$I_N = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-i(x, Ax)} dx, \quad (1)$$

где $dx = dx_1 dx_2 \dots dx_N$.

Теорема 3. Если матрица A положительно определенная, то

$$I_N = \frac{\pi^{N/2}}{|A|^{1/2}}. \quad (2)$$

Приведем два доказательства этой теоремы.

Первое доказательство. Пусть T — ортогональная матрица, приводящая матрицу A к диагональному виду. Введем замену переменных $x = Ty$, тогда

$$(x, Ax) = (y, Ay). \quad (3)$$

Далее, $\prod_{i=1}^N dx_i = \prod_{i=1}^N dy_i$, поскольку якобиан замены $x = Ty$ равен $|T|$, который можно взять равным $+1$. Так как матрица T устанавливает взаимно однозначное соответствие, то

$$I_N = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\lambda_1 y_1^2 - \lambda_2 y_2^2 - \dots - \lambda_N y_N^2} dy = \\ = \prod_{i=1}^N \int_{-\infty}^{\infty} e^{-\lambda_i y_i^2} dy_i = \frac{\pi^{N/2}}{(\lambda_1 \lambda_2 \dots \lambda_N)^{1/2}}. \quad (4)$$

Учитывая, что $\prod_{i=1}^N \lambda_i = |A|$, а

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \pi^{1/2},$$

получаем формулу (2).

Второе доказательство. В § 3 гл. 5 было показано, что

$$(x, Ax) = \sum_{k=1}^N \frac{D_k}{D_{k-1}} y_k^2, \quad D_0 = 1, \quad (5)$$

если матрица A положительно определенная. В формуле (5)

$$y_k = x_k + \sum_{j=k+1}^N b_{kj} x_j, \quad k = 1, 2, \dots, N. \quad (6)$$

Сделаем замену переменных, определенную формулами (6). Якобиан опять равен 1, а соответствие между старыми и новыми переменными взаимно однозначное. Следовательно,

$$I_N = \left(\int_{-\infty}^{\infty} e^{-D_1 y_1^2} dy_1 \right) \left(\int_{-\infty}^{\infty} e^{-D_2 y_2^2 / D_1} dy_2 \right) \dots \left(\int_{-\infty}^{\infty} e^{-D_N y_N^2 / D_{N-1}} dy_N \right) = \\ = \frac{\pi^{N/2}}{D_N^{1/2}} = \frac{\pi^{N/2}}{|A|^{1/2}}. \quad (7)$$

Упражнения

1. Показать, что если A — положительно определенная матрица, а B симметрическая, то

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-(x, Ax) - i(x, Bx)} dx = \frac{\pi^{N/2}}{|A + iB|^{1/2}}.$$

Для вычисления интеграла следует

(а) положить $x = Ty$, где $T'AT = \Lambda$, а Λ — диагональная матрица;

(б) сделать еще одну замену переменных $y_k = z_k/\lambda_k^{1/2}$;

(в) в получающемся подинтегральном выражении

$$- \sum_{k=1}^N z_k^2 - l(x, Cx)$$

матрицу C привести к диагональному виду ортогональным преобразованием $z = Sw$;

(г) вычислить полученный интеграл.

2. Вычислить интеграл

$$I_{mn} = \int_0^{\infty} \int_0^{\infty} e^{-a_{11}x_1^2 - 2a_{12}x_1x_2 - a_{22}x_2^2} x_1^m x_2^n dx_1 dx_2$$

для целых неотрицательных значений m и n . Учтеь при этом, что

$$I_{00} = \frac{\pi}{(a_{11}a_{22} - a_{12}^2)^{1/2}}$$

а

$$I_{2m, 2n} = \left(\frac{\partial}{\partial a_{11}} \right)^m \left(\frac{\partial}{\partial a_{22}} \right)^n I_{00}$$

Как исследовать общий случай: m или n нечетные.

3. Вычислить интеграл

$$I_N = \int_{-\infty}^{\infty} e^{-(x, ABx)} dx$$

для положительно определенных матриц A и B .

4. Вычислить интеграл

$$I_N = \int_{-\infty}^{\infty} e^{-(x, Ax) + y^l(x, y)} dx.$$

§ 10. Аналог для эрмитовых матриц. По аналогии с (9.1) рассмотрим интеграл

$$J(H) = \int_{-\infty}^{\infty} e^{-\bar{y}(H, Hx)} dx dy, \quad (1)$$

где $dx = dx_1 dx_2 \dots dx_N$, $dy = dy_1 dy_2 \dots dy_N$, а H — положительно определенная эрмитова матрица.

Представим H в виде $H = A + iB$, где матрица A действительная и положительно определенная, а матрица B действительная кососимметрическая. Тогда

$$J(H) = \int_{-\infty}^{\infty} e^{-(x, Ax) - y(Bx, y) - (y, Ay)} dx dy. \quad (2)$$

Поскольку этот интеграл сходится абсолютно, то можно интегрировать сначала по x , а затем по y .

Используя соотношение

$$(x, Ax) + 2(Bx, y) = (x, Ax) - 2(x, By) = \\ = (A(x - A^{-1}By), x - A^{-1}By) - (By, A^{-1}By), \quad (3)$$

получаем

$$\int_{-\infty}^{\infty} e^{-(x, Ax) - 2(Bx, y)} dx = \frac{\pi^{N/2} e^{(By, A^{-1}By)}}{|A|^{1/2}}. \quad (4)$$

Следовательно,

$$J(H) = \frac{\pi^{N/2}}{|A|^{1/2}} \int_{-\infty}^{\infty} e^{-[(y, Ay) + (y, BA^{-1}By)]} dy = \\ = \frac{\pi^{N/2}}{|A|^{1/2}} \cdot \frac{\pi^{N/2}}{|A - BA^{-1}B|^{1/2}} = \frac{\pi^N}{|A| |I + A^{-1}BA^{-1}B|^{1/2}}. \quad (5)$$

§ 11. Связь между $J(H)$ и $|H|$. Остается теперь выразить $J(H)$ через $|H|$. Имеем

$$|H| = |A + iB| = |A| |I + iA^{-1}B|, \\ |H'| = |A - iB| = |A| |I - iA^{-1}B|. \quad (1)$$

Так как $|H| = |H'|$, то

$$|H|^2 = |A|^2 |(I + iA^{-1}B)(I - iA^{-1}B)| = |A|^2 |I + A^{-1}BA^{-1}B|. \quad (2)$$

Сравнивая эти соотношения, получаем

$$J(H) = \pi^N / |H|. \quad (3)$$

Упражнения к гл. 6

1. Показать, что характеристические числа матрицы $\rho(A)$, где $\rho(\lambda) =$ — полином, равны $\rho(\lambda_i)$, и, следовательно, $|\rho(A)| = \prod_i \rho(\lambda_i)$

2. При каких условиях аналогичный результат имеет место для рациональных функций?

3. Найти обратные матрицы к следующим матрицам:

$$C = \begin{vmatrix} x & -y \\ y & x \end{vmatrix}, \quad Q = \begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2 & x_1 & -x_4 & x_3 \\ -x_3 & x_4 & x_1 & -x_2 \\ -x_4 & -x_3 & x_2 & x_1 \end{vmatrix}.$$

4. Пусть $A_0 = I$, а последовательность скаляров и матриц определяется формулами $c_k = \text{Sp}(AA_{k-1})/k$, $k = 1, 2, \dots$; $A_k = AA_{k-1} - c_k I$. Тогда $A^{-1} = A_{n-1}/c_n$. (Фрейм.)

5. Показать, что $|A + tB| = |A| (1 + t \text{Sp}(A^{-1}B) + \dots)$.

6. Используя представление

$$\begin{aligned} |A + \varepsilon B + \lambda I| &= \varphi_N(A + \varepsilon B) + \lambda \varphi_{N-1}(A + \varepsilon B) + \dots = \\ &= |A + \lambda I| (1 + \varepsilon \operatorname{Sp}((A + \lambda I)^{-1}B) + \dots), \end{aligned}$$

получить соотношение

$$\varphi_{N-1}(A + \varepsilon B) = \varphi_{N-1}(A) - \varepsilon \varphi_N(A) \operatorname{Sp}(A^{-1}B) + \dots;$$

получить аналогичный результат для каждого $\varphi_k(A)$.

7. Показать, что выполнение условия $\operatorname{Sp}(AB) > 0$ для всех положительно определенных матриц B необходимо и достаточно для того, чтобы $A \geq 0$ и $A \neq 0$.

8. Показать, что

$$|I + \lambda A| = e^{\sum_{k=1}^n \frac{(-1)^{k-1}}{k} \lambda^k \operatorname{Sp}(A^k)}$$

вывести отсюда соотношения, связывающие $\varphi_k(A)$ и $\operatorname{Sp}(A^k)$.

9. Показать, что матрицы $AB + B$ и $BA + B$ имеют одинаковый определитель, для чего установить формулу $\operatorname{Sp}((AB + B)^n) = \operatorname{Sp}((BA + B)^n)$ для $n = 1, 2, \dots$.

10. Пусть $B_0 = I$, $B_i = AB_{i-1} + k_i I$, где $k_i = -\operatorname{Sp}(AB_{i-1})/i$, тогда

$$(-1)^n |A - \lambda I| = \lambda^n + k_1 \lambda^{n-1} + \dots \quad (\text{Леверье.})$$

11. Пусть A, B, \dots — перестановочные матрицы и пусть $f(x_1, x_2, \dots)$ — произвольная рациональная функция. Показать, что характеристические числа a_1, a_2, \dots, a_n матрицы A , b_1, b_2, \dots, b_n матрицы B, \dots можно упорядочить таким образом, чтобы характеристические числа матрицы $f(A, B, \dots)$ имели вид $f(a_1, b_1, \dots)$, $f(a_2, b_2, \dots)$ и т. д.

12. Показать, что каждая невырожденная матрица может быть представлена в виде произведения двух, не обязательно действительных, симметрических матриц и притом бесконечным числом способов. (Восс.)

13. Если H — неотрицательно определенная эрмитова матрица, то существует треугольная матрица T такая, что $H = TT^*$. (Теплиц.)

14. Для любой невырожденной матрицы A существует треугольная матрица T такая, что матрица TA унитарная. (Э. Шмидт.)

15. Пусть P, Q, R, X — матрицы второго порядка. Показать, что всякое характеристическое число матрицы X , являющейся решением уравнения $PX^2 + QX + R = 0$, есть корень уравнения $|P\lambda^2 + Q\lambda + R| = 0$ (**). (Сильвестр.)

16. Если матрицы A, B и C положительно определенные, то корни уравнения $|A\lambda^2 + B\lambda + C| = 0$ имеют отрицательные действительные части (***)

17. Рассмотрим матрицы A и B такие, что $AB = r_1 BA$, где r_1 есть корень q -й степени из единицы. Показать, что характеристические числа матриц A и B , λ_i и μ_i , можно упорядочить таким образом, что величины $(\lambda_i^q + \mu_i^q)^{1/q}$ и $r_1^{(q-1)/2} \lambda_i \mu_i$ будут характеристическими числами матриц $A + B$ и AB соответственно. Для различных значений i , возможно, потребуются выделить различные ветви корня q -й степени. (Поттер.)

18. Используя представление

$$\frac{x^{N/2}}{|A + \varepsilon B|^{1/2}} = \int_{-\infty}^{\infty} e^{-x(\lambda, Ax) - \varepsilon(x, Bx)} \prod dx_i$$

*) Равенство $|AB + B| = |BA + B|$ очевидно. (Прим. ред.)

**) См. упражнение 8 к гл. 11. (Прим. ред.)

***) См. упражнение 25 к гл. 4 и § 8 гл. 13. (Прим. ред.)

и разлагая обе части этого равенства по степеням ϵ , получить выражения для коэффициентов разложения $|A + \epsilon B|$ по степеням ϵ ; получить отсюда выражения для функций $\varphi_k(A)$.

19. Пусть A и B — унитарные матрицы, обладающие тем свойством, что при всех значениях скаляров c_1 и c_2 матрица $c_1 A + c_2 B$ имеет своими характеристическими числами величины $c_1 \lambda_l + c_2 \mu_l$, где λ_l и μ_l — характеристические числа матриц A и B соответственно. Тогда $AB = BA$. (Мощкин.)

20. Показать, что матрица A невырожденная, если

$$|a_{ll}| > \sum_{k \neq l} |a_{lk}|, \quad l = 1, 2, \dots, N.$$

21. Показать, что собственные векторы матрицы A являются собственными векторами $p(A)$ для любого полинома p . Обратное утверждение не всегда имеет место.

22. Если матрицы A и B симметрические и такие, что

$$|I - \lambda A| |I - \mu B| = |I - \lambda A - \mu B|$$

при всех λ и μ , то $AB = 0$. (Крейг — Хотеллинг.)

23. Пусть $AB = 0$, и пусть $p(A, B)$ — полином со свободным членом, равным нулю. Тогда $|\lambda I - p(A, B)| = \lambda^{-N} |\lambda I - p(A, 0)| \cdot |\lambda I - p(B, 0)|$. (Г. Шнайдер.)

24. Обобщить этот результат на полиномы вида $p(A_1, A_2, \dots, A_n)$, где $A_i A_j = 0$, $i < j$. (Г. Шнайдер.)

25. Пусть A и B — положительно определенные матрицы. Используя соотношение $|AB| = |A^{1/2}| |B| |A^{1/2}|$, показать, что

$$\frac{(\sqrt{\pi})^N}{|AB|^{1/2}} = \int_{-\infty}^{\infty} e^{-(A^{1/2}x, BA^{1/2}x)} \prod_{i=1}^N dx_i.$$

26. Показать, что

$$\int_{-\infty}^{\infty} e^{-(x, Ax)/(x+(y, x))} \prod_i dx_i = (2\pi)^{N/2} |A|^{-1/2} e^{-(y, A^{-1}y)/2}.$$

27. Пусть Y — положительно определенная, а X — симметрическая матрицы второго порядка. Рассмотрим интеграл

$$J(Y) = \int_{X > 0} e^{-Sp(XY)} |X|^{s-1/2} dx_{11} dx_{12} dx_{22},$$

где область интегрирования определяется неравенствами $x_{11} > 0$, $x_{11}x_{22} - x_{12}^2 > 0$. Другими словами, это область, где матрица X является положительно определенной. Тогда

$$J(Y) = \frac{\pi^{1/2} \Gamma(s) \Gamma(s-1/2)}{|Y|^s}$$

при $\text{Re}(s) > 1/2$. (Ингам — Зигель.) (См. комментарий к § 9, где имеется информация о дальнейших результатах этого типа.)

28. Пусть A — матрица с комплексными элементами порядка N : $A = \|a_{rs}\| = \|b_{rs} + ic_{rs}\|$. Пусть, далее, B — действительная матрица порядка $2N$, полученная из матрицы A заменой элементов $b_{rs} + ic_{rs}$ их матричными эквивалентами

$$b_{rs} + ic_{rs} \rightarrow \begin{vmatrix} b_{rs} & c_{rs} \\ -c_{rs} & b_{rs} \end{vmatrix}$$

Показать, что A^{-1} можно получить, обратив B и воспользовавшись указанным соответствием.

29. Пусть A — симметрическая матрица и $|A| \neq 0$; тогда окаймленный определитель

$$B(x) = \begin{vmatrix} a_{11} & \dots & a_{1N} & x_1 \\ a_{21} & \dots & a_{2N} & x_2 \\ \dots & \dots & \dots & \dots \\ a_{N1} & \dots & a_{NN} & x_N \\ x_1 & \dots & x_N & 0 \end{vmatrix}.$$

умноженный на второй главный минор матрицы A , представляет собой квадрат некоторой линейной функции аргументов x_i . (Бернсайд—Пэнтон.)

30. Следовательно, если матрица A симметрическая и ее первый главный минор равен нулю, то определитель и его второй главный минор имеют противоположные знаки.

31. Пусть X и A — квадратные матрицы второго порядка. Найти все решения уравнения $X^2 = A$, имеющие вид $X = c_1 I + c_2 A$, где c_1 и c_2 — скаляры.

32. Подобно этому, рассмотрим случай, когда X и A — квадратные матрицы третьего порядка, а $X = c_1 I + c_2 A + c_3 A^2$.

33. Пусть $f(t)$ — полином степени, не большей чем $N-1$, и пусть $\lambda_1, \lambda_2, \dots, \lambda_N$ представляют собой N различных характеристических чисел матрицы A . Тогда

$$f(A) = \sum_{i=1}^N f(\lambda_i) \prod_{\substack{j=1 \\ j \neq i}}^N \left[\frac{A - \lambda_j I}{\lambda_i - \lambda_j} \right].$$

(Интерполяционная формула Сильвестра.)

Чему равна правая часть этого соотношения, если $f(\lambda)$ — полином степени, большей или равной N ?

34. Справедлива ли эта формула для $f(A) = A^{-1}$?

35. Показать, что

$$\int_{(x, Ax) < 1} dx_1 dx_2 \dots dx_N = \frac{\pi^{N/2} |A|^{-1/2}}{\Gamma(N/2 + 1)}.$$

36. Если элементы a_{ij} квадратной матрицы A порядка N действительны и удовлетворяют условиям $a_{ii} > \sum_{i \neq j} |a_{ij}|$, $i = 1, 2, \dots, N$, то свободный член в разложении Лорана функции

$$f(z_1, z_2, \dots, z_N) = \prod_{i=1}^N \left(\sum_{k=0}^{\infty} a_{i,k} z_k / z_i \right)^{-1}$$

на контуре $|z_1| = |z_2| = \dots = |z_N|$ равен $|A|^{-1}$.

Другими словами,

$$\frac{1}{|A|} = \frac{1}{(2\pi i)^N} \oint_{|z_1|=1} \dots \oint_{|z_N|=1} \prod_{i=1}^N \left[\frac{dz_i}{\sum_k a_{i,k} z_k} \right].$$

Этот результат остается справедливым и для комплексных матриц таких, что $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$. Некоторые дальнейшие результаты имеются в работе Уитла¹⁾.

Приведенные соотношения принадлежат Якоби; некоторые их приложения можно найти в книгах Вейля²⁾ и Форсайта³⁾.

37. Показать, что если матрицы A и B неотрицательно определенные, то $\sum_{i,j} a_{ij} b_{ij} \geq 0$. Вывести отсюда теорему Шура, приведенную в § 7. (Фейер.)

38. Пусть $\{A_i\}$ есть набор симметрических матриц порядка $N \times N$. Введем скалярное произведение по формуле

$$(A_i, A_j) = \text{Sp}(A_i A_j).$$

Пусть, далее, этот набор состоит из $M = N(N+1)/2$ линейно независимых симметрических матриц. Используя введенное скалярное произведение, построить ортонормированный набор $\{Y_i\}$ линейно независимых матриц и показать, что любая симметрическая матрица X порядка N может быть пред-

ставлена в виде $X = \sum_{i=1}^M c_i Y_i$, где $c_i = (X, Y_i)$.

39. В условиях предыдущего упражнения рассмотреть следующий набор матриц второго порядка:

$$A_1 = \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix}, \quad A_2 = \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix}, \quad A_3 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}.$$

и построить соответствующий ортонормированный набор.

40. Рассмотреть аналогичный набор матриц $\{A_i\}$ в общем случае и построить матрицы $\{Y_i\}$.

41. Теорема Гамильтона—Кэли утверждает, что всякая квадратная матрица A удовлетворяет своему характеристическому уравнению. Минимальным многочленом матрицы A называется многочлен $q(\lambda)$ минимальной степени, обладающий тем свойством, что $q(A) = 0$. Показать, что $q(\lambda)$ является делителем характеристического многочлена.

42. Найти якобиан преобразования $Y = X^{-1}$.

43. Вычислить функцию от матрицы

$$\frac{1}{2\pi i} \int_C f(\lambda) (\lambda I - A)^{-1} d\lambda$$

при различных предположениях относительно матрицы A и контура интегрирования C . Пуанкаре первым использовал контурный интеграл

$\frac{1}{2\pi i} \int_C f(\lambda) (\lambda I - A)^{-1} d\lambda$ для представления функции от матрицы $f(A)$.

(H. Poincaré, Sur les groupes continus, Trans. Cambridge Phil. Soc. 18 (1899), 220—225 and Oeuvres, vol. 3, 173—212.)

¹⁾ P. Whittle, Some Combinatorial Results for Matrix Powers, Quart. J. Math. 7 (1956), 316—320.

²⁾ H. Weyl, The Classical Groups ..., Princeton University Press, N. J., 1946. [Русский перевод: Классические группы, ИЛ, 1947.]

³⁾ A. R. Forsyth, Lectures Introductory to the Theory of Functions of Two Complex Variables, Cambridge University Press, New York, 1914.

⁴⁾ Олкин (I. Olkin), Note on the Jacobians of Certain Matrix Transformations Useful in Multivariate Analysis, Biometrika 40 (1953), 43—46.

44. Доказать, что если уравнение $Ax = b$ имеет решение при всех b , то существует обратная матрица A^{-1} .

45. Для любых двух положительных констант k_1 и k_2 существуют матрицы A и B такие, что каждый элемент матрицы $AB - I$ по абсолютной величине меньше k_1 , а у матрицы $BA - I$ есть элементы, большие k_2 . (Хаусхолдер.)

46. Если матрица A_1 невырожденная, то

$$\begin{vmatrix} A_1 & A_2 \\ A_2 & A_1 \end{vmatrix} = |A_1| |A_1 - A_2 A_1^{-1} A_2|.$$

47. Если $BC = CB$ и

$$A = \begin{vmatrix} I & 0 & \dots & \dots \\ -B & I & 0 & \dots \\ C - B & I & 0 & \dots \\ 0 & C - B & I & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}.$$

то

$$A^{-1} = \begin{vmatrix} I & 0 & \dots & \dots \\ D_1 & I & 0 & \dots \\ D_2 & D_1 & I & 0 & \dots \\ D_3 & D_2 & D_1 & I & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}.$$

Найти рекуррентные соотношения для матриц D_i .

48. Если A — прямоугольная матрица размера $N \times M$ и ранга r , $r < N$, C — прямоугольная матрица размера $M \times N$, удовлетворяющая условию $ACA = kA$, где k — скаляр, а B — прямоугольная матрица размера $M \times N$, то характеристическое уравнение матрицы AB имеет вид $\lambda^{N-r} \varphi(\lambda) = 0$, а $\lambda^{N-r} \varphi(\lambda - k) = 0$ является характеристическим уравнением матрицы $A(B+C)$. (Паркер.)

49. Выразить X через Y , если $Y = (AX + B)(CX + B)^{-1}$.

50. Ранее мы ввели соответствие

$$a + bi \sim \begin{vmatrix} a & b \\ -b & a \end{vmatrix}.$$

Показать, что равенство $\overline{f(z)} = f(\bar{z})$, где $f(z)$ — степенной ряд, является необходимым и достаточным условием, при котором выполняется соотношение

$$f(a + bi) \sim f\left(\begin{vmatrix} a & b \\ -b & a \end{vmatrix}\right).$$

(Обобщение этого результата читатель найдет в работе D. W. Robinson, Math. Mag. 32 (1959), 218—215.)

51. Если $F(z) = \sum_{n=0}^{\infty} c_n z^n$, где $c_n \geq 0$ и $F(z) \neq 0$, то $F(A)$ положительно определенная всякий раз, когда матрица A положительно определенная.

52. Если матрица $F(A)$ положительно определенная для любой матрицы вида $\|a_{i-j}\|$, то $F(z) = \sum_{n=0}^{\infty} c_n z^n$, где $c_n \geq 0$. (У. Рудин.) (Более ранний результат принадлежит Шёнбергу. См. Schoenberg, Duke Math. J., 1942.)

53. Матрица $A = \|a_{ij}\|$ называется *матрицей Якоби*, если $a_{ij} = 0$ при $|i-j| \geq 2$. В какой степени определены элементы a_{ij} симметрической якобиевой матрицы A , если известны ее характеристические числа. Эта задача представляет собой конечномерный вариант задачи нахождения функции $\phi(x)$ по спектру уравнения $u'' + \lambda \phi(x) u = 0$ при различных краевых условиях¹⁾.

54. Пусть матрица A имеет вид

$$A = \begin{vmatrix} a_1 & 1 & 0 & \dots & \dots \\ 1 & a_2 & 1 & 0 & \dots \\ 0 & 1 & a_3 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{vmatrix}.$$

Как выражаются числа a_i через характеристические числа матрицы A ?

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 1. Подробное рассмотрение функций от матриц имеется в книге

Мак-Даффи (С. С. MacDuffee), The Theory of Matrices, Ergebnisse der Mathematik, Reprint, Chelsea Publishing Co., New York, 1946*).

Более поздние статьи по этому вопросу:

Райнхарт (R. F. Rinehart), The Derivative of a Matrix Function, Proc. Amer. Math. Soc. 7 (1956), 2—5;

Рихтер (H. Richter), Über Matrixfunktionen, Math. Ann. 122 (1950), 16—34;

Афрайэт (S. N. Afriat), Analytic Functions of Finite Dimensional Linear Transformations. Proc. Cambridge Phil. Soc. 55 (1959), 51—56.

Прекрасный обзор многих вопросов, связанных с функциями от матриц, дан Райнхартом:

Райнхарт (R. F. Rinehart), The Equivalence of Definitions of a Matrix Function, Amer. Math. Monthly 62 (1955), 395—414.

См. также

Двайр и Макфейл (P. S. Dwyer and M. S. Macphail), Symbolic Matrix Derivatives, Ann. Math. Stat. 19 (1948), 517—534.

¹⁾ Борг (G. Borg), Eine Umkehrung der Sturm-Liouvilleschen Eigenwerte Aufgabe, Acta Math. 78 (1946), 1—96. [См. также: М. Г. Крейн, Решение обратной задачи Штурма—Лиувилля, ДАН СССР 76, № 1 (1951), 21—23; М. Г. Крейн, Определение плотности неоднородной симметричной струны по спектру ее частот, ДАН СССР 76, № 3 (1951), 345—348; И. М. Гельфанд и Б. М. Левитан, Об определении дифференциального уравнения по его спектральной функции, ИАН СССР, сер. матем. 15, № 4 (1951), 309—365. См., кроме того, непосредственно относящееся к данному вопросу дополнение II в книге Ф. Р. Гантмахера и М. Г. Крейна, «Оscилляционные матрицы и ядра и малые колебания механических систем», Гостехиздат, 1950. (Прим. ред.)]

²⁾ См. также Ф. Р. Гантмахер. Теория матриц, «Наука», 1967, гл. V. (Прим. ред.)

Рассматриваемые нами вопросы тесно связаны с функциями гиперкомплексных величин. В частности, для кватернионов, которые можно рассматривать как матрицы вида

$$Q = \begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2 & x_1 & -x_4 & x_3 \\ -x_3 & x_4 & x_1 & -x_2 \\ -x_4 & -x_3 & x_2 & x_1 \end{vmatrix}.$$

развита довольно полная теория, во многом аналогичная обычной теории функций комплексного переменного. Кстати, последнюю можно рассматривать как теорию действительных матриц специального вида

$$z = \begin{vmatrix} x & y \\ -y & x \end{vmatrix}.$$

См. статью

Фуэтер (R. Fueter), *Functions of a Hyper Complex Variable*, University of Zurich, 1948—1949, reprinted by Argonne National Laboratory, 1959.

См. также

Фуэтер (R. Fueter), *Commentarii Math. Helvetici* 20, 419—420,

где имеются указания на связанную с этими вопросами литературу.

Упомянем, наконец, работу

Рот (W. E. Roth), *A Solution of the Matrix Equation $P(X) = A$* , *Trans. Amer. Math. Soc.* 30 (1928), 597—599,

где изложены ранние результаты Сильвестра, Кэли и Фробениуса, связанные с решением полиномиальных уравнений этого типа.

Дальнейшие ссылки на работы по теории функций от матриц будут приведены в конце гл. 10.

§ 3. Понятие обобщенного обращения для вырожденных матриц ввел

Мур (E. H. Moore), *General Analysis, Part. I*, *Mem. Amer. Phil. Soc.* 1 (1935), 197,

а позднее, независимо от Мура,

Пенроуз (R. Penrose), *A Generalized Inverse for Matrices*, *Proc. Cambridge Phil. Soc.* 51 (1953), 406—413 *).

Эти и другие вопросы рассматриваются в работе

Хестинс (M. R. Hestenes), *Inversion of Matrices by Biorthogonalization*, *J. Soc. Indust. Appl. Math.* 6 (1958), 51—90.

*) См. по этому поводу Ф. Р. Гантмахер, *Теория матриц*, «Наука», 1967, гл. I, стр 32. (*Прим. ред.*).

Изложение вычислительных аспектов обращения матриц и связанных с ними различных аналитических вопросов можно найти в работе

фон Нейман и Голдстейн (J. von Neumann and H. Goldstine), Numerical Inverting of Matrices of High Order, Bull. Amer. Math. Soc. 53 (1947), 1021—1099.

§ 4. Прекрасное изложение различных методов обращения матриц, а также большое число ссылок читатель найдет в работе

Гринспен (D. Greenspan), Method of Matrix Inversion, Amer. Math. Monthly 62 (1955), 303—319*.

В численном анализе чрезвычайно важную роль играет задача нахождения реализуемых методов обращения матриц, которые отличаются от теоретических. Эти вопросы будут подробно рассмотрены Дж. Форсайтом в последующем томе.

Очень часто необходимо знать, что исследуемая матрица является невырожденной. Удобно поэтому иметь несколько простых, не связанных с громоздкими вычислениями критериев, гарантирующих отличие от нуля определителя матрицы.

По-видимому, наиболее полезным из них является следующий (упражнение 20 к этой главе):

Если матрица A действительная и $|a_{ii}| > \sum_{i \neq j} |a_{ij}|$, $i = 1, 2, \dots, N$, то $|A| \neq 0$.

Историю этого результата, а также ряд его обобщений читатель найдет в работе

Таусски (O. Taussky), A Recurring Theorem on Determinants, Amer. Math. Monthly 56 (1949), 672—676**).

Тесно связана с предыдущей задача построения оценок характеристических чисел матрицы по ее элементам. С этой задачей мы еще встретимся в гл. 16. Подробное же рассмотрение этой задачи мы отложим до одного из последующих томов нашей серии.

Имеется библиография по этой теме:

Таусски (O. Taussky), Bibliography on Bounds for Characteristic Roots of Finite Matrices, National Bureau of Standards Report, September 1951.

Отметим еще, что из предыдущего результата следует полезный результат Гершгорина, состоящий в том, что все характеристические числа матрицы A лежат внутри кругов с центрами a_{ii} и радиусами, равными $\sum_{i \neq j} |a_{ij}|$, $i = 1, 2, \dots, N$. За ори-

* См. также Д. К. Фаддеев и В. Н. Фаддеева, Вычислительные методы линейной алгебры, Физматгиз, 1963. (Прим. ред.)

** См. также Ф. Р. Гантмахер, Теория матриц, «Наука», 1967, гл. XIV. (Прим. ред.)

гинальным изложением этого результата мы отсылаем читателя к работе

Гершгорин С., Über die Abgrenzung der Eigenwerte einer Matrix, ИАН СССР 7 (1931), 749—754.

§ 6. Как уже упоминалось ранее, с помощью эллиптических функций можно получить другое представление квадратных матриц третьего порядка; по этому поводу см. работы

Каспари (F. Caspary), Zur Theorie der Thetafunktionen mit zwei Argumenten, Kronecker J. 44, 74—86;

Каспари (F. Caspary), Sur les systèmes orthogonaux formés par les fonctions thêta, Comptes Rendus 54, 490—493,

а также ряд других статей этого же автора, написанных в тот же период.

§ 7. См. статью

Шур (I. Schur), Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen, J. Math. 140 (1911), 1—28.

См. также

Фейер (L. Fejer), Über die Eindeutigkeit der Lösung partiellen Differentialgleichung zweiter Ordnung, Math. Z. 1 (1918), 70—79,

где имеется ссылка на более раннюю работу Мотарда. См. еще работу

Левин (H. Lewy), Composition of Solutions of Linear Partial Differential Equations in Two Independent Variables, J. Math. and Mech. 8 (1959), 185—192.

Скалярное произведение матриц, определяемое формулой $(A, B) = \sum a_{ij} b_{ij}$, использовалось Р. Ольденбургером для других целей в работе

Ольденбургер (R. Oldenburger), Expansions of Quadratic Forms, Bull. Amer. Math. Soc. 49 (1943), 136—141.

§ 9. Этот несобственный интеграл играет важную роль в различных разделах анализа. Мы будем существенно опираться на этот интеграл и его обобщение в гл. 8, посвященной неравенствам. Далеко идущие обобщения формулы (2) могут быть получены с помощью интегралов, введенных Ингамом и Зигелем; см. работы

Ингам (A. E. Ingham), An Integral which Occurs in Statistics, Proc. Cambridge Phil. Soc. 29 (1933), 271—276;

Зигель (C. L. Siegel), Über die analytische Theorie der quadratischen Formen, Ann. Math. 36 (1935), 527—606.

Другие обобщения и связанные с ними результаты можно найти в работах

Беллман (R. Bellman), A Generalization of Some Integral Identities Due to Ingham and Siegel, Duke Math. J. 24 (1956), 571—578;

- Герц (C. S. Herz), Bessel Functions of Matrix Argument, Ann. Math. 61 (1955), 474—523;
- Олкин (I. Olkin), A Class of Integral Identities with Matrix Argument, Duke Math. J. 26 (1959), 207—213;
- Бохнер (S. Bochner), Group Invariance of Cauchy's Formula in Several Variables, Ann. Math. 45 (1944), 686—707;
- Беллман (R. Bellman), Generalized Eisenstein Series and Nonanalytic Automorphic Functions, Proc. Nat. Acad. Sci. U. S. A. 36 (1950), 356—359;
- Маас (H. Maass), Zur Theorie der Kugelfunktionen einer Matrixvariablen, Math. Z. 135 (1953), 391—416;
- Поттер (H. S. A. Potter), The Volume of a Certain Matrix Domain, Duke Math. J. 18 (1951), 391—397.

Интегралы этого типа, их аналоги и обобщения очень естественно возникают в различных задачах многомерного статистического анализа. См., например,

- Андерсон и Гиршик (T. W. Anderson and M. A. Girschick), Some Extensions of the Wishart Distributions, Ann. Math. Stat. 15 (1944), 345—357;
- Раш (G. Rasch), A Functional Equation for Wishart's Distribution, Ann. Math. Stat. 19 (1948), 262—266;
- Ситгривс (R. Sitgreaves), On the Distribution of Two Random Matrices Used in Classification Procedures, Ann. Math. Stat. 23 (1952), 263—270.

В связи с функциями от матриц уместно упомянуть исследования Лёвнера в различных областях математической физики и прикладной математики. Как мы уже отмечали в нескольких упражнениях, из соотношения $A \geq B$, понимаемого в том смысле, что разность $A - B$ двух симметрических матриц A и B является неотрицательно определенной матрицей, не следует, что $A^2 \geq B^2$, хотя и верно, что $B^{-1} \geq A^{-1}$, если $A \geq B > 0$. Впервые этот вопрос обсуждался в общем виде в работе

Лёвнер (C. Loewner), Math. Z. 88 (1934), 177—216.

См. также

Добш (R. Dobsch), Math. Z. 43 (1937), 353—388.

В своей работе Лёвнер рассматривал задачу определения класса функций, для которых $f(A) \geq f(B)$, если $A \geq B$. Эти функции обладают тем свойством, что $\operatorname{Re} [f(z)] \geq 0$ всякий раз, когда $\operatorname{Re}(z) \geq 0$. Они играют важную роль в современной теории цепей; см., например, работы

Вайнберг и Слепян (L. Weinberg and P. Slepian), Positive Real Matrices, Hughes Research Laboratories, Culver City, Calif., 1955;

Эфферц (F. H. Effertz), On the Synthesis of Networks Containing Two Kinds of Elements, Proc. Symposium on Modern Network Synthesis, Polytechnic Institute of Brooklyn, 1955,

где дан прекрасный обзор полученных в этой области результатов, и

Даффин (R. J. Duffin), Elementary Operations Which Generate Network Matrices, Proc. Amer. Math. Soc. 6 (1955), 335—339.

Кроме того, эти функции имеют важное значение в некоторых разделах современной физики. По этому поводу см. работы

Лейн и Томас (A. M. Lane and R. G. Thomas), R-matrix Theory of Nuclear Reactions, Revs. Mod. Phys. 39 (1958), 257—352 [русский перевод: Теория ядерных реакций при низких энергиях, ИЛ, 1960];

Вигнер (E. Wigner), On a Class of Analytic Functions from the Quantum Theory of Collisions, Ann. Math. 53 (1951), 36—67.

Дальнейшее рассмотрение этих вопросов содержится в более поздней работе

Лёвнер (C. Loewner), Some Classes of Functions Defined by Difference or Differential Inequalities, Bull. Amer. Math. Soc. 56 (1950), 308—319,

и в статье

Бендат и Шерман (J. Bendat and S. Sherman), Monotone and Convex Operator Functions, Trans. Amer. Math. Soc. 79 (1955), 58—71.

Здесь же рассматривается понятие выпуклой функции от матрицы, введенное Краусом в работе

Краус (F. Kraus), Math. Z. 41 (1936), 18—42.

§ 10. Этот результат содержится в статье

Беллман (R. Bellman), Representation Theorems and Inequalities for Hermitian Matrices, Duke Math. J., 1959.

**ВАРИАЦИОННОЕ ОПИСАНИЕ
ХАРАКТЕРИСТИЧЕСКИХ ЧИСЕЛ**

§ 1. Введение. В предыдущих главах мы имели возможность убедиться в том, что как наибольшее, так и наименьшее характеристические числа имеют весьма простую геометрическую интерпретацию, отправляясь от которой, мы пришли к вариационной формулировке задачи нахождения этих чисел.

Как мы увидим далее, те же самые геометрические соображения приводят к соответствующим вариационным задачам для остальных характеристических чисел. Однако упомянутый вариационный подход значительно уступает как в эффективности, так и в красоте другому подходу, которому мы обязаны Р. Куранту и Э. Фишеру. Оба упомянутых подхода опираются на каноническую форму, полученную в гл. 4.

Используя подход Куранта и Фишера, мы выведем ряд интересных свойств симметрических матриц.

§ 2. Отношение Релея. Как мы знаем, значения, принимаемые квадратичной формой (x, Ax) на сфере $(x, x) = 1$, совпадают со значениями, принимаемыми квадратичной формой $(y, \Lambda y) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_N y_N^2$ на сфере $(y, y) = 1$, где $\Lambda = T'AT$, $y = Tx$, а матрица T ортогональная. В дальнейшем мы будем предполагать, что

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N. \quad (1)$$

Учитывая это условие, легко убедиться в справедливости неравенств

$$\begin{aligned} (y, \Lambda y) &\geq \lambda_N (y_1^2 + y_2^2 + \dots + y_N^2), \\ (y, \Lambda y) &\leq \lambda_1 (y_1^2 + y_2^2 + \dots + y_N^2), \end{aligned} \quad (2)$$

из которых сразу же получаем представления для λ_1 и λ_N :

$$\begin{aligned} \lambda_1 &= \max_y \frac{(y, \Lambda y)}{(y, y)} = \max_x \frac{(x, Ax)}{(x, x)}, \\ \lambda_N &= \min_y \frac{(y, \Lambda y)}{(y, y)} = \min_x \frac{(x, Ax)}{(x, x)}. \end{aligned} \quad (3)$$

Отношение

$$q(x) = \frac{(x, Ax)}{(x, x)} \quad (4)$$

часто называют *отношением Релея*.

Из выражений (3) следует, что при всех x справедливы неравенства

$$\lambda_1 \geq \frac{(x, Ax)}{(x, x)} \geq \lambda_N. \quad (5)$$

Соотношения такого типа очень важны в приложениях, когда требуется быстро оценить характеристические числа. Замечательно, что простой выбор значений координат x_i часто дает весьма точные приближения для λ_1 и λ_N . Это обстоятельство способствовало успеху многих важных результатов теоретической физики.

Интерес к характеристическим числам (и в меньшей степени к собственным векторам) вызван тем, что во многих прикладных вопросах физики λ_i представляют собой собственные частоты. Более общо можно сказать, что в приложениях λ_i являются наблюдаемыми величинами, с помощью которых можно провести сравнение теоретических выводов с экспериментом.

Вопросы этого характера будут детально рассматриваться в отдельной книге настоящей серии, посвященной вычислительным аспектам теории матриц; мы же лишь слегка коснемся их ниже в § 12.

§ 3. Вариационные свойства характеристических чисел. Начнем наше рассмотрение вариационных свойств с доказательства следующей теоремы:

Теорема 1. Пусть x^i — набор из N собственных векторов, соответствующих характеристическим числам λ_i . Тогда при $k = 1, 2, \dots, N$

$$\lambda_k = \max_{R_k} (x, Ax) / (x, x), \quad (1)$$

где R_k — область x -пространства, определяемая соотношениями ортогональности

$$(x, x^i) = 0, \quad i = 1, 2, \dots, k-1, \quad x \neq 0. \quad (2)$$

Геометрический смысл этого результата ясен. Для того чтобы определить, например, вторую по величине полуось эллипсоида, нужно найти максимальное расстояние от начала координат в плоскости, перпендикулярной наибольшей полуоси, до поверхности эллипсоида.

Докажем формулу (1) аналитически. Представим вектор x в виде

$$x = \sum_{k=1}^N u_k x^k. \quad (3)$$

Тогда

$$(x, Ax) = \sum_{k=1}^N \lambda_k u_k^2, \quad (4)$$

$$(x, x) = \sum_{k=1}^N u_k^2.$$

Соотношение (1) для λ_1 совпадает с результатом, установленным в § 2. Рассмотрим выражение для λ_2 . Условие $(x, x^1) = 0$ равносильно условию $u_1 = 0$. Ясно поэтому, что максимум выражения $\sum_{k=2}^N \lambda_k u_k^2$ при условии $\sum_{k=2}^N u_k^2 = 1$ равен λ_2 .

Точно таким же образом легко установить справедливость (1) для всех остальных k .

§ 4. Обсуждение. Как хорошо известно из физики, увеличение жесткости стержня или пружины приводит к увеличению всех собственных частот. Аналитический эквивалент этого явления состоит в следующем: характеристические числа суммы $A + B$ равномерно больше характеристических чисел матрицы A , если матрица B положительно определенная.

Легко показать справедливость этого утверждения для λ_1 и λ_N . Используя очевидные обозначения, имеем

$$\lambda_1(A+B) = \max_x \frac{(x, (A+B)x)}{(x, x)} = \max_x \left[\frac{(x, Ax)}{(x, x)} + \frac{(x, Bx)}{(x, x)} \right] \geq$$

$$\geq \max_x \frac{(x, Ax)}{(x, x)} + \min_x \frac{(x, Bx)}{(x, x)} \geq \lambda_1(A) + \lambda_N(B). \quad (1)$$

Поскольку по предположению $\lambda_N(B) > 0$, мы видим, что $\lambda_1(A+B) > \lambda_1(A)$, если матрица B положительно определенная. Доказательство того, что $\lambda_N(A+B) > \lambda_N(A)$, проводится аналогично.

Если, однако, попытаться подобным же образом провести доказательство для характеристических чисел $\lambda_2, \lambda_3, \dots, \lambda_{N-1}$, то мы столкнемся с тем в данном случае неприятным обстоятельством, что вариационное описание характеристических чисел в теореме 1 индуктивно.

В формулу для λ_2 входит собственный вектор x^1 и т. д.

$\mu_2 \leq \lambda_2$. Для этого в области, определенной соотношением $(y, y) = 1$, выделим совокупность, состоящую из одного вектора с $y_1 = 1$. Такой выбор вектора y приводит к тому, что первая компонента z_1 вектора z должна быть равна нулю. Поскольку минимум по подмножеству всегда больше минимума по множеству, содержащему в себе это подмножество, или равен ему, мы имеем

$$\begin{aligned} \mu_2 &\leq \max_{z_1=0} \{(\lambda_1 z_1^2 + \lambda_2 z_2^2 + \dots + \lambda_N z_N^2) / (z_1^2 + z_2^2 + \dots + z_N^2)\} \leq \\ &\leq \max_z \{(\lambda_2 z_2^2 + \dots + \lambda_N z_N^2) / (z_2^2 + \dots + z_N^2)\} = \lambda_2. \end{aligned} \quad (9)$$

Следовательно, $\mu_2 = \lambda_2$.

Воспользуемся теперь тем же самым приемом для того, чтобы установить этот результат для общего случая. Рассмотрим величину

$$\begin{aligned} \mu_k &= \min_{\substack{(y^i, y^i)=1 \\ i=1, 2, \dots, k-1}} \max_{(x, y^i)=0} (x, Ax) / (x, x) = \\ &= \min_{i=1, 2, \dots, k-1} \max_{(x, y^i)=0} \left\{ \frac{\sum_{k=1}^N \lambda_k z_k^2}{\sum_{k=1}^N z_k^2} \right\}. \end{aligned} \quad (10)$$

Будем искать максимум по подобласти

$$S: z_{k+1} = z_{k+2} = \dots = z_N = 0 \quad \text{и} \quad (z, y^i) = 0. \quad (11)$$

Легко видеть, что

$$\mu_k \geq \min_{\substack{(y^i, y^i)=1 \\ i=1, 2, \dots, k-1}} \max_S \left\{ \frac{\sum_{i=1}^k \lambda_i z_i^2}{\sum_{i=1}^k z_i^2} \right\} \geq \lambda_k. \quad (12)$$

так как $\sum_{i=1}^k \lambda_i z_i^2 / \sum_{i=1}^k z_i^2 \geq \lambda_k$ для всех z .

Для того чтобы показать, что $\mu_k \leq \lambda_k$, рассмотрим минимизацию по совокупности, состоящей из $(k-1)$ векторов

$$y^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad y^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad y^{k-1} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}. \quad (13)$$

Условия ортогональности $(y^i, z) = 0$ эквивалентны в этом случае равенствам

$$z_1 = z_2 = \dots = z_{k-1} = 0. \quad (14)$$

Как и ранее, мы видим, что $\mu_k \leq \lambda_k$, а следовательно, $\mu_k = \lambda_k$.

§ 7. Монотонное поведение $\lambda_k(A)$. Представление характеристических чисел λ_k , данное в теореме 2, делает очевидным следующий результат.

Теорема 3. Пусть A и B — симметрические матрицы, причем матрица B неотрицательно определенная. Тогда

$$\lambda_k(A+B) \geq \lambda_k(A), \quad k=1, 2, \dots, N. \quad (1)$$

Если матрица B положительно определенная, то

$$\lambda_k(A+B) > \lambda_k(A), \quad k=1, 2, \dots, N. \quad (2)$$

Упражнение

1. Найти оценку снизу для разности $\lambda_k(A+B) - \lambda_k(A)$.

§ 8. Теорема отделения Штурма. Теорема 2 позволяет также доказать следующую теорему.

Теорема 4. Рассмотрим последовательность симметрических матриц

$$A_r = \|a_{ij}\|, \quad i, j=1, 2, \dots, r, \quad (1)$$

где $r=1, 2, \dots, N$.

Пусть $\lambda_k(A_r)$, $k=1, 2, \dots, r$, обозначает k -е характеристическое число матрицы A_r , где в соответствии с ранее принятым обозначением

$$\lambda_1(A_r) \geq \lambda_2(A_r) \geq \dots \geq \lambda_r(A_r). \quad (2)$$

Тогда

$$\lambda_{k+1}(A_{i+1}) \leq \lambda_k(A_i) \leq \lambda_k(A_{i+1}). \quad (3)$$

Доказательство. Мы лишь наметим доказательство, оставив читателю в качестве полезного упражнения дополнить его деталями. Мы докажем неравенства

$$\lambda_2(A_{i+1}) \leq \lambda_1(A_i) \leq \lambda_1(A_{i+1}). \quad (4)$$

Имеем

$$\begin{aligned} \lambda_1(A_i) &= \max_x (x, A_i x) / (x, x), \\ \lambda_1(A_{i+1}) &= \max_x (x, A_{i+1} x) / (x, x), \end{aligned} \quad (5)$$

где x , фигурирующий в первом выражении, представляет собой i -мерный вектор, а x , фигурирующий во втором выражении

представляет собой $(i+1)$ -мерный вектор. Рассматривая вариацию по множеству $(i+1)$ -мерных векторов x с $(i+1)$ -й компонентой, равной нулю, мы убеждаемся, что $\lambda_1(A_i) \leq \lambda_1(A_{i+1})$. Для того чтобы получить неравенство $\lambda_2(A_{i+1}) \leq \lambda_1(A_i)$, мы воспользуемся соотношениями (6.2). Тогда

$$\begin{aligned} \lambda_1(A_i) &= \max_{(y^k, y^k)=1} \min_{(x, y^k)=0} (x, A_i x)/(x, x), \\ \lambda_2(A_{i+1}) &= \max_{(y^k, y^k)=1} \min_{(x, y^k)=0} (x, A_{i+1} x)/(x, x). \end{aligned} \quad (6)$$

Все векторы, фигурирующие в выражении для $\lambda_1(A_i)$, являются i -мерными, тогда как в выражении для $\lambda_2(A_{i+1})$ входят $(i+1)$ -мерные векторы. В последнем выражении рассмотрим только те векторы x , у которых $(i+1)$ -я компонента равна нулю. Ясно, что для векторов x этого типа можно максимизацию вести только по таким векторам y^k , у которых $(i+1)$ -я компонента равна нулю. Следовательно, $\lambda_2(A_{i+1}) \leq \lambda_1(A_i)$.

§ 9. Необходимое и достаточное условие положительной определенности матрицы A . Используя предыдущий результат, мы можем получить другое доказательство того факта, что необходимым и достаточным условием положительной определенности матрицы A является положительность всех ее главных миноров: $|A_k| > 0$, $k = 1, 2, \dots, N$.

Как мы знаем, для того чтобы матрица A была положительно определенной, необходимо и достаточно, чтобы $\lambda_k(A) > 0$, $k = 1, 2, \dots, N$. Если A положительно определенная, а следовательно, и $\lambda_k(A) > 0$, то в силу доказанной теоремы отделения $\lambda_k(A_i) > 0$,

$k = 1, 2, \dots, i$. Таким образом, $|A_i| = \prod_{k=1}^i \lambda_k(A_i) > 0$.

С другой стороны, если определители $|A_k|$ положительны при всех k , то, в частности, при $k=1$ мы имеем: $\lambda_1(A_1) > 0$. Поскольку далее

$$\lambda_1(A_2) \geq \lambda_1(A_1) \geq \lambda_2(A_2), \quad (1)$$

то из условия $|A_2| > 0$ вытекает, что $\lambda_2(A_2) > 0$. Индукцией устанавливаем, что $\lambda_k(A_i) > 0$, $k = 1, 2, \dots, i$, при $i = 1, 2, \dots, N$. Таким путем мы приходим к доказательству сформулированного утверждения. Детали доказательства мы опять оставляем читателю в качестве упражнения.

§ 10. Теорема отделения Пуанкаре. Установим теперь следующий результат, полезный как для вычислительных, так и для аналитических целей.

Теорема 5. Пусть $\{y^k\}$, $k = 1, 2, \dots, K$, — набор из K ортогональных векторов, и пусть $x = \sum_{k=1}^K u_k y^k$, так что

$$(x, Ax) = \sum_{k,l=1}^K u_k u_l (y^k, Ay^l). \quad (1)$$

Рассмотрим матрицу

$$B_K = \|(y^k, Ay^l)\|, \quad k, l = 1, 2, \dots, K. \quad (2)$$

Тогда

$$\begin{aligned} \lambda_j(B_K) &\leq \lambda_j(A), & i = 1, 2, \dots, K, \\ \lambda_{K-j}(B_K) &\geq \lambda_{N-j}(A), & j = 0, 1, 2, \dots, K-1. \end{aligned} \quad (3)$$

Этот результат сразу же следует из теоремы 4.

§ 11. Теорема о представлении. Введем обозначение

$$|A|_k = \lambda_N \lambda_{N-1} \dots \lambda_{N-k+1}. \quad (1)$$

Докажем теперь следующую теорему.

Теорема 6. Если матрица A положительно определенная, то

$$\frac{\pi^{k/2}}{|A|_k^{1/2}} = \max_{R_k} \int_{R_k} e^{-(x, Ax)} dV_k; \quad (2)$$

здесь интегрирование проводится по k -мерному линейному подпространству R_k N -мерного пространства R , объемными элементами которого служат dV_k , а максимум берется по всем R_k .

Доказательство. Легко видеть, что для доказательства достаточно взять (x, Ax) в виде $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_N x_N^2$. Следовательно, мы должны показать, что

$$\frac{\pi^{k/2}}{(\lambda_N \lambda_{N-1} \dots \lambda_{N-k+1})^{1/2}} = \max_{R_k} \int_{R_k} e^{-(\lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_N x_N^2)} dV_k. \quad (3)$$

Обозначим через $V_a(\rho)$ объем области, заданной соотношениями

$$\begin{aligned} \lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_N x_N^2 &\leq \rho, \\ (x, a^i) &= 0, \quad i = 1, 2, \dots, N-k, \end{aligned} \quad (4)$$

где векторы a^i линейно независимы.

Ясно, что

$$V_a(\rho) = \rho^{k/2} V_a(1). \quad (5)$$

Тогда

$$\int_{R_k} = \int_{(x, a^j)=0} e^{-\lambda_1 x_1^2 - \lambda_2 x_2^2 - \dots - \lambda_N x_N^2} dV_k = \int_0^{\infty} e^{-\rho} dV_a(\rho) = \frac{kV_a(1)}{2} \int_0^{\infty} e^{-\rho} \rho^{k/2-1} d\rho. \quad (6)$$

Чтобы завершить доказательство, мы должны показать, что максимум объема $V_a(1)$ достигается тогда, когда соотношения $(x, a^j) = 0$ сводятся к равенствам $x_1 = 0, x_2 = 0, \dots, x_k = 0$. Это, однако, является следствием формулы для объема эллипсоида

$$1 = \sum_{i=N-k+1}^N \lambda_i x_i^2 \text{ и неравенств, содержащихся в теореме 5*}.$$

§ 12. Приближенные методы. Задача определения минимума функционала $\int_0^1 u'^2 dt$ на функциях, удовлетворяющих условиям

$$\int_0^1 q(t) u^2 dt = 1, \quad (1a)$$

$$u(0) = u(1) = 0, \quad (16)$$

является одной из тех, которые можно решать с помощью методов вариационного исчисления **). Используя стандартную процедуру, мы приходим к задаче Штурма—Лиувилля:

требуется найти те значения параметра λ , при которых уравнение

$$u'' + \lambda q(t) u = 0, \quad u(0) = u(1) = 0, \quad (2)$$

имеет нетривиальное решение.

Поскольку в общем случае решения этого уравнения не могут быть выражены через элементарные функции, мы вынуждены прибегать к различным приближенным методам.

Вместо того чтобы рассматривать точное вариационное уравнение (2) и использовать приближенный метод для его решения,

*) Поскольку объем k -мерной сферы радиуса единица равен $\pi^{k/2}/\Gamma(\frac{k}{2}+1)$

(см., например, Д. К. Фаддеев и И. С. Соминский, Сборник задач по высшей алгебре, «Наука», 1972, стр. 283), то правая часть (6) равна

$\frac{\pi^{k/2}}{(\lambda_1 \lambda_2 \dots \lambda_k)^{1/2}}$, где $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$ — собственные значения матрицы $A_k = \|y^s, Ay^s\|$, $s, t = 1, 2, \dots, k$, а y^s образуют ортонормированный базис в R_k . (Прим. ред.)

**) Следует предположить $q(x) > 0$. (Прим. ред.)

Теорема 5. Пусть $\{y^k\}$, $k = 1, 2, \dots, K$, — набор из K ортогональных векторов, и пусть $x = \sum_{k=1}^K u_k y^k$, так что

$$(x, Ax) = \sum_{k, l=1}^K u_k u_l (y^k, Ay^l). \quad (1)$$

Рассмотрим матрицу

$$B_K = \|(y^k, Ay^l)\|, \quad k, l = 1, 2, \dots, K. \quad (2)$$

Тогда

$$\begin{aligned} \lambda_j(B_K) &\leq \lambda_j(A), & i = 1, 2, \dots, K, \\ \lambda_{K-j}(B_K) &\geq \lambda_{N-j}(A), & j = 0, 1, 2, \dots, K-1. \end{aligned} \quad (3)$$

Этот результат сразу же следует из теоремы 4.

§ 11. Теорема о представлении. Введем обозначение

$$|A|_k = \lambda_N \lambda_{N-1} \dots \lambda_{N-k+1}. \quad (1)$$

Докажем теперь следующую теорему.

Теорема 6. Если матрица A положительно определенная, то

$$\frac{\pi^{k/2}}{|A|_k^{1/2}} = \max_{R_k} \int_{R_k} e^{-(x, Ax)} dV_k; \quad (2)$$

здесь интегрирование проводится по k -мерному линейному подпространству R_k N -мерного пространства R , объемными элементами которого служат dV_k , а максимум берется по всем R_k .

Доказательство. Легко видеть, что для доказательства достаточно взять (x, Ax) в виде $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_N x_N^2$. Следовательно, мы должны показать, что

$$\frac{\pi^{k/2}}{(\lambda_N \lambda_{N-1} \dots \lambda_{N-k+1})^{1/2}} = \max_{R_k} \int_{R_k} e^{-(\lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_N x_N^2)} dV_k. \quad (3)$$

Обозначим через $V_\alpha(\rho)$ объем области, заданной соотношениями

$$\begin{aligned} \lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_N x_N^2 &\leq \rho, \\ (x, a^i) &= 0, \quad i = 1, 2, \dots, N-k, \end{aligned} \quad (4)$$

где векторы a^i линейно независимы.

Ясно, что

$$V_\alpha(\rho) = \rho^{k/2} V_\alpha(1). \quad (5)$$

Тогда

$$\int_{R_k} = \int_{(x, a^j)=0} e^{-\lambda_1 x_1^2 - \lambda_2 x_2^2 - \dots - \lambda_N x_N^2} dV_k = \\ = \int_0^{\infty} e^{-\rho} dV_a(\rho) = \frac{kV_a(1)}{2} \int_0^{\infty} e^{-\rho} \rho^{k/2-1} d\rho. \quad (6)$$

Чтобы завершить доказательство, мы должны показать, что максимум объема $V_a(1)$ достигается тогда, когда соотношения $(x, a^j) = 0$ сводятся к равенствам $x_1 = 0, x_2 = 0, \dots, x_k = 0$. Это, однако, является следствием формулы для объема эллипсоида

$$1 = \sum_{i=N-k+1}^N \lambda_i x_i^2 \text{ и неравенств, содержащихся в теореме 5*}.$$

§ 12. Приближенные методы. Задача определения минимума функ-

ционала $\int_0^1 u'^2 dt$ на функциях, удовлетворяющих условиям

$$\int_0^1 q(t) u^2 dt = 1, \quad (1a)$$

$$u(0) = u(1) = 0, \quad (16)$$

является одной из тех, которые можно решать с помощью методов вариационного исчисления**). Используя стандартную процедуру, мы приходим к задаче Штурма—Лиувилля:

требуется найти те значения параметра λ , при которых уравнение

$$u'' + \lambda q(t) u = 0, \quad u(0) = u(1) = 0, \quad (2)$$

имеет нетривиальное решение.

Поскольку в общем случае решения этого уравнения не могут быть выражены через элементарные функции, мы вынуждены прибегать к различным приближенным методам.

Вместо того чтобы рассматривать точное вариационное уравнение (2) и использовать приближенный метод для его решения,

*) Поскольку объем k -мерной сферы радиуса единица равен $\pi^{k/2}/\Gamma(\frac{k}{2}+1)$

(см., например, Д. К. Фаддеев и И. С. Соминский, Сборник задач по высшей алгебре, «Наука», 1972, стр. 283), то правая часть (6) равна

$\frac{\pi^{k/2}}{(\tilde{\lambda}_1 \tilde{\lambda}_2 \dots \tilde{\lambda}_k)^{1/2}}$, где $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \dots \geq \tilde{\lambda}_k$ — собственные значения матрицы $A_k =$

$= \|y^s, Ay^s\|$, $s, t = 1, 2, \dots, k$, а y^s образуют ортонормированный базис в R_k . (Прим. ред.)

**) Следует предположить $q(x) > 0$. (Прим. ред.)

мы всегда можем заменить исходную вариационную задачу *приближенной вариационной задачей*, а затем применить *точный метод* для решения этой новой задачи. Опишем один из таких путей.

Пусть $\{u_i(t)\}$ — последовательность линейно независимых на отрезке $[0, 1]$ функций, удовлетворяющих условиям (16), т. е. $u_i(0) = u_i(1) = 0$. Мы попытаемся найти приближенное решение исходной вариационной задачи в виде

$$u = \sum_{i=1}^N x_i u_i(t). \quad (3)$$

Задача, которая теперь стоит перед нами, конечномерная; в нее входят неизвестные x_1, x_2, \dots, x_N . Нам нужно минимизировать квадратичную форму

$$\sum_{i,j=1}^N x_i x_j \int_0^1 u_i'(t) u_j'(t) dt \quad (4)$$

при условии

$$\sum_{i,j=1}^N x_i x_j \int_0^1 q(t) u_i(t) u_j(t) dt = 1. \quad (5)$$

Ясно, что мы получим существенные упрощения, если выберем последовательность $\{u_i(t)\}$ такой, что

$$\int_0^1 u_i'(t) u_j'(t) dt = \delta_{ij}, \quad \text{либо} \quad \int_0^1 q(t) u_i(t) u_j(t) dt = \delta_{ij}. \quad (6)$$

Первому условию обычно легче удовлетворить. Так, например, мы можем взять

$$u_k(t) = \sin \pi k t, \quad (7)$$

нормированную соответствующим образом, или

$$u_k(t) = P_k(t), \quad (8)$$

где $P_k(t)$ — k -й полином Лежандра, опять же подходящим образом нормированный. Последний из указанных вариантов особенно удобен, если $q(t)$ является полиномом по t , так как интегралы

вида $\int_0^1 t^k u_i(t) u_j(t) dt$ легко вычисляются, если $u_i(t)$ заданы формулой (8).

Описанная процедура сопряжена с целым рядом интересных и важных проблем. Первостепенное значение имеет вопрос о сходимости решения конечномерной задачи к решению исходной задачи.

Кроме того, представляет интерес скорость сходимости (вопрос большой практической важности), а также характер сходимости: является ли она монотонной, колебательной и т. д. Здесь, однако, мы не будем углубляться в эти вопросы.

Упражнение

1. Пусть $\lambda_1^{(N)}, \lambda_2^{(N)}, \dots, \lambda_N^{(N)}$ — характеристические числа, возникающие в задаче, постановка которой описывается формулами (4) и (5) при $N=2, 3, \dots$. Какие неравенства связывают $\lambda_i^{(N)}$ и $\lambda_i^{(N-1)}$?

Упражнения к гл. 7

1. Пусть A и B — эрмитовы матрицы, характеристические числа которых соответственно равны $\lambda_1 \geq \lambda_2 \geq \dots; \mu_1 \geq \mu_2 \geq \dots$, и пусть $\nu_1 \geq \nu_2 \geq \dots$ — характеристические числа $A+B$. Показать, что $\lambda_i + \mu_j \geq \nu_{i+j-1}$ для $i+j \leq N+1$ (Вейль).

2. Приведением A к диагональной форме дать индуктивное доказательство теоремы отделения Пуанкаре (см. теорему 5 настоящей главы).

3. В чем состоит необходимое и достаточное условие того, что $a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \geq 0$ при всех $x_1, x_2 \geq 0$?

4. Указать необходимое и достаточное условие, при котором

$$\sum_{i,j=1}^3 a_{ij}x_i x_j \geq 0 \text{ для всех } x_1, x_2, x_3 \geq 0?$$

5. Можно ли получить соответствующие условия для $\sum_{i,j=1}^N a_{ij}x_i x_j$?

6. Показать, что

$$A(r) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

— положительно определенная матрица для $r > 0$, если $a_{ij} \geq 0$, и матрица $A = \|a_{ij}\|$ является положительно определенной.

7. Справедливо ли предыдущее утверждение для 3×3 -матриц?

8. Доказать, что если $a_{ii} > 0$ и матрица $\| |a_{ij}| \|$ положительно определенная, то матрица $\| a_{ij} \|$ не обязательно положительно определенная.

9. Показать, что характеристические числа матрицы $A^{1/2}BA^{1/2}$ меньше характеристических чисел матрицы A , если B — симметрическая матрица, все собственные значения которой расположены между 0 и 1, а матрица A положительно определенная.

10. Справедливо ли неравенство $(Tx, Ax) \leq (x, Ax)$ для всех действительных x , если A — положительно определенная матрица, а T — ортогональная матрица?

11. Определим ранг симметрической матрицы как разность между ее порядком и числом нулевых собственных значений. Используя результаты этой главы, показать эквивалентность этого определения и определения, данного в приложении А.

12. Предположим, что мы имеем некоторое семейство $\{A(q)\}$ действительных симметрических матриц, зависящих от параметра q , скалярного или векторного. Мы хотим определить максимальное характеристическое число

*) См. Ф. Рисс и Б. Секефальви Надь, Лекции по функциональному анализу, ИЛ, 1954, гл. VI, § 1. (Прим. ред.)

каждой матрицы, которое обозначим через $f(q)$, а затем найти максимум по q этой функции. Мы поступим следующим образом. Выберем некоторое начальное значение параметра q , скажем q_0 , и пусть x^0 — собственный вектор, принадлежащий характеристическому числу $f(q_0)$. Следующее значение параметра мы выберем, максимизируя выражение $(x^0, A(q)x^0)/(x^0, x^0)$ по всем значениям q . Обозначим одно из значений параметра, дающее максимум, через q_1 . Пусть x^1 — собственный вектор, принадлежащий характеристическому числу $f(q_1)$. Продолжая таким образом, мы получим последовательность $f(q_0), f(q_1), \dots$. Показать, что

$$f(q_0) < f(q_1) < f(q_2) < \dots$$

13. Показать на примере, что описанная процедура не всегда позволяет получить $\max_q f(q)$.

14. При каких условиях, наложенных на матрицу $A(q)$, мы можем гарантировать достижение абсолютного максимума?

15. Определим спектральный радиус $r(A)$ квадратной матрицы A как наибольший из модулей ее характеристических чисел. Пусть H — положительно определенная эрмитова матрица, и пусть

$$g(A, H) = \max_x [(Ax, HAx)/(x, Hx)]^{1/2}.$$

Показать, что $r(A) = \min_H g(A, H)$.

16. Показать, что если матрица A действительная, то $r(A) = \min_S g(A, S)$, где теперь минимум берется по всем действительным симметрическим матрицам. [Осборн (H. Osborn).]

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 1. Р. Курант обобщил минимаксное представление на случай дифференциальных операторов с частными производными. Это обобщение легло в основу современной теории собственных значений операторов. Обсуждение этих вопросов можно найти в монографии

Курант, Гильберт (R. Courant, D. Hilbert), *Methods of Mathematical Physics*, New York, Interscience Publishers, 1953. [Русский перевод второго немецкого издания: *Методы математической физики*, т. I, II, Гостехиздат, 1951; см. также перевод американского издания 1961 г.: Курант, Уравнения с частными производными, «Мир», 1964.]

§ 2. Отношение Релея позволяет весьма просто получать верхние границы для наименьшего характеристического числа. Задача нахождения нижних границ связана со значительно большими трудностями. А. Вайнштейн развил сильный метод для случая более общих операторов; этот метод обобщили Н. Ароншайн и др. Изложение и дальнейшие ссылки имеются в статье

Днас (J. B. Diaz), *Upper and Lower Bounds for Eigenvalues*, *Calculus of Variations and Its Applications*, L. M. Graves (ed.), McGraw-Hill Book Company, Inc., New York, 1959.

См. также работу

Темпл и Бикли (G. Temple and W. G. Bickley), *Rayleigh's Principle and Its Applications to Engineering*, Oxford, 1933

Задача определения интервалов, внутри которых могут или, напротив, не могут лежать характеристические числа, имеет важное значение. См. работы

Като (Т. Kato), *On the Upper and Lower Bounds of Eigenvalues*, J. Phys. Soc. Japan 4 (1949), 334—339;

Бейтмен (H. Bateman), *Trans. Cambridge Phil. Soc.* 20 (1908), 371—382;

Темпл (G. Temple), *An Elementary Proof of Kato's Lemma*, *Matematika* 2 (1955), 39—41.

§ 6. См. цитированную выше монографию Куранта и Гильберта, а также

Фишер (E. Fisher), *Über quadratische Formen mit reellen Koeffizienten*, *Monatsh. Math. u. Physik* 16 (1905), 234—249.

Обсуждение и некоторые приложения можно найти в статье

Поля (G. Polya), *Estimates for Eigenvalues*, *Studies in Mathematics and Mechanics*, presented to R. Von Mises, Academic Press, Inc., New York, 1954.

Получены существенные обобщения этого результата. Некоторые из них, принадлежащие Фань Цзы, приведены в гл. 7, §§ 10, 11 и 13. См.

Фань Цзы (Ku Fan), *On a Theorem of Weyl Concerning Eigenvalues of Linear Transformations, I*, *Proc. Nat. Acad. Sci. U. S. A.* 35 (1955), 652—655.

Другие обобщения получены Виландтом:

Виландт (H. Wielandt), *An Extremum Property of Sums of Eigenvalues*, *Proc. Amer. Math. Soc.* 6 (1955), 106—110*).

См. далее

Амир Моез (Ali R. Amir-Moez), *Extreme Properties of Eigenvalues of a Hermitian Transformations and Singular Values of the Sum and Product of a Linear Transformation*, *Duke Math. J.* 23 (1956), 463—477;

Маркус и Томпсон (M. Marcus and R. Thompson), *A Note on Symmetric Functions of Eigenvalues*, *Duke Math. J.* 24 (1957), 43—46;

Маркус (M. Marcus), *Convex Functions of Quadratic Forms*, *Duke Math. J.* 24 (1957), 321—325.

* См. также А. С. Маркус, *Собственные и сингулярные числа суммы и произведения линейных операторов*. УМН 19, № 4 (118) (1965), и Ф. Р. Гантмахер, *Теория матриц*, «Наука», 1967, Дополнение. (Прим. ред.)

§ 8. Результаты этого типа можно получить классическим методом Штурма. См. книгу

Бернсайда и Пэнтон (W. S. Burnside and A. W. Panton), *Theory of Equations*, vol. 11, Longmans, Green & Co., New York, 1928.

§ 10.

Пуанкаре (H. Poincaré), *Sur les equations aux derivées partielles de la physique mathématique*, Amer. J. Math. 12 (1890), 211—294.

§ 11.

Беллман, Гликсберг и Гросс (R. Bellman, I. Glicksberg and O. Gross), *Notes on Matrix Theory—IV*, Amer. Math. Monthly 62 (1955), 571—572.

§ 12. Детальное изложение вычислительных методов этого типа содержится в книге

Коллатц (L. Collatz), *Eigenwertproblemen und ihre numerische Behandlung*, New York, 1948 [русский перевод: *Задачи на собственные значения*, «Мир», 1968].

Упомянем, наконец, монографию

Пароди (M. Parodi), *Sur quelques propriétés valeurs caractéristiques des matrices carrées*, fascicule CXVII, Mem. sci. mathem. 1952. [См. также М. Пароди, *Локализация характеристических чисел матриц и ее применения*, ИЛ, 1960.]

Задача нахождения соответствующего вариационного описания характеристических чисел комплексных матриц, возникающих при исследовании линейных диссипативных систем, чрезвычайно трудна. Некоторые шаги в этом направлении описаны в статье

Даффин (R. J. Duffin), *A Minimax Theory for Overdamped Networks*, J. Rat. Mech. Analysis 4 (1955), 221—233.

См. также результаты Долфа и др., на которые мы ссылались выше.

НЕРАВЕНСТВА

§ 1. Введение. В этой главе мы установим ряд интересных неравенств, касающихся характеристических чисел и детерминантов симметрических матриц. Нашими основными орудиями будут интегральное тождество

$$\frac{\pi^{N/2}}{|A|^{1/2}} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-i(x, Ax)} dx, \quad (1)$$

справедливое для положительно определенных матриц, его обобщение, данное в теореме 6 гл. 7, а также некоторые обобщения минимаксного представления Куранта—Фишера, принадлежащие Фань Цзы.

Прежде чем вывести некоторые неравенства, непосредственно относящиеся к матрицам, мы установим стандартные неравенства Коши—Шварца и Гёльдера. Далее мы докажем неравенство о среднем арифметическом и среднем геометрическом, так как оно понадобится нам для одного из результатов этой главы.

§ 2. Неравенство Коши—Шварца. Первый результат, который мы установим, уже встречался в упражнениях. Однако он вполне заслуживает того, чтобы сформулировать его еще раз.

Теорема 1. Для любых двух действительных векторов x и y имеет место неравенство

$$(x, y)^2 \leq (x, x)(y, y). \quad (1)$$

Доказательство. Рассмотрим квадратичную форму по переменным u и v

$$Q(u, v) = (ux + vy, ux + vy) = u^2(x, x) + 2uv(x, y) + v^2(y, y). \quad (2)$$

Поскольку форма $Q(u, v)$, очевидно, неотрицательна, то неравенство (1) выполняется с необходимостью. Заметим, что неравенство (1) является хотя и частным, но наиболее важным случаем неотрицательности определителя Грама, установленного в § 4 гл. 4.

§ 3. Интегральный вариант. Точно таким же образом можно установить интегральный аналог предыдущего неравенства.

Теорема 2. Пусть $f(x)$ и $g(x)$ — функции аргумента x , определенные на некоторой области R . Тогда

$$\left(\int_R fg \, dV\right)^2 \leq \left(\int_R f^2 \, dV\right) \left(\int_R g^2 \, dV\right). \quad (1)$$

Доказательство. Так как $(f \pm g)^2 \geq 0$, то мы имеем

$$2|fg| \leq f^2 + g^2. \quad (2)$$

Следовательно, произведение fg интегрируемо, если интегрируемы f^2 и g^2 . Рассмотрим теперь интеграл

$$\int_R (fu + gv)^2 \, dV, \quad (3)$$

который в силу предыдущего неравенства существует. Относительно u и v интеграл (3) представляет собой неотрицательно определенную квадратичную форму. Следовательно, как и выше, имеет место (1).

Упражнения

1. Можно также поступить так:

$$(x, x)(y, y) - (x, y)^2 = (y, y) \left[(x, x) - 2 \frac{(x, y)^2}{(y, y)} + \frac{(x, y)^2}{(y, y)} \right] = \\ = (y, y) \left(x - y \frac{(x, y)}{(y, y)}, x - y \frac{(x, y)}{(y, y)} \right) \geq 0.$$

2. Доказать, что $\left(\sum_{k=1}^N x_k y_k\right)^2 < \left(\sum_{k=1}^N \lambda_k x_k^2\right) \left(\sum_{k=1}^N y_k^2 / \lambda_k\right)$, если числа λ_k

положительны.

3. Показать, используя предыдущий результат, что $(x, Ax)(y, A^{-1}y) \geq (x, y)^2$, где матрица A положительно определенная. Установить этот результат, не опираясь на диагональную форму.

§ 4. Неравенство Гёльдера. Следующая теорема обобщает неравенство (1) § 2.

Теорема 3. Пусть $p > 1$ и $q = p/(p-1)$; тогда

$$\sum_{k=1}^N x_k y_k \leq \left(\sum_{k=1}^N x_k^p\right)^{1/p} \left(\sum_{k=1}^N y_k^q\right)^{1/q} \quad (1)$$

для всех $x_k, y_k \geq 0$.

Доказательство. Рассмотрим кривую

$$v = u^{p-1}, \quad (2)$$

где $p > 1$.

Ясно, что площадь прямоугольника $OvRu$ меньше или равна сумме площадей OPu и OQv :

$$uv \leq \int_0^u u_1^{p-1} du_1 + \int_0^v v_1^{(p-1)} dv_1, \quad (3)$$

причем равенство выполняется только в том случае, если $v = u^{p-1}$. Следовательно, если $u, v \geq 0, p > 1$, то имеет место неравенство

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}. \quad (4)$$

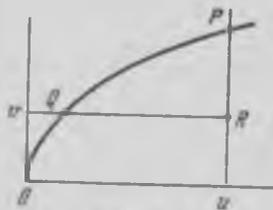


Рис. 2.

Теперь, последовательно полагая

$$u = x_k / \left(\sum_{k=1}^N x_k^p \right)^{1/p}, \quad v = y_k / \left(\sum_{k=1}^N y_k^q \right)^{1/q}, \quad (5)$$

$k = 1, 2, \dots, N$, и суммируя по k , мы приходим к следующему результату:

$$\frac{\sum_{k=1}^N x_k y_k}{\left(\sum_{k=1}^N x_k^p \right)^{1/p} \left(\sum_{k=1}^N y_k^q \right)^{1/q}} \leq \frac{1}{p} \frac{\sum_{k=1}^N x_k^p}{\sum_{k=1}^N x_k^p} + \frac{1}{q} \frac{\sum_{k=1}^N y_k^q}{\sum_{k=1}^N y_k^q}. \quad (6)$$

Так как $1/p + 1/q = 1$, то получаем неравенство (1).

Полагая в (4)

$$u = \frac{f(x)}{\left(\int_R f(x)^p dV \right)^{1/p}}, \quad v = \frac{g(x)}{\left(\int_R g(x)^q dV \right)^{1/q}}, \quad (7)$$

мы имеем

$$\frac{\int_R f(x) g(x) dV}{\left(\int_R f(x)^p dV \right)^{1/p} \left(\int_R g(x)^q dV \right)^{1/q}} \leq \frac{1}{p} \frac{\int_R f(x)^p dV}{\int_R f(x)^p dV} + \frac{1}{q} \frac{\int_R g(x)^q dV}{\int_R g(x)^q dV}. \quad (8)$$

Интегрируя по области R , получаем неравенство

$$\int_R f(x) g(x) dV \leq \left(\int_R f(x)^p dV \right)^{1/p} \left(\int_R g(x)^q dV \right)^{1/q}, \quad (9)$$

справедливое при условии, что $f(x), g(x) \geq 0$ и интегралы в правой части существуют.

§ 5. Вогнутость $|A|$. Теперь мы выведем некоторые следствия из интегрального тождества (1.1). Первыми из них будет

Теорема 4. Если матрицы A и B положительно определенные, то

$$|\lambda A + (1-\lambda)B| \geq |A|^\lambda |B|^{1-\lambda} \quad (1)$$

при $0 \leq \lambda \leq 1$.

Доказательство. Имеем

$$\frac{\pi^{N/2}}{|\lambda A + (1-\lambda)B|^{1/2}} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\lambda(x, Ax) - (1-\lambda)(x, Bx)} dx. \quad (2)$$

Используем теперь интегральную формулу неравенства Гёльдера, положив $p = 1/\lambda$, $q = 1/(1-\lambda)$. Тогда:

$$\begin{aligned} \frac{\pi^{N/2}}{|\lambda A + (1-\lambda)B|^{1/2}} &\leq \left(\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-(x, Ax)} dx \right)^\lambda \left(\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-(x, Bx)} dx \right)^{1-\lambda} = \\ &= \frac{\lambda^{N/2} \pi^{N(1-\lambda)/2}}{|\lambda|^{N/2} |B|^{(1-\lambda)/2}}. \end{aligned} \quad (3)$$

Этот результат после упрощений совпадает с (1); в свою очередь неравенство (1) является частным случаем более общего результата, который мы выведем позже.

Упражнения

1. Доказать, что $|A + iB| \geq |A|$, если матрица A положительно определенная, а матрица B действительная симметрическая.
2. Показать, что $\lambda x + (1-\lambda)y \geq x^\lambda y^{1-\lambda}$ при $x, y > 0$, $0 < \lambda \leq 1$.

§ 6. Одно полезное неравенство. Впоследствии мы воспользуемся результатом, который содержится в следующей теореме:

Теорема 5. Если матрица A положительно определенная, то

$$|A| \leq a_{11} a_{22} \dots a_{NN}. \quad (1)$$

Приведем два доказательства этой теоремы.

Первое доказательство. $|A|$ можно представить в виде

$$|A| = a_{11} \begin{vmatrix} a_{11} & \dots & a_{1N} \\ a_{21} & \dots & a_{2N} \\ \vdots & & \vdots \\ a_{N1} & \dots & a_{NN} \end{vmatrix} + \begin{vmatrix} 0 & a_{11} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{vmatrix}. \quad (2)$$

Так как A положительно определенная, то и матрица $\|a_{ij}\|$, $i, j = 2, \dots, N$, тоже положительно определенная, следовательно, квадратичная форма по переменным $a_{12}, a_{13}, \dots, a_{1N}$, каковой является второе слагаемое в правой части (2), отрицательно определенная*). В силу этого

$$|A| \leq a_{11} \begin{vmatrix} a_{22} & \dots & a_{2N} \\ a_{32} & \dots & a_{3N} \\ \dots & \dots & \dots \\ a_{N2} & \dots & a_{NN} \end{vmatrix}, \quad (3)$$

откуда, применяя индукцию, получаем доказываемый результат.

Второе доказательство. Рассмотрим равенство (1.1), положив в нем $N = 3$. Заменяя x_1 на $-x_1$ и складывая, получаем

$$\frac{\pi^{3/2}}{|A|^{1/2}} = \int_{-\infty}^{\infty} e^{-a_{11}x_1^2 - 2a_{12}x_1x_2 - a_{22}x_2^2 - a_{11}x_1^2 \left(\frac{z+z^{-1}}{2}\right)} dx_1 dx_2 dx_3, \quad (4)$$

где

$$z = e^{-2a_{11}x_1x_2 - 2a_{12}x_1x_3}. \quad (5)$$

Так как $z + z^{-1} \geq 2$ для всех $z \geq 0$, то

$$\begin{aligned} \frac{\pi^{3/2}}{|A|^{1/2}} &\geq \left(\int_{-\infty}^{\infty} e^{-(a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2)} dx_1 dx_2 \right) \left(\int_{-\infty}^{\infty} e^{-a_{11}x_1^2} dx_1 \right) = \\ &= \frac{\pi^{1/2}}{|a_{11}|^{1/2}} \frac{\pi}{|A_2|^{1/2}}, \end{aligned} \quad (6)$$

где $A_2 = \|a_{ij}\|$, $i, j = 2, 3$. Следовательно,

$$|a_{11}|^{1/2} |A_2|^{1/2} \geq |A|^{1/2}, \quad (7)$$

откуда следующим шагом получаем искомый результат для $N = 3$. Для произвольного N доказательство проводится аналогичным образом.

Упражнение

1. Пусть $D_1 = |a_{ij}|$, $i, j = 1, 2, \dots, n_1$, $D_2 = |a_{ij}|$, $i, j = n_1 + 1, \dots, n_2, \dots$, $D_r = |a_{ij}|$, $i, j = n_{r-1} + 1, \dots, N$. Тогда $|A| \leq D_1 D_2 \dots D_r$.

§ 7. Неравенство Адамара. Наиболее известное из детерминантных неравенств принадлежит Адамару.

*) См. § 6 гл. 5 настоящей книги. (Прим. ред.)

Теорема 6. Пусть B — произвольная невырожденная действительная квадратная матрица. Тогда

$$|B|^n \leq \prod_{i=1}^N \left(\sum_{k=1}^N b_{ik}^2 \right). \quad (1)$$

Для доказательства следует применить теорему 5 к положительно определенной квадратной матрице $A = BB'$.

§ 8. Вогнутость произведения $\lambda_N \lambda_{N-1} \dots \lambda_k$. Определим функцию $|A|_k$ элементов матрицы A формулой

$$|A|_k = \lambda_N \lambda_{N-1} \dots \lambda_k. \quad (1)$$

Можно доказать следующую теорему:

Теорема 7. Если матрицы A и B положительно определены, то при $0 \leq \lambda \leq 1$, $k = 1, 2, \dots, N$ имеем

$$|\lambda A + (1-\lambda)B|_k \geq |A|_k^\lambda |B|_k^{1-\lambda}. \quad (2)$$

Доказательство следует из теоремы 6 гл. 7 в точности тем же путем, которым была получена теорема 4 из соотношения (1.1).

Упражнения

1. Показать, что

$$\varphi_i(A) = \frac{|A|}{|A_i|} = \min_x (x, Ax),$$

где x удовлетворяет условию $x_i = 1$, а A_i — матрица, получающаяся из матрицы A после вычеркивания i -й строки и i -го столбца*).

2. Отправляясь от результата предыдущего упражнения, показать, что при условии $A > 0$ и $B > 0$ справедливо неравенство

$$\varphi_i[\lambda A + (1-\lambda)B] \geq \varphi_i(A)^\lambda \varphi_i(B)^{1-\lambda},$$

$0 \leq \lambda \leq 1$. (Неравенство Бергстрема.)

3. Пусть A и B — две положительно определенные матрицы порядка N , и пусть $C = \lambda A + (1-\lambda)B$, $0 \leq \lambda \leq 1$. Пусть A^j , $j = 1, 2, \dots, N$, обозначает подматрицу матрицы A , получаемую вычеркиванием первых $j-1$ строк и столбцов. Если k_1, k_2, \dots, k_N представляют собой N положительных чисел

таких, что $\sum_{i=1}^N k_i \geq 0$, то

$$\prod_{i=1}^N |C^i|^{k_i} \geq \prod_{i=1}^N |A^i|^{\lambda k_i} |B^i|^{(1-\lambda)k_i}. \quad (\text{Фань Цзы.})$$

4. Доказать неравенство Адамара для эрмитовых матриц.

* Матрица A предполагается положительно определенной. (Прим. ред.)

5. Если A — положительно определенная матрица размера $N \times N$ и P_k обозначает произведение всех главных миноров k -го порядка матрицы A , то

$$P_1 \geq P_2^{1/C_{N-1}} \geq P_3^{1/C_{N-1}^2} \geq \dots \geq P^{1/C_{N-1}^{N-2}} \geq P_N. \quad (\text{Сас.})$$

См. работу Мирского¹⁾.

§ 9. Аддитивные неравенства, вытекающие из мультипликативных. Сейчас мы приведем пример, иллюстрирующий математический принцип, согласно которому каждое мультипликативное неравенство имеет аддитивный аналог.

Применим теорему 7 к матрицам

$$A = I + eX, \quad B = I + eY, \quad e > 0, \quad (1)$$

где X и Y — произвольные симметрические матрицы. Если e достаточно мало, то A и B будут положительно определенными.

Пусть $x_1 \geq x_2 \geq \dots \geq x_N$, $y_1 \geq y_2 \geq \dots \geq y_N$ — характеристические числа матриц X и Y соответственно и $z_1 \geq z_2 \geq \dots \geq z_N$ — характеристические числа матрицы $\lambda X + (1 - \lambda)Y$. В этом случае применение теоремы 7 дает

$$(1 + ez_N)(1 + ez_{N-1}) \dots (1 + ez_k) \geq \\ \geq [(1 + ex_N)(1 + ex_{N-1}) \dots (1 + ex_k)]^\lambda \times \\ \times [(1 + ey_N)(1 + ey_{N-1}) \dots (1 + ey_k)]^{1-\lambda}. \quad (2)$$

Для членов первого порядка относительно e это соотношение дает

$$1 + e(z_N + z_{N-1} + \dots + z_k) \geq 1 + \lambda e(x_N + x_{N-1} + \dots + x_k) + \\ + (1 - \lambda)e(y_N + y_{N-1} + \dots + y_k) + o(e^2). \quad (3)$$

Поскольку e произвольно мало, то

$$z_N + z_{N-1} + \dots + z_k \geq \lambda(x_N + \dots + x_k) + \\ + (1 - \lambda)(y_N + y_{N-1} + \dots + y_k). \quad (4)$$

Сформулируем этот результат, принадлежащий Фань Цзы, в следующей форме:

Теорема 8. Пусть A — симметрическая матрица и

$$S_k(A) = \lambda_N + \lambda_{N-1} + \dots + \lambda_k. \quad (5)$$

Тогда при $0 \leq \lambda \leq 1$, $k = 1, 2, \dots, N$ имеем

$$S_k[\lambda A + (1 - \lambda)B] \geq \lambda S_k(A) + (1 - \lambda)S_k(B). \quad (6)$$

Заменив A на $-A$, мы получим следующее утверждение:

¹⁾ L. Mirsky, On a Generalisation of Hadamard's Determinantal Inequality Due to Szasz, Arch. Math. 8 (1957), 274—275.

Теорема 9. Если

$$T_k(A) = \lambda_1 + \lambda_2 + \dots + \lambda_k, \quad (7)$$

где A — симметрическая матрица, то

$$T_k[\lambda A + (1-\lambda)B] \leq \lambda T_k(A) + (1-\lambda)T_k(B) \quad (8)$$

при $0 \leq \lambda \leq 1$, $k = 1, 2, \dots, N$.

§ 10. Другой путь. Получим эти результаты иным способом, который позволит нам попутно выявить некоторые свойства, представляющие самостоятельный интерес. Начнем с доказательства следующей теоремы:

Теорема 10. Если матрица A положительно определенная, то

$$\lambda_1 \lambda_2 \dots \lambda_{N-k+1} = \max_R |(z^i, Az^j)|, \quad (1)$$

$$\lambda_N \lambda_{N-1} \dots \lambda_k = \min_R |(z^i, Az^j)|,$$

где R — область, определенная соотношениями

$$(z^i, z^j) = \delta_{ij}, \quad i, j = 1, 2, \dots, N-k+1. \quad (2)$$

Другими словами, минимум берется по всем совокупностям из $N-k+1$ ортонормированных векторов.

Доказательство. Рассмотрим определитель

$$D_1(x, y) = \begin{vmatrix} (x, Ax) & (x, Ay) \\ (x, Ay) & (y, Ay) \end{vmatrix} = \begin{vmatrix} \sum_{k=1}^N \lambda_k u_k^2 & \sum_{k=1}^N \lambda_k u_k v_k \\ \sum_{k=1}^N \lambda_k u_k v_k & \sum_{k=1}^N \lambda_k v_k^2 \end{vmatrix}. \quad (3)$$

где $x = \sum_{k=1}^N u_k x^k$, $y = \sum_{k=1}^N v_k x^k$, а x^k — собственные векторы матрицы A .

Тождество Лагранжа дает

$$D_1(x, y) = \frac{1}{2} \sum_{i, j=1}^N \lambda_i \lambda_j (u_i v_j - u_j v_i)^2. \quad (4)$$

Члены с $i = j$ сюда не входят, так как тогда $u_i v_j - u_j v_i = 0$.

Легко видеть, что

$$\frac{1}{2} \lambda_1 \lambda_2 \sum_{i, j=1}^N (u_i v_j - u_j v_i)^2 \geq D_1(x, y) \geq \frac{1}{2} \lambda_N \lambda_{N-1} \sum_{i, j=1}^N (u_i v_j - u_j v_i)^2. \quad (5)$$

Следовательно,

$$\lambda_1 \lambda_2 \left| \begin{array}{cc} \sum_{i=1}^N u_i^2 & \sum_{i=1}^N u_i v_i \\ \sum_{i=1}^N u_i v_i & \sum_{i=1}^N v_i^2 \end{array} \right| \geq D_2(x, y) \geq \lambda_N \lambda_{N-1} \left| \begin{array}{cc} \sum_{i=1}^N u_i^2 & \sum_{i=1}^N u_i v_i \\ \sum_{i=1}^N u_i v_i & \sum_{i=1}^N v_i^2 \end{array} \right|, \quad (6)$$

или

$$\lambda_1 \lambda_2 \geq D_2(x, y) \geq \lambda_N \lambda_{N-1}. \quad (7)$$

Это доказывает результат для $k = N - 1$. В общем случае следует воспользоваться тождеством для определителя Грама, приведенным в § 5 гл. 4, а не тождеством Лагранжа *).

Упражнение

1. Используя теоремы 9 и 4, доказать теорему 7.

§ 11. Более простое выражение для $\lambda_N \lambda_{N-1} \dots \lambda_k$. Так как определитель является относительно сложной функцией, желательно получить представление более простое, чем установленное в теореме 10. С этой целью мы докажем теорему 11.

Теорема 11. Если матрица A является положительно определенной, то

$$\lambda_N \lambda_{N-1} \dots \lambda_k = \min_R (z^1, Az^1)(z^2, Az^2) \dots (z^{N-k+1}, Az^{N-k+1}), \quad (1)$$

где R определяется так же, как и в (10.2).

Доказательство. Мы знаем, что из положительной определенности матрицы A вытекает положительная определенность матрицы (z^i, Az^j) , $i, j = 1, 2, \dots, N - k + 1$. Тогда из теоремы 5 получаем

$$|(z^i, Az^j)| \leq \prod_{i=1}^{N-k+1} (z^i, Az^i). \quad (2)$$

Следовательно,

$$\min_R |(z^i, Az^j)| \leq \min_R \prod_{i=1}^{N-k+1} (z^i, Az^i). \quad (3)$$

Выбирая $z^i = x^i$, $i = N, N - 1, \dots, N - k + 1$, мы видим, что

$$\min_R \prod_{i=1}^{N-k+1} (z^i, Az^i) \leq \lambda_N \lambda_{N-1} \dots \lambda_k. \quad (4)$$

Этот результат совместно с теоремой 10 дает теорему 11.

*) По существу автор предлагает новое доказательство соотношений (1), поскольку они прямо следуют из теоремы 5 гл. 7 (стр. 142). (Прим. ред.)

Упражнения

1. Используя предыдущий результат, доказать теорему 7.
 2. Рассмотреть случай $k=N$ и с помощью специального выбора z^1 показать, что

$$\lambda_N < \min_i a_{ii}.$$

3. В общем случае показать, что

$$\lambda_N \lambda_{N-1} \dots \lambda_k < a_{NN} a_{N-1, N-1} \dots a_{kk}.$$

4. Показать, что $\min_B \text{Sp}(AB)/N = |A|^{1/N}$, где минимум берется по всем B , удовлетворяющим условиям $B > 0$, $|B|=1$, а матрица A положительно определенная.

5. Отправляясь от результата предыдущего упражнения, показать, что $|A+B|^{1/N} \geq |A|^{1/N} + |B|^{1/N}$, если матрицы A и B положительно определенные. (А. Минковский.)

§ 12. Неравенство между средним арифметическим и средним геометрическим. В дальнейшем нам понадобится фундаментальное неравенство, касающееся среднего арифметического и среднего геометрического.

Теорема 12. Если $x_i \geq 0$, то

$$\sum_{i=1}^N x_i / N \geq (x_1 x_2 \dots x_N)^{1/N}. \quad (1)$$

Доказательство. Хотя этот результат можно легко установить методами анализа, мы приведем доказательство, близкое по духу тем методам, на которые мы опирались до сих пор¹⁾. Отправляясь от неравенства $(a_1^{1/2} - a_2^{1/2})^2 \geq 0$, или, что то же,

$$\frac{a_1 + a_2}{2} \geq (a_1 a_2)^{1/2}, \quad (2)$$

положим

$$a_1 = (b_1 + b_2)/2, \quad a_2 = (b_3 + b_4)/2. \quad (3)$$

В результате, принимая во внимание (2), получаем

$$\frac{b_1 + b_2 + b_3 + b_4}{4} \geq \left(\frac{b_1 + b_2}{2}\right)^{1/2} \left(\frac{b_3 + b_4}{2}\right)^{1/2} \geq (b_1 b_2 b_3 b_4)^{1/4}. \quad (4)$$

Продолжая таким образом, убеждаемся, что (1) справедливо для $N=2^k$, $k=1, 2, \dots$. Чтобы закончить доказательство, мы воспользуемся обратной индукцией и покажем, что неравенство (1) выполняется для $N-1$, если оно справедливо для N . Для этого положим

$$y_1 = x_1, \quad y_2 = x_2, \quad \dots, \quad y_{N-1} = x_{N-1}, \quad y_N = \frac{x_1 + x_2 + \dots + x_{N-1}}{N-1} \quad (5)$$

¹⁾ Это доказательство принадлежит Коши.

и воспользуемся неравенством (1). В результате после несложных преобразований получаем то же неравенство для $N-1$.

Прослеживая этапы доказательства, мы видим, что равенство имеет место только при $x_1 = x_2 = \dots = x_N$.

§ 13. Мультипликативные неравенства, вытекающие из аддитивных. Ранее мы показали, как вывести аддитивные неравенства из мультипликативных. Теперь мы покажем, как можно провести обратную дедукцию.

Теорема 13. Из неравенства *)

$$\sum_{i=k}^N \lambda_i \leq \sum_{i=k}^N (x^i, Ax^i), \quad (1)$$

$k=1, 2, \dots, N$, справедливого для любого ортонормированного набора векторов $\{x^i\}$, вытекает неравенство

$$\prod_{i=k}^N \lambda_i \leq \prod_{i=k}^N (x^i, Ax^i). \quad (2)$$

Доказательство. Рассмотрим сумму

$$\begin{aligned} c_k \lambda_k + c_{k+1} \lambda_{k+1} + \dots + c_N \lambda_N &= \\ &= c_k (\lambda_k + \lambda_{k+1} + \dots + \lambda_N) + \\ &+ (c_{k+1} - c_k) (\lambda_{k+1} + \dots + \lambda_N) + \dots + (c_N - c_{N-1}) \lambda_N. \end{aligned} \quad (3)$$

Пусть

$$0 < c_k \leq c_{k+1} \leq \dots \leq c_N, \quad \prod_{i=k}^N c_i = 1. \quad (4)$$

Тогда, используя (1), получаем неравенство

$$\begin{aligned} c_k \lambda_k + c_{k+1} \lambda_{k+1} + \dots + c_N \lambda_N &\leq \\ &\leq c_k (x^k, Ax^k) + c_{k+1} (x^{k+1}, Ax^{k+1}) + \dots + c_N (x^N, Ax^N), \end{aligned} \quad (5)$$

выполняющееся для любого ортонормированного набора векторов $\{x^i\}$ и скалярных констант, удовлетворяющих (4). Из последнего неравенства получаем

$$\min_R \left(\sum_{i=k}^N c_i \lambda_i \right) \leq \min_R \left[\sum_{i=k}^N c_i (x^i, Ax^i) \right], \quad (6)$$

где R — множество в c -пространстве, определяемое соотношениями (4).

*) Неравенство (1) следует из теоремы 5 гл. 7, поскольку правая его часть равна сумме собственных значений матрицы $B_{N-k+1} = \|(x^i, Ax^j)\|$, $i, j = k, k+1, \dots, N$. (Прим. ред.)

Из теоремы 12 следует, что

$$\frac{1}{N-k+1} \sum_{i=k}^N c_i \lambda_i \geq (c_k c_{k+1} \dots c_N)^{1/(N-k+1)} (\lambda_k \lambda_{k+1} \dots \lambda_N)^{1/(N-k+1)} =$$

$$= (\lambda_k \lambda_{k+1} \dots \lambda_N)^{1/(N-k+1)}, \quad (7)$$

причем равенство имеет место только тогда, когда все $\lambda_i c_i$ равны между собой.

С учетом второго условия (4) мы приходим к заключению, что

$$c_i = (\lambda_k \lambda_{k+1} \dots \lambda_N)^{1/(N-k+1)} / \lambda_i. \quad (8)$$

Поскольку из неравенства $\lambda_k \geq \lambda_{k+1} \geq \dots \geq \lambda_N > 0$ следует, что набор (8) принадлежит множеству R , то при указанных c_i достигается минимум в левой части (6).

Предположим временно, что ортонормированные векторы $\{x^i\}$ удовлетворяют условию

$$(x^k, Ax^k) \leq (x^{k+1}, Ax^{k+1}) \leq \dots \leq (x^N, Ax^N). \quad (9)$$

Тогда, рассуждая аналогично, мы придем к заключению, что правая часть (6) равна

$$((x^k, Ax^k)(x^{k+1}, Ax^{k+1}) \dots (x^N, Ax^N))^{1/(N-k+1)}$$

и, следовательно,

$$\lambda_k \lambda_{k+1} \dots \lambda_N \leq (x^k, Ax^k)(x^{k+1}, Ax^{k+1}) \dots (x^N, Ax^N). \quad (10)$$

Ясно, однако, что любой набор векторов $\{x^i\}$ можно упорядочить таким образом, чтобы выполнялись неравенства (9). Следовательно, (10) справедливо для всех ортонормированных $\{x^i\}$.

Упражнения к гл. 8

1. Пусть $n(A) = \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2}$. Показать, что

$$\left(\sum_{j=1}^N |\lambda_j|^{2k} \right)^{1/2} \leq n(A^k) \leq c_0 N^{m-1} \left(\sum_{j=1}^N |\lambda_j|^{2k} \right)^{1/2},$$

где m — максимальная кратность любого характеристического числа матрицы A (*). (Гочи.)

2. Если λ_i, μ_i, ν_i — занумерованные, как обычно, характеристические числа соответственно матриц A^*A, B^*B и $(A+B)^*(A+B)$, то

$$\sum_{i=1}^k \nu_i^{1/2} \leq \sum_{i=1}^k \lambda_i^{1/2} + \sum_{i=1}^k \mu_i^{1/2}, \quad k=1, 2, \dots, N^2. \quad (\text{Фань Цзы.})$$

При тех же предположениях имеет место неравенство

$$(\nu_{i+j-1})^{1/2} \leq \lambda_i^{1/2} + \mu_j^{1/2}. \quad (\text{Фань Цзы.})$$

*) Предполагается, что $|\lambda_j| > 0$ ($j=1, 2, \dots, N$). (Прим. ред.)

3. Если $H = A + iB$ — эрмитова матрица^{*)}, то для того, чтобы H была положительно определенной, необходимо и достаточно, чтобы характеристические числа матрицы $iA^{-1}B$ были действительны и не превосходили единицы по модулю. (Робертсон — О. Таусски.)

4. Если $H = A + iB$ — положительно определенная матрица, где матрицы A и B действительные, то $|A| \geq |H|$, причем равенство имеет место тогда и только тогда, когда $B = 0$. (О. Таусски.)

5. Если H_1 — положительно определенная эрмитова матрица, а H_2 — эрмитова матрица, то $H_1 + H_2$ положительно определенная тогда и только тогда, когда все характеристические числа матрицы $H_1^{-1}H_2$ больше -1 . (Фань Цзы — О. Таусски.)

6. Пусть K_1 — положительно определенная матрица и K_2 — такая матрица, что матрица K_1K_2 эрмитова. Тогда K_1K_2 положительно определена в том и только в том случае, когда все характеристические числа K_2 действительны и положительны. (Фань Цзы — О. Таусски.)

7. Если A и B — симметрические матрицы, то характеристические числа матрицы $AB - BA$ чисто мнимые. Показать далее, что $\text{Sp}((AB)^2) \leq \text{Sp}(A^2B^2)$.

8. Пусть A и B — две произвольные матрицы; показать, что квадрат абсолютной величины произвольного характеристического числа матрицы AB больше или равен произведению минимального характеристического числа матрицы AA' на минимальное характеристическое число матрицы BB' .

9. Пусть $H = A + iB$ — положительно определенная эрмитова матрица; тогда $|A| > |B|$. (Робертсон.)

10. Если $A = B + C$, где B — положительно определенная матрица, а C — кососимметрическая матрица, то $|A| \geq |B|$. (О. Таусски.)

11. Если A — симметрическая матрица, A и $I - A$ — неотрицательно определенные матрицы, а O — ортогональная матрица, то $|I - AO| \geq |I - A|$. (О. Таусски.)

12. Доказать неравенство Шура¹⁾:

$$\sum_{i=1}^N |\lambda_i|^2 < \sum_{i,j=1}^N |a_{ij}|^2.$$

13. Пусть A — симметрическая матрица, имеющая k нулевых характеристических чисел. Тогда

$$\sum_{i=1}^N \lambda_i = \sum_{i=1}^{N-k} \lambda_i, \quad \sum_{i=1}^N \lambda_i^2 = \sum_{i=1}^{N-k} \lambda_i^2.$$

где суммирование проводится по всем ненулевым характеристическим числам. Из неравенства Коши

$$\left(\sum_{i=1}^{N-k} \lambda_i \right)^2 < (N-k) \left(\sum_{i=1}^{N-k} \lambda_i^2 \right)$$

вывести неравенство

$$(\text{Sp } A)^2 < (N-k) \text{Sp } (A^2),$$

т. е.

$$k < \frac{N \text{Sp } (A^2) - (\text{Sp } A)^2}{\text{Sp } (A^2)}.$$

14. Воспользовавшись результатом предыдущего упражнения, показать, что ранг симметрической матрицы A больше или равен отношению $(\text{Sp } A)^2 / \text{Sp } (A^2)$.

^{*)} A и B — вещественные матрицы. (Прим. ред.)

¹⁾ I. Schur, Math. Ann. 66 (1909), 488—510.

15. Получить аналогичные соотношения в терминах $\text{Sp}(A^k)$ для $k=1, 2, 3, \dots$ (Использованный нами прием впервые применил Шнирельман в связи с проблемой представления каждого целого числа в виде суммы не более чем фиксированного числа простых чисел.)

16. Пусть λ_N — наименьшее, а λ_1 — наибольшее характеристическое число положительно определенной матрицы A . Тогда

$$(x, x)^2 \leq (Ax, x)(A^{-1}x, x) \leq \frac{(\lambda_1 + \lambda_N)^2}{4\lambda_1\lambda_N} (x, x)^2.$$

Это частный случай более общего результата, полученного В. Гройбом и В. Райнболтом¹⁾.

17. Если матрица X положительно определенная, то $X + X^{-1} \geq 2I$.

18. Если матрица X положительно определенная, то существует единственная матрица Y , положительно определенная и имеющая заранее заданные диагональные элементы, которая минимизирует величину $\text{Sp}(XY^{-1})$.

19. Эта матрица Y удовлетворяет уравнению вида $X = Y\Lambda Y$, где Λ — диагональная матрица с положительными элементами λ_i .

20. Минимальное значение величины $\text{Sp}(XY^{-1})$ равно $\sum_{i=1}^N \lambda_i / \lambda_i$ (П. Уитл.)

21. Справедливо ли неравенство $BC + CB \geq 2B^{1/2}CB^{1/2}$, если B и C — положительно определенные матрицы?

22. Справедливо ли неравенство $(A+B)^{1/2}C(A+B)^{1/2} \leq A^{1/2}CA^{1/2} + B^{1/2}CB^{1/2}$, если A, B и C — положительно определенные матрицы?

23. Пусть A и B — матрицы размера $N \times N$ с комплексными элементами такие, что матрицы $I - A^*A$ и $I - B^*B$ неотрицательно определенные; тогда $|I - A^*B| \geq |I - A^*A| |I - B^*B|$. (Н у а, Acta Math. Sinica 5 (1955), 463—470. См. Math. Rev. 17 (1956), 703.) Обобщения, опирающиеся на представление, полученное в § 10 гл. 6, можно найти в работе Беллман (R. Bellman), Representation Theorems and Inequalities for Hermitian Matrices, Duke Math. J., 1959, другие обобщения содержатся в работе Маркус (M. Marcus), On a Determinantal Inequality, Amer. Math. Monthly 65 (1958), 266—268.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 1. Классическая книга

Харди, Литлвуд, Поля (G. H. Hardy, J. E. Littlewood, G. Polya), Inequalities, Cambridge University Press, New York, 1934. [Русский перевод: Неравенства, ИЛ, 1948.]

послужила началом современной теории неравенств, а также явилась стимулом для непрекращающегося большого интереса к этой области. Некоторые более поздние результаты, а также их различные применения излагаются в книге

Беккенбах и Беллман (E. F. Beckenbach and R. Bellman), Inequalities, Ergeb. Math. 1960. [Русский перевод: Неравенства, «Мир», 1965.]

¹⁾ В. Гройб и В. Райнболт (W. Greub and W. Rheinboldt), On a Generalization of an Inequality of L. V. Kantorovich, Proc. Amer. Math. Soc., 1959.

В этой книге содержится гораздо более обширный материал о неравенствах, связанных с матрицами и характеристическими числами.

Основные результаты этой главы были получены в статьях Фань Цзы:

Фань Цзы (К у Ф а н), Some Inequalities Concerning Positive-definite Matrices, Proc. Cambridge Phil. Soc. 51 (1955), 414—421;

Фань Цзы (К у Ф а н), On a Theorem of Weyl Concerning the Eigenvalues of Linear Transformations, I, Proc. Nat. Acad. Sci. U. S. A. 35 (1949), 652—655;

Фань Цзы (К у Ф а н), On a Theorem of Weyl Concerning the Eigenvalues of Linear Transformations, II, Proc. Nat. Acad. Sci. U. S. A. 36 (1950), 31—35;

Фань Цзы (К у Ф а н), Maximum Properties and Inequalities for the Eigenvalues of Completely Continuous Operators, Proc. Nat. Acad. Sci. U. S. A. 37 (1951), 760—766;

Фань Цзы (К у Ф а н), Problems 4429 and 4430, Amer. Math. Monthly 58 (1951), 194 и 80 (1953), 48—50 (здесь можно найти теоремы 7 и 11).

Нами использовались в основном совершенно другие методы. Мы стремились показать, насколько полезными являются интегральные представления при выводе неравенств.

§§ 2—4. Глубокое рассмотрение этих классических неравенств содержится в книге Харди, Литлвуда и Поля, цитированной выше.

§ 5. Этот результат, насколько это известно автору, принадлежит Фань Цзы; см. третью и пятую из цитированных выше работ. См. также статьи

Оппенгейм (A. Oppenheim), Inequalities Connected with Definite Hermitian Forms, J. London Math. Soc. 5 (1930), 114—119;

Оппенгейм (A. Oppenheim), Inequalities Connected with Definite Hermitian Forms, II, Amer. Math. Monthly 61 (1954), 463—466,

где приводятся дальнейшие ссылки.

Подобные применения представления, выведенного в § 10 гл. 6, можно найти в работе

Беллман (R. Bellman), Hermitian Matrices and Representation Theorems, Duke Math. J., 1959,

где дано обобщение некоторых результатов, полученных Хуа; см.

Хуа (L. K. Hua), Inequalities Involving Determinants, Acta Math. Sinica 5 (1955), 463—470; Math. Rev. 17 (1956), 703.

Хуа использует другой тип интегрального представления, вытекающий из общей теории представления групп.

§ 6. См. работу

Беккенбах (E. F. Beckenbach), An Inequality for Definite Hermitian Determinants, Bull. Amer. Math. Soc., 1929, 325—329.

§ 7. Неравенство Адамара является одним из наиболее обоснованных результатов анализа: в литературе имеется свыше ста его доказательств. Приведенное доказательство следует работе

Беллман (R. Bellman) Notes on Matrix Theory, II, Amer. Math. Monthly 40 (1953), 174—175.

Относительно своего неравенства Адамар сделал очень интересное замечание в своей увлекательной книге

Адамар (J. Hadamard), The Psychology of Invention in the Mathematical Field, Princeton University Press, Princeton, N. J., 1949.

За оригинальным изложением мы отсылаем читателя к статье

Адамар (J. Hadamard), Resolution d'une question relative aux determinants, Bull. sci. math. 2 (1893), 240—248.

Обобщение неравенства Адамара можно найти в работе

Фишер (E. Fischer), Über den Hadamardschen Determinantsatz, Arch. Math. u. Phys. (3), 13 (1908), 32—40.

Другие далеко идущие обобщения даны Шуром:

Шур (I. Schur), Über endliche Gruppen und Hermitesche Formen, Math. Z. 1 (1918), 184—207.

См. также

Вильямсон (J. Williamson), Note on Hadamard's Determinant Theorem, Bull. Amer. Math. Soc. 53 (1947), 608—613.

§ 8. Этот результат и связанная с ним теорема 11 содержатся в пятой из указанных выше работ Фань Цзы.

Приведенное доказательство следует работе

Беллман, Гликсберг и Гросс (R. Bellman, I. Glicksberg and O. Gross), Notes on Matrix Theory, VI, Amer. Math. Monthly 62 (1955), 571—572.

Результат, сформулированный в упражнении 2, принадлежит Бергстрому, см. работу

Бергстром (H. Bergstrom), A Triangle Inequality for Matrices, Den 11te Skandinaviske Matematiker-kongress, Trondheim, 1949, Oslo 1952, 264—267,

а в доказательстве мы следуем работе

Беллман (R. Bellman), Notes on Matrix Theory, IV: An Inequality Due to Bergstrom, Amer. Math. Monthly 62 (1955), 172—173.

Результат упражнения 3 приводится в статье

Беллман (R. Bellman), Notes on Matrix Theory, IX, Amer. Math. Monthly 64 (1957), 189—191,

где даны два метода: один опирается на обобщение представления Ингама—Зигеля, на которое мы ссылались выше, другой — на неравенство Бергстрома.

§ 9. Этот параграф иллюстрирует важный прием. Иногда мультипликативные неравенства легче получить *ab initio*. Здесь мы следуем работе

Беллман (R. Bellman), Notes on Matrix Theory, VII, Amer. Math. Monthly 62 (1955), 647—648.

§ 12. В книге Харди, Литлвуда и Поля приводится история неравенства, связывающего среднее арифметическое и среднее геометрическое, а также большое число его доказательств. Приведенное доказательство импонирует тем, что оно опирается на обратную индукцию.

§ 13. Мы опустили доказательство неравенства (1), которое содержится во второй статье Фань Цзы.

Приведенный вывод неравенства (2) из неравенства (1) дан в работе

Беллман (R. Bellman), Notes on Matrix Theory, XV: Multiplicative Properties from Additive Properties, Amer. Math. Monthly.

Неравенства, связанные с различными функциями от характеристических чисел, широко исследовались различными авторами; см. следующие статьи, где можно найти обширную библиографию по этим вопросам:

Островский (A. Ostrowski), Sur quelques applications des fonctions convexes et concaves au sens de I. Schur, J. math. pures et appl. 31, no 9 (1952) 253—292;

Маркус и Лопес (M. D. Marcus and L. Lopes), Inequalities for Symmetric Functions and Hermitian Matrices, Canad. J. Math. 8 (1956). 524—531;

Маркус и Мак-Грегор (M. D. Marcus and J. L. McGregor), Extremal Properties of Hermitian Matrices, Canad. J. Math. 8 (1956);

Маркус и Мойлс (M. Marcus and B. N. Moyls), Extreme Value Properties of Hermitian Matrices, Department of Mathematics, University of British Columbia, Vancouver, Canada, 1956;

Маркус, Мойлс и Вествик (M. Marcus, B. N. Moyls and R. Westwick), Extremal Properties of Hermitian Matrices, Canad. J. Math.;

Маркус, Мойлс и Вествик (M. Marcus, B. N. Moyls and R. Westwick), Some Extreme Value Results for Indefinite Hermitian Matrices, II, III, III. J. Math. 2 (1958), 408—414.

Мы полностью опустили такой интересный и важный вопрос, как неравенства для характеристических чисел, выраженные через элементы матрицы A . Детальное рассмотрение этих вопросов можно найти в статьях

Брауэр (A. Brauer), Limits for the Characteristic Roots of a Matrix, IV, Duke Math. J. 19 (1952), 75—91;

Браун (E. T. Browne), Limits to the Characteristic Roots of Matrices, Amer. Math. Monthly 46 (1939), 252—265;

Паркер (W. V. Parker), The Characteristic Roots of Matrices, Duke Math. J. 12 (1945), 519—526.

Упомянем еще работу

Де Брёйни и Ван Данциг (N. G. De Bruijn and D. Van Dantzig), *Inequalities Concerning Minors and Eigenvalues*, *Nieuw. Arch. Wisk.* 4, no. 3 (1956), 18—35,

где систематически излагается большое число детерминантных неравенств.

Ряд интересных матричных неравенств приводится в работах Масани и Винера, Хелсона и Лауденслера, цитируемых в конце гл. 11, а также в статье

Фань Цзы и Гоффман (Ku Fan and A. J. Hoffman), *Some Metric Inequalities in the Space of Matrices*, *Proc. Amer. Math. Soc.* 6 (1955), 111—116.

Наконец, см.

Таусски (O. Taussky), *Bibliography on Bounds for Characteristic Roots of Finite Matrices*, National Bureau of Standards, Rept. 1162, September 1951.

Неравенства, совершенно отличные от приведенных нами, можно найти в статье

Мирский (L. Mirsky), *Inequalities for Normal and Hermitian Matrices*, *Duke Math. J.* 24 (1957), 591—599.

См. также

Хейнсворт (E. V. Haynsworth), *Note on Bounds for Certain Determinants*, *Duke Math. J.* 24 (1957), 313—320.

Если N достаточно велико, то на решение этой системы требуется затратить много усилий. Поэтому вместо того, чтобы применять к решению этой системы уравнений один из многочисленных существующих способов, мы последуем совершенно другим путем.

§ 3. **Функциональное уравнение.** Рассмотрим последовательность задач минимизации, минимизируя по всем x_k каждое из выражений:

$$Q_r(x) = \sum_{k=r}^N c_k (x_k - x_{k-1})^2 + \sum_{k=r}^N d_k (x_k - a_k)^2, \quad (1)$$

где $x_{r-1} = u$ предполагается заданным. Очевидно, что искомый минимум зависит от величины u и, конечно, от r . Введем последовательность функций $\{f_r(u)\}$ посредством равенства

$$f_r(u) = \min_x Q_r(x), \quad (2)$$

где $-\infty < u < \infty$, $r = 1, 2, \dots, N$. Заметим, что функция

$$f_N(u) = \min_{x_N} [c_N (x_N - u)^2 + d_N (x_N - a_N)^2] \quad (3)$$

может быть легко определена. Для нахождения рекуррентного соотношения, связывающего $f_r(u)$ с $f_{r+1}(u)$ при $r = 1, 2, \dots, N-1$, поступим следующим образом. Заметим, что

$$\begin{aligned} f_r(u) &= \min_{x_r} \min_{x_{r+1}} \dots \min_{x_N} \left[\sum_{k=r}^N c_k (x_k - x_{k-1})^2 + d_k (x_k - a_k)^2 \right] = \\ &= \min_{x_r} \left\{ c_r (x_r - u)^2 + d_r (x_r - a_r)^2 + \right. \\ &\quad \left. + \min_{x_{r+1}} \dots \min_{x_N} \left[\sum_{k=r+1}^N c_k (x_k - x_{k-1})^2 + d_k (x_k - a_k)^2 \right] \right\} = \\ &= \min_{x_r} [c_r (x_r - u)^2 + d_r (x_r - a_r)^2 + f_{r+1}(x_r)]. \quad (4) \end{aligned}$$

Поскольку $f_N(u)$ определена равенством (3), это рекуррентное соотношение в принципе определяет $f_{N-1}(u)$, $f_{N-2}(u)$, ..., $f_1(u)$, т. е. в конце концов функцию, которую мы хотели минимизировать с самого начала.

Используя цифровую вычислительную машину, можно вычислить в соответствии с (4) все члены последовательности $\{f_i(u)\}$.

Таким образом, вместо того чтобы решать частную задачу для фиксированного N , нашей целью было погрузить эту вариационную задачу в семейство задач подобного типа. Если

даже частная задача является сложной, она может быть достаточно просто связана с другими задачами семейства. Это свойство характерно для большого класса вариационных проблем, из которых задача, рассмотренная выше, представляет весьма частный пример.

§ 4. Рекуррентные соотношения. Принимая во внимание аналитическую структуру каждого члена последовательности $f_r(u)$, мы можем получить значительно большее. Начнем с доказательства по индукции того, что каждый член последовательности является квадратичной функцией u :

$$f_r(u) = u_r + v_r u + w_r u^2, \quad (1)$$

коэффициенты которой u_r , v_r и w_r зависят от r , но не от u . Легко видеть, что этот результат верен для $r = N$, поскольку минимум (1) достигается в точке

$$x_N = \frac{c_N u + d_N a_N}{c_N + d_N}. \quad (2)$$

Подставляя это значение x_N в правую часть (3), мы можем убедиться, что функция $f_N(u)$ действительно квадратична. Подобный же анализ для $f_{N-1}(u)$ и далее по индукции для $f_r(u)$ показывает, что все они имеют вид (1). Это означает, что для определения функций $f_r(u)$ достаточно найти последовательность коэффициентов $\{u_r, v_r, w_r\}$. Поскольку нам известны u_N, v_N и w_N , достаточно получить рекуррентные соотношения, связывающие u_r, v_r, w_r с u_{r+1}, v_{r+1} и w_{r+1} . Для их вывода обратимся к (3.4) и запишем

$$u_r + v_r u + w_r u^2 = \min_{x_r} [d_r (x_r - a_r)^2 + c_r (x_r - u)^2 + u_{r+1} + v_{r+1} x_r + w_{r+1} x_r^2]. \quad (3)$$

Величина x_r , дающая минимум (3), определяется формулой

$$(c_r + d_r + w_{r+1}) x_r = c_r u + d_r a_r - \frac{1}{2} v_{r+1}. \quad (4)$$

Подставляя это значение x_r в (3) и приравнявая коэффициенты при одинаковых степенях u , получаем желаемые рекуррентные соотношения, которые оказываются нелинейными. Дальнейшие детали мы оставляем читателю в качестве упражнения.

§ 5. Более сложный пример. В § 2 мы поставили задачу приближения последовательности величин $\{a_k\}$ последовательностью $\{x_k\}$, имеющей малую разность $x_{k+1} - x_k$. Сделаем дальнейший

шаг и рассмотрим проблему минимизации квадратичной функции

$$Q(x) = \sum_{k=1}^N c_k (x_k - x_{k-1})^2 + \sum_{k=1}^N d_k (x_k - a_k)^2 + \sum_{k=1}^N e_k (x_k - 2x_{k-1} + x_{k-2})^2, \quad (1)$$

где $x_0 = u$, $x_{-1} = v$.

Как и ранее, введем функцию двух переменных $f_r(u, v)$ посредством равенства

$$f_r(u, v) = \min_x \left[\sum_{k=r}^N c_k (x_k - x_{k-1})^2 + \sum_{k=r}^N d_k (x_k - a_k)^2 + \sum_{k=r}^N e_k (x_k - 2x_{k-1} + x_{k-2})^2 \right], \quad (2)$$

где $r = 1, 2, \dots, N$, $-\infty < u, v < \infty$, $x_{r-1} = u$, $x_{r-2} = v$. Нетрудно видеть, как и ранее, что для $r = 1, 2, \dots, N$

$$f_r(u, v) = \min_{x_r} [c_r (x_r - u)^2 + d_r (x_r - a_r)^2 + e_r (x_r - 2u + v)^2 + f_{r-1}(x_r, u)] \quad (3)$$

и что

$$f_r(u, v) = u_r + v_r u + w_r u^2 + v'_r v + w'_r v^2 + z_r uv.$$

Коэффициенты u_r, v_r, \dots, z_r в последнем равенстве зависят только от номера r . Комбинируя эти результаты, мы без труда получим рекуррентное соотношение, связывающее соседние члены последовательности $\{u_r, v_r, w_r, v'_r, w'_r, z_r\}$. Поскольку мы можем легко получить эти значения для $r = N$, мы имеем простой метод нахождения всей последовательности.

§ 6. Проблема Штурма — Лиувилля. Проблемой, имеющей большое значение в теоретической физике и прикладной математике, является задача определения величин λ , при которых однородное уравнение

$$\begin{aligned} u'' + \lambda \varphi(t) u &= 0, \\ u(0) &= u(1) = 0, \end{aligned} \quad (1)$$

имеет нетривиальное решение. Предполагается, что функция $\varphi(t)$ является вещественной, непрерывной и положительной на интервале $[0, 1]$.

В § 12 гл. 7 мы наметили один из путей получения приближенного решения этой задачи. Укажем сейчас другое решение, снова оставляя в стороне строгие обоснования. Отметим, однако, что вопрос о скорости сходимости при $N \rightarrow \infty$ имеет большое значение, поскольку от ответа на этот вопрос зависит эффективность метода. Не пытаясь разыскивать функцию $u(t)$ при $0 \leq t \leq 1$,

2. Применить это представление для получения другой аппроксимирующей матричной записи, а именно:

$$u_i = \lambda \sum_{j=1}^N k_{ij} \varphi_j u_j, \quad i=1, 2, \dots, N.$$

§ 7. Функциональные уравнения. Поскольку матрица A , определенная равенством (6.4), является вещественной и симметрической, мы знаем, что ее собственные значения могут быть определены при минимизации функции

$$\sum_{k=1}^N (u_k - u_{k-1})^2 \quad (1)$$

при ограничениях

$$u_0 = u_N = 0, \quad (2a)$$

$$\sum_{k=1}^{N-1} \varphi_k u_k^2 = 1. \quad (2b)$$

Будем определять вместо этого минимум функции

$$\sum_{k=r}^N (u_k - u_{k-1})^2 \quad (3)$$

при ограничениях

$$u_{r-1} = v, \quad (4a)$$

$$u_N = 0, \quad (4b)$$

$$\sum_{k=r}^{N-1} \varphi_k u_k^2 = 1, \quad (4c)$$

где $r=1, 2, \dots, N-1$, $-\infty < v < \infty$.

Минимальное значение (3) для всех r является функцией v , $f_r(v)$. Нетрудно видеть, что

$$f_{N-1} = (1/\sqrt{\varphi_{N-1}} - v)^2. \quad (5)$$

Аналогично предыдущему для $r=1, 2, \dots, N-1$ получим

$$f_r(v) = \min_{u_r} [(u_r - v)^2 + (1 - \varphi_r u_r)^2 f_{r+1}(v/\sqrt{1 - \varphi_r u_r^2})]. \quad (6)$$

Последовательность функций $\{f_r(v)\}$, очевидно, не обладает простой аналитической структурой. Однако рекуррентное соотношение (6) может быть использовано при числовых расчетах.

Упражнение

1. Рассмотреть подобным же образом задачу определения максимума и минимума функций:

$$(a) \quad (ax_1)^2 + (x_1 + ax_2)^2 + \dots + (x_1 + x_2 + \dots + x_{N-1} + ax_N)^2$$

при ограничении $x_1^2 + x_2^2 + \dots + x_N^2 = 1$;

$$(б) \quad x_1^2 + (x_1 + ax_2)^2 + \dots + (x_1 + ax_2 + a^2x_3 + \dots + a^{N-1}x_N)^2$$

при ограничении $x_1^2 + x_2^2 + \dots + x_N^2 = 1$;

$$(в) \quad x_1^2 + (x_1 + ax_2)^2 + [x_1 + ax_2 + (a+b)x_3]^2 + \dots \\ \dots + [x_1 + ax_2 + (a+b)x_3 + \dots + [a + (N-2)b]x_N]^2$$

при ограничении $x_1^2 + x_2^2 + \dots + x_N^2 = 1$.

Другие методы при рассмотрении этих же вопросов обсуждаются в статье Островский (А. М. Ostrovski), On the Bounds for a One-parameter Family of Matrices, J. für Math. 200 (1958), 190—200.

§ 8. Матрицы Якоби. Под матрицей Якоби будем понимать матрицу, обладающую тем свойством, что отличными от нуля являются только элементы, расположенные вдоль главной диагонали и двух прилежащих к ней диагоналей. Поэтому

$$a_{ij} = 0 \quad \text{при} \quad |i - j| \geq 2. \quad (1)$$

Рассмотрим теперь вопрос о решении системы уравнений

$$A\mathbf{x} = \mathbf{c}, \quad (2)$$

где A — положительно определенная матрица Якоби. Как нам известно, эта задача эквивалентна минимизации неоднородной квадратичной формы

$$Q(\mathbf{x}) = (\mathbf{x}, A\mathbf{x}) - 2(\mathbf{c}, \mathbf{x}). \quad (3)$$

Раскрывая это равенство, увидим, что следует минимизировать

$$Q(\mathbf{x}) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + \\ + 2a_{23}x_2x_3 + \dots + 2a_{N-1,N}x_{N-1}x_N + \\ + a_{NN}x_N^2 - 2c_1x_1 - 2c_2x_2 - \dots - 2c_Nx_N. \quad (4)$$

Рассмотрим задачу отыскания минимума функций

$$Q_k(\mathbf{x}, z) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + \\ + 2a_{23}x_2x_3 + \dots + 2a_{k-1,k}x_{k-1}x_k + a_{kk}x_k^2 - 2c_1x_1 - 2c_2x_2 - \dots - 2zx_k \quad (5)$$

для $k = 1, 2, \dots, N$. При этом

$$Q_1(\mathbf{x}, z) = a_{11}x_1^2 - 2zx_1. \quad (6)$$

Определим

$$f_k(z) = \min Q_k(\mathbf{x}, z). \quad (7)$$

Легко показать далее, что

$$f_k(z) = \min_{x_k} [a_{kk}x_k^2 - 2zx_k + f_{k-1}(c_{k-1} + a_{k-1, k}x_k)] \quad (8)$$

для $k = 2, 3, \dots$

Снова можно показать, что каждая функция $f_k(z)$ квадратична по z , так что

$$f_k(z) = u_k + v_k z + w_k z^2, \quad k = 2, 3, \dots, \quad (9)$$

причем

$$f_1 = \frac{z^2}{a_{11}}.$$

Подставляя в (8) последние выражения, получим рекуррентные соотношения, связывающие u_k, v_k, w_k с u_{k-1}, v_{k-1} и w_{k-1} . Оставляем их вывод читателю в качестве упражнения.

Упражнения

1. Обобщить рассмотренную процедуру на случай, когда матрица A обладает свойством $a_{ij} = 0$ при $|i - j| \geq 3$.

2. Получить соотношения, соответствующие приведенным в § 7, для наибольшего и наименьшего собственных значений симметричной матрицы A , у которой (1) $a_{ij} = 0$ при $|i - j| \geq 2$ и (2) $a_{ij} = 0$ при $|i - j| \geq 3$.

§ 9. Аналитическое продолжение. Мы получили возможность применить вариационные методы к решению уравнения $Ax = c$ ценой допущения, что матрица A является симметрической и положительно определенной. Возникает соблазн доказать, что точные решения, которые мы нашли, остаются в силе прежде всего для симметрических матриц, которые не являются знакоопределенными при условии, разумеется, что ни один из встречающихся знаменателей не обращается в нуль, а затем, при подходящих предположениях, — и для несимметрических матриц. Аргументацию можно провести следующим путем. Выражения для x_i являются линейными функциями c_i и рациональными функциями a_{ij} . Равенство вида

$$\sum_{j=1}^N a_{ij} x_j = c_i, \quad (1)$$

справедливое в той области пространства a_{ij} , где матрица A является знакоопределенной, несомненно должно выполняться для всех a_{ij} , которые удовлетворяют условию симметрии.

Очевидно, однако, что строгое обоснование доказательства подобного рода должно увести нас в трудную область функций

нескольких комплексных переменных. Поэтому мы откажемся от этого пути, утверждая лишь, что он существует, и предложим другой, который требует только знания элементарных фактов, касающихся аналитического продолжения функций одной комплексной переменной.

Рассмотрим вместо симметрической матрицы A матрицу $Iz + A$, где z —скаляр.

При достаточно больших положительных z эта матрица является положительно определенной и, кроме того, она является матрицей Якоби, если A —матрица Якоби.

Теперь мы хотим применить принцип аналитического продолжения. Величины x , найденные по формулам, выведенным ранее, являются рациональными функциями z , аналитическими при достаточно большой вещественной части z . Следовательно, тождества, справедливые в этой области, должны выполняться и при $z=0$, если только ни одна из встречающихся функций не имеет особенностей при $z=0$. Этими особенностями могут быть только нули знаменателей. Другими словами, формулы законны до тех пор, пока они не теряют смысла.

§ 10. Несимметрические матрицы. Чтобы применить технику вариационного исчисления к несимметрическим матрицам, мы должны использовать другой подход.

Рассмотрим выражение

$$f(x, y) = (x, Bx) + 2(x, Ay) + (y, By) - 2(a, x) - 2(b, y), \quad (1)$$

в котором по предположению B —симметрическая и положительно определенная матрица, а матрица A вещественна. Минимизируя (1) по x и y , мы получим уравнения

$$Bx + Ay = a, \quad A'x + By = b. \quad (2)$$

Чтобы гарантировать положительную определенность функции $f(x, y)$ как по x , так и по y , мы должны подчинить матрицу B некоторым дополнительным условиям. Поскольку мы хотим также использовать принцип аналитического продолжения, наиболее коротким путем к цели, пожалуй, будет выбор B в виде $B = Iz$, где z —достаточно большое положительное число.

Теперь, как и ранее, может быть использован аппарат функциональных уравнений. Мы оставляем читателю разработку деталей и проведение аналитического продолжения.

§ 11. Случай комплексной матрицы A . Посмотрим, что нужно делать в случае, когда матрица A является симметрической, но комплексной: $A = B + iC$, B и C вещественны. Уравнение $Ax = c$

принимает вид

$$(B + iC)(x + iy) = a + ib, \quad (1)$$

или после приравнивания вещественной и мнимой частей

$$\begin{aligned} Bx - Cy &= a, \\ Cx + By &= b. \end{aligned} \quad (2)$$

Заметим, что эти соотношения устанавливают естественное соответствие между комплексной матрицей $B + iC$ N -го порядка и матрицей

$$\begin{vmatrix} B & -C \\ C & B \end{vmatrix}, \quad (3)$$

имеющей порядок $2N$. Это соотношение встречалось раньше в примерах.

Для того чтобы связать эту матрицу и систему (2) с вариационной задачей, мы рассмотрим квадратичную форму

$$f(x, y) = (x, Cx) + 2(x, By) - (y, Cy) - 2(b, x) - 2(a, y). \quad (4)$$

Если принять условие, что C положительно определена, то окажется, что $f(x, y)$ является выпуклой функцией по x и вогнутой по y . Как следствие общих теорем, которые мы рассмотрим в гл. 16, имеем

$$\min_x \max_y f(x, y) = \max_y \min_x f(x, y). \quad (5)$$

Поскольку функция $f(x, y)$ является в данном случае квадратичной, нет необходимости привлекать эти общие результаты: обе части равенства (5) могут быть точно вычислены и после сравнения окажутся равными. Мы оставляем возможность провести это сравнение читателю.

Перед применением метода функциональных уравнений нам необходим дополнительный результат, заключающийся в том, что минимум по x_i и максимум по y_i могут определяться в любом порядке. В частности,

$$\min_{(x_1, x_2, \dots, x_N)} \max_{(y_1, y_2, \dots, y_N)} = \min_{x_1, y_1} \max_{(x_2, \dots, x_N)} \min_{(y_2, \dots, y_N)}. \quad (6)$$

Доказательство этого мы также оставляем читателю. Проблемы подобного рода являются фундаментальными в теории игр, которую мы обсудим коротко в гл. 16.

Комбинируя предыдущие результаты, мы можем рассмотреть случай комплексной матрицы A . В случае необходимости можно рассмотреть матрицу $A + izI$, где z — достаточно большая поло-

жительная величина, введенная, как и ранее, для целей аналитического продолжения.

§ 12. Слабо связанные системы. Рассмотрим теперь задачу решения системы линейных уравнений вида

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= c_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= c_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + b_1x_4 &= c_3, \\ b_1x_4 + a_{44}x_4 + a_{45}x_5 + a_{46}x_6 &= c_4, \\ a_{54}x_4 + a_{55}x_5 + a_{56}x_6 &= c_5, \\ a_{64}x_4 + a_{65}x_5 + a_{66}x_6 + b_2x_7 &= c_6, \\ &\dots \\ b_{N-1}x_{2N-3} + a_{2N-3, 2N-3}x_{2N-3} + a_{2N-3, 2N-1}x_{2N-1} + a_{2N-3, 2N}x_{2N} &= c_{2N-3}, \\ a_{2N-1, 2N-2}x_{2N-2} + a_{2N-1, 2N-1}x_{2N-1} + a_{2N-1, 2N}x_{2N} &= c_{2N-1}, \\ a_{2N, 2N-3}x_{2N-3} + a_{2N, 2N-1}x_{2N-1} + a_{2N, 2N}x_{2N} &= c_{2N}. \end{aligned} \quad (1)$$

Если все коэффициенты b_i равны нулю, то эта задача разделяется на N простых трехмерных задач. Разумно предположить, что имеется путь, позволяющий использовать почти диагональную структуру системы (1). Назовем матрицы типа матриц, сформированных из коэффициентов системы (1), *слабо связанными*. Они появляются во многих физических, инженерных и экономических задачах анализа систем с большим количеством переменных, в которых имеется «слабая связь» между отдельными частями системы.

Для упрощения обозначений введем матрицы

$$A_k = \|a_{i+2k-3, j+2k-3}\|, \quad i, j = 1, 2, 3, \quad (2)$$

и векторы

$$x^k = \begin{Bmatrix} x_{2k-1} \\ x_{2k} \\ x_{2k+1} \end{Bmatrix}, \quad c^k = \begin{Bmatrix} c_{2k-1} \\ c_{2k} \\ c_{2k+1} \end{Bmatrix}, \quad (3)$$

где $k = 1, 2, \dots$

Первоначально мы предположим, что матрицы коэффициентов (2) положительно определенные. Решение линейной системы эквивалентно задаче определения минимума неоднородной квадратичной формы

$$\begin{aligned} (x^1, A_1x^1) + (x^2, A_2x^2) + \dots + (x^N, A_Nx^N) - \\ - 2(c^1, x^1) - 2(c^2, x^2) - \dots - 2(c^N, x^N) + \\ + 2b_1x_6x_4 + 2b_2x_6x_7 + \dots + 2b_{N-1}x_{2N-3}x_{2N-2} \end{aligned} \quad (4)$$

Будем разыскивать этот минимум с помощью метода функционального уравнения. Введем последовательность функций $\{f_k(z)\}$,

определенную соотношением

$$f_k(z) = \min_{x^i} \left[\sum_{i=1}^k [(x^i, Ax^i) - 2(c^i, x^i)] + 2 \sum_{i=1}^{k-1} b_i x_{1+2i} x_{2i} + 2zx_{2k} \right] \quad (5)$$

для $-\infty < z < \infty$, $k = 1, 2, \dots$

Следуя обычным методам, получим рекуррентное соотношение

$$f_k(z) = \min_{R_k} [(x^k, A_k x^k) + 2zx_{2k} - 2(c^k, x^k) + f_{k-1}(b_{k-1}x_{2k-2})], \quad (6)$$

где через R_k обозначена трехмерная область

$$-\infty < x_{2k}, x_{2k-2}, x_{2k-4} < \infty.$$

§ 13. Упрощения — I. Мы можем записать (12.6) в форме

$$f_k(z) = \min_{x_{2k-2}} \left\{ \min_{x_{2k}, x_{2k-1}} [(x^k, A_k x^k) + 2zx_{2k} - 2(c^k, x^k)] + f_{k-1}(b_{k-1}x_{2k-2}) \right\}. \quad (1)$$

Введем последовательность функций

$$g_k(z, y) = \min_{x_{2k}, x_{2k-1}} [(x^k, A_k x^k) + 2zx_{2k} - 2(c^k, x^k)], \quad (2)$$

где $x_{2k-2} = y$. Тогда получаем простое одномерное рекуррентное соотношение

$$f_k(z) = \min_y [g_k(z, y) + f_{k-1}(b_{k-1}y)]. \quad (3)$$

§ 14. Упрощения — II. Как и в параграфе, посвященном матрицам Якоби, мы можем пойти дальше, если заметим, что $f_k(z)$ для каждого k квадратична по z :

$$f_k(z) = u_k + 2v_k z + w_k z^2, \quad (1)$$

где u_k, v_k, w_k не зависят от z . Подставляя это выражение в (13.3), мы без труда получим рекуррентные соотношения для последовательности $\{u_k, v_k, w_k\}$.

Упражнения

1. Получить рекуррентные соотношения, позволяющие находить наибольшее и наименьшее собственные значения слабо связанных матриц.

2. Распространить предыдущие результаты на случай, когда не все A_k имеют одинаковые размеры.

§ 15. Уравнение $Ax = y$. Теперь мы обсудим простую тему. Мы знаем, что если матрица A является положительно определенной, то решение уравнения

$$Ax = y, \quad (1)$$

которое дается формулой $x = A^{-1}y$, может быть также получено как решение задачи минимизации квадратичной формы

$$Q(x) = (x, Ax) - 2(x, y). \quad (2)$$

Минимальное значение (2) равно $-(y, A^{-1}y)$. Сравнивая два подхода к проблеме решения (1), мы можем получить некоторые интересные тождества. Введем функцию N переменных

$$f_N(y_1, y_2, \dots, y_N) = \min_{x_i} \left[\sum_{i,j=1}^N a_{ij} x_i x_j - 2 \sum_{i=1}^N x_i y_i \right]. \quad (3)$$

Запишем

$$\begin{aligned} \sum_{i,j=1}^N a_{ij} x_i x_j - 2 \sum_{i=1}^N x_i y_i &= \\ &= a_{NN} x_N^2 + \sum_{i,j=1}^{N-1} a_{ij} x_i x_j - 2 \sum_{i=1}^{N-1} x_i (y_i - a_{iN} x_N) - 2 x_N y_N. \end{aligned} \quad (4)$$

Из этого выражения получим следующее функциональное уравнение:

$$\begin{aligned} f_N(y_1, y_2, \dots, y_N) &= \min_{x_N} [a_{NN} x_N^2 - 2 x_N y_N + \\ &+ f_{N-1}(y_1 - a_{1N} x_N, y_2 - a_{2N} x_N, \dots, y_{N-1} - a_{N-1,N} x_N)]. \end{aligned} \quad (5)$$

Чтобы конструктивным образом использовать это соотношение, напомним, что $f_N(y_1, y_2, \dots, y_N) = -(y, A^{-1}y)$. Поэтому можно записать

$$f_N(y_1, y_2, \dots, y_N) = \sum_{i,j=1}^N c_{ij}(N) y_i y_j. \quad (6)$$

Возвращаясь к (5), получим

$$\begin{aligned} \sum_{i,j=1}^N c_{ij}(N) y_i y_j &= \min_{x_N} \left[a_{NN} x_N^2 - 2 x_N y_N + \sum_{i,j=1}^{N-1} c_{ij}(N-1) \times \right. \\ &\quad \left. \times (y_i - a_{iN} x_N)(y_j - a_{jN} x_N) \right], \end{aligned} \quad (7)$$

или, группируя соответствующие члены,

$$\begin{aligned} \sum_{i,j=1}^N c_{ij}(N) y_i y_j &= \min_{x_N} \left[x_N^2 \left\{ a_{NN} + \sum_{i,j=1}^{N-1} a_{iN} a_{jN} c_{ij}(N-1) \right\} + \right. \\ &\quad \left. + x_N \left\{ -2 y_N - 2 \sum_{i,j=1}^{N-1} (y_i a_{jN}) c_{ij}(N-1) \right\} + \sum_{i,j=1}^{N-1} c_{ij}(N-1) y_i y_j \right]. \end{aligned} \quad (8)$$

Теперь просто выполнить минимизацию

$$x_N = \frac{y_N + \sum_{l,j=1}^{N-1} a_{lN} y_l c_{lj}(N-1)}{a_{NN} + \sum_{l,j=1}^{N-1} a_{lN} a_{jN} c_{lj}(N-1)}, \quad (9)$$

вследствие чего сам минимум дается равенством

$$\frac{\left(\sum_{l,j=1}^{N-1} c_{lj}(N-1) y_l y_j \right) \left(a_{NN} + \sum_{l,j=1}^{N-1} a_{lN} a_{jN} c_{lj}(N-1) \right) - \left(y_N + \sum_{l,j=1}^{N-1} y_l a_{lN} c_{lj}(N-1) \right)^2}{a_{NN} + \sum_{l,j=1}^{N-1} a_{lN} a_{jN} c_{lj}(N-1)}, \quad (10)$$

Приравнявая коэффициенты при $y_l y_j$ в сумме $\sum_{l,j=1}^N c_{lj}(N) y_l y_j$ и в (10), получим рекуррентные соотношения, связывающие последовательности $\{c_{lj}(N)\}$ и $\{c_{lj}(N-1)\}$.

§ 16. Квадратичное уклонение. Задача минимизации по всем x_k квадратичной формы

$$Q_N(\mathbf{x}) = \int_0^T \left[f(t) - \sum_{k=1}^N x_k g_k(t) \right]^2 dt \quad (1)$$

в принципе решается весьма просто. Пусть $\{h_k(t)\}$ есть ортогональная последовательность функций, полученная по $g_k(t)$ с помощью процедуры ортогонализации Грама—Шмидта. Очевидно, что без ограничения общности функции $g_k(t)$ можно считать линейно независимыми. Тогда

$$\begin{aligned} \min_x Q_N(\mathbf{x}) &= \min_y \int_0^T \left[f(t) - \sum_{k=1}^N y_k h_k(t) \right]^2 dt = \\ &= \int_0^T f^2(t) dt - \sum_{k=1}^N \left(\int_0^T f(t) h_k(t) dt \right)^2. \end{aligned} \quad (2)$$

Этот результат может быть записан как

$$\int_0^T f^2(t) dt - \int_0^T \int_0^T f(s) f(t) k_N(s, t) ds dt, \quad (3)$$

где

$$k_N(s, t) = \sum_{k=1}^N h_k(s) h_k(t), \quad (4)$$

Получим теперь рекуррентное соотношение, связывающее члены последовательности $\{k_N(s, t)\}$. Введем квадратичный функционал—функцию от функции $f(t)$:

$$\varphi_N(f) = \min_x \int_0^T \left(f - \sum_{k=1}^N x_k g_k \right)^2 dt. \quad (5)$$

Используя технику функциональных уравнений, получим для $N = 2, 3, \dots$

$$\varphi_N(f) = \min_{x_N} \varphi_{N-1}(f - x_N g_N) \quad (6)$$

"

$$\varphi_1(f) = \min_{x_1} \int_0^T (f - x_1 g_1)^2 dt. \quad (7)$$

Чтобы раскрыть (6), мы воспользуемся приемом, неоднократно применявшимся ранее и связанным с квадратичным характером $\varphi_N(f)$.

Итак,

$$\begin{aligned} \varphi_N(f) = & \min_{x_N} \left[\int_0^T (f - x_N g_N)^2 dt - \right. \\ & \left. - \int_0^T \int_0^T (f(s) - x_N g_N(s))(f(t) - x_N g_N(t)) k_N(s, t) ds dt \right] = \\ & = \min_{x_N} \left\{ \int_0^T f^2 dt - \int_0^T \int_0^T k_{N-1}(s, t) f(s) f(t) ds dt - \right. \\ & \left. - 2x_N \left[\int_0^T f g_N dt - \int_0^T \int_0^T g_N(s) f(t) k_{N-1}(s, t) ds dt \right] + \right. \\ & \left. + x_N^2 \left[\int_0^T g_N^2 dt - \int_0^T \int_0^T k_{N-1}(s, t) g_N(s) g_N(t) ds dt \right] \right\}. \quad (8) \end{aligned}$$

Находя величину x_N , минимизирующую (8), и вычисляя минимум $\varphi_N(f)$, получим рекуррентное соотношение

$$\begin{aligned} k_N(s, t) = & k_{N-1}(s, t) + \\ & + \frac{1}{d_N} \left\{ g_N(s) g_N(t) - g_N(s) \int_0^T g_N(s_1) k_{N-1}(s_1, t) ds_1 - \right. \\ & \quad \left. - g_N(t) \int_0^T g_N(t_1) k_{N-1}(t, t_1) dt_1 + \right. \\ & \left. + \int_0^T \int_0^T g_N(s_1) g(t_1) k_{N-1}(s_1, s) k_{N-1}(t_1, t) dt_1 ds_1 \right\}, \quad (9) \end{aligned}$$

где

$$d_N = \int_0^T g_N^*(t) dt + \int_0^T \int_0^T k_{N-1}(s, t) g_N(s) g_N(t) ds dt. \quad (10)$$

§ 17. Результат Стильеса. В предыдущих разделах мы подчеркивали связь между решением уравнения $Ax = b$ и минимизацией квадратичной формы $(x, Ax) - 2(b, x)$ (где A — положительно определенная матрица). Используем ту же идею для того, чтобы установить интересный результат, принадлежащий Стильесу. Обобщение этого результата мы получим в гл. 16.

Теорема 1. Если A — положительно определенная матрица, обладающая тем свойством, что $a_{ij} < 0$ при $i \neq j$, то все элементы обратной матрицы A^{-1} положительны.

Доказательство. Рассмотрим задачу минимизации $Q(x) = (x, Ax) - 2(b, x)$ в предположении, что все компоненты вектора b положительны. Предположим, что в точке минимума $x_1, x_2, \dots, x_k < 0$, а $x_{k+1}, x_{k+2}, \dots, x_N \geq 0$. Переписывая $Q(x)$ в виде

$$\begin{aligned} a_{11}x_1^2 + a_{22}x_2^2 + \dots + a_{NN}x_N^2 + \sum_{i,j=1}^k a_{ij}x_ix_j + \sum_{i=1}^k \sum_{j=k+1}^N a_{ij}x_ix_j + \\ + \sum_{i=k+1}^N \sum_{j=1}^k a_{ij}x_ix_j + \sum_{i,j=k+1}^N a_{ij}x_ix_j - 2 \sum_{i=1}^k b_ix_i - 2 \sum_{i=k+1}^N b_ix_i, \end{aligned} \quad (1)$$

мы видим, что вследствие отрицательности a_{ij} при $i \neq j$ можно получить меньшее значение $Q(x)$, если хотя бы один из x_i ($i = k+1, \dots, N$) отличен от нуля. Это уменьшение $Q(x)$ достигается заменой x_i на $-x_i$ ($i = 1, 2, \dots, k$) при сохранении неизменными всех остальных переменных. Во всяком случае, мы видим, что все x_i ($i = 1, 2, \dots, N$) являются неотрицательными.

Чтобы показать, что в действительности все они положительны при положительных b_i , мы заметим, что по крайней мере один из x_i должен быть больше нуля. В случае если в точке минимума

$$x_1 = x_2 = \dots = x_{N-1} = 0,$$

то x_N , определяемый как величина, которая минимизирует $a_{NN}x_N^2 - 2b_Nx_N$, равен $\frac{b_N}{a_{NN}}$ и, следовательно, положителен. Поскольку в точке минимума

$$a_{ii}x_i = b_i - \sum_{j \neq i} a_{ij}x_j \quad (i = 1, 2, \dots, N-1), \quad (2)$$

то и все $x_i > 0$. Мы видим, что при любых положительных b_i все составляющие вектора $A^{-1}b$ положительны. Это доказывает неотрицательность элементов матрицы A^{-1} .

Чтобы показать, что на самом деле все элементы матрицы A^{-1} положительны, достаточно установить, что вектор $A^{-1}b$ имеет положительные составляющие всякий раз, когда все коэффициенты b_i неотрицательны, а по крайней мере один из них положителен.

Обращаясь к (2), мы видим, что неравенство $a_{ij} < 0$ при $i \neq j$ обеспечивает выполнение этого условия.

Упражнения к гл. 9

1. Даны две последовательности $\{a_k\}$ и $\{b_k\}$, $k=0, 1, 2, \dots, N$. Часто нужно определить конечную последовательность $\{x_k\}$, $k=0, 1, 2, \dots, M$, которая выражает b_k наиболее «точно» в форме

$$b_k = \sum_{i=0}^M a_{k-i} x_i, \quad a_s = 0 \quad \text{при } s < 0.$$

Чтобы оценить близость приближения, мы используем сумму

$$Q_{N, M}(x) = \sum_{k=0}^N \left(b_k - \sum_{i=0}^M x_i a_{k-i} \right)^2.$$

Рассмотрим квадратичную форму по переменным y_0, y_1, \dots, y_N , определенную как

$$f_{N, M}(y_0, y_1, \dots, y_N) = \min_x \{ (y_0 - a_0 x_0)^2 + (y_1 - x_0 a_1 - x_1 a_0)^2 + \dots \\ \dots + (y_N - x_0 a_N - x_1 a_{N-1} - \dots - x_M a_{N-M})^2 \}.$$

Показать, что

$$f_{N, 1}(y_0, y_1, \dots, y_N) = \left(\sum_{k=0}^N y_k^2 \right) \left(\sum_{k=0}^N a_k^2 \right) - \left(\sum_{k=0}^N a_k y_k \right)^2$$

и что в общем случае

$$f_{N, M}(y_0, y_1, \dots, y_N) = \\ = \min_{x_0} [(y_0 - x_0 a_0)^2 + f_{N-1, M-1}(y_1 - x_0 a_1, y_2 - x_0 a_2, \dots, y_N - x_0 a_N)].$$

2. Записать

$$f_{N, M}(y_0, y_1, \dots, y_N) = \sum_{i, j=1}^N c_{ij}(N, M) y_i y_j$$

и использовать предыдущее рекуррентное соотношение для получения связи $c_{ij}(N, M)$ с $c_{ij}(N-1, M-1)$.

3. Записать $f_{N, M}(y_0, y_1, \dots, y_N) = (y, A_{M, N} y)$ и получить рекуррентное соотношение в матричной форме.

Дополнительное обсуждение этой задачи, являющейся исходной точкой теории экстраполяции Колмогорова—Винера, см. в книге Винера (Wiener), *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*, John Wiley & Sons, 1942, и в особенности приложение к этой книге, написанное Н. Левинсоном.

4. Пусть A и B — две положительные определенные матрицы порядка N , причем $AB \neq BA$, а c — заданный N -мерный вектор. Рассмотрим вектор $x_N = Z_N Z_{N-1} \dots Z_1 Z_1 c$, где каждая из матриц Z_i есть A или B . Матрицы Z_i ($i = 1, 2, \dots, N$) должны быть выбраны из условия максимизации скалярного произведения (x_N, b) , где b — фиксированный N -мерный вектор.

Для $N = 1, 2, \dots$ и всех c определить функцию

$$f_N(c) = \max_{\{Z_i\}} (x_N, b).$$

Тогда

$$f_1(c) = \max \{(Ac, b), (Bc, b)\},$$

$$f_N(c) = \max \{f_{N-1}(Ac), f_{N-1}(Bc)\}, N = 2, 3, \dots$$

5. Существует ли скаляр λ такой, что $f_N(c) \sim \lambda^N g(c)$ при $N \rightarrow \infty$?

6. Каков ответ на этот вопрос, если $AB = BA$?

7. Пусть C — заданная положительно определенная матрица, а Z_i должны быть выбраны так, чтобы максимизировать максимальное собственное значение матрицы $Z_N Z_{N-1} \dots Z_1 Z_1 C$. Обозначим через $g_N(C)$ этот максимум максиморум. Тогда

$$g_1(C) = \max \{\varphi(AC), \varphi(BC)\},$$

$$g_N(C) = \max \{g_{N-1}(AC), g_{N-1}(BC)\}, N = 2, 3, \dots$$

где через $\varphi(X)$ обозначено максимальное собственное значение матрицы X .

8. Существует ли скаляр λ такой, что $g_N(C) \sim \lambda^N h(C)$ при $N \rightarrow \infty$? См. Бом (C. Bohm) Sulla minimizzazione di una funzione del prodotto di enti non commutati, *Lincei-Rend. Sci. fis. Mat. e Not.* 23 (1957), 386—388.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 1. Теория динамического программирования в действительности является аппаратом изучения многошаговых процессов принятия решений. Оказывается, что многие вариационные задачи могут быть сформулированы на языке динамического программирования. Эта интерпретация дает новый подход, являющийся полезным как для аналитического рассмотрения, так и для целей вычисления. В этой главе мы интересовались, следуя нашей общей программе, лишь аналитическими аспектами. Для детального ознакомления с общей теорией читатель отсылается к книге

Беллман (R. Bellman), *Dynamic Programming*, Princeton University Press, Princeton, N. J., 1957. [Русский перевод: Р. Беллман, *Динамическое программирование*, ИЛ, 1960.]

§ 2. Задача, рассмотренная в этом параграфе, является простым примером так называемой проблемы «сглаживания», ср.

Беллман (R. Bellman), *Dynamic Programming and its Application to Variational Problems in Mathematical Economics*, *Calculus of Variations*, Proc. of Symp. in Appl. Math., McGraw-Hill, New York, 1958, т. VIII, 115—139.

Обсуждение проблемы сглаживания другого типа см. в следующих работах:

Шёнберг (I. J. Schoeneberg), On Smoothing Operations and Their Generating Functions, *Bull. Amer. Math. Soc.* 59 (1953), 199—230;

Винер (N. Winer), *Cybernetics*, John Wiley & Sons, N. Y., 1951. [Русский перевод: Н. Винер, Кибнетика, «Советское радио», 1968.]

§ 3. Вывод этого рекуррентного соотношения является частным примером применения принципа оптимальности (см. книгу, указанную в ссылке к § 1). Результаты, обсуждавшиеся в §§ 3—5, впервые были опубликованы в статье

Беллман (R. Bellman), On a Class of Variation Problems, *Quart. Appl. Math.* 14 (1957), 353—359.

§ 7. Эти результаты опубликованы в работе

Беллман (R. Bellman), Eigenvalues and Functional Equations, *Proc. Amer. Math. Soc.* 8 (1957), 68—72.

§ 8. Это обсуждение следует статье

Беллман (R. Bellman), On a Class of Variational Problems, *Quart. Appl. Math.* 19 (1957), 353—359.

Значительное число результатов, связанных с тем же фактом, что матрицы Якоби появляются при рассмотрении процессов случайного блуждания, содержится в работах

Карлин и Мак-Грегор (S. Karlin, J. McGregor), Coincident Properties of Birth and Death Processes, *Techn. Rep. 9*, Stanford University, 1958;

Дин (P. Dean), The Special Distribution of a Jacobian Matrix, *Proc. Cambridge Phil. Soc.* 52 (1956), 752—755.

§ 9. Это рассуждение, проведенное для случая конечных матриц, может быть легко распространено на более общий класс операторов.

Применение этих методов описано в статье

Беллман и Леман (R. Bellman, S. Lehman), Functional Equations in the Theory of Dynamic Programming, X: Resolvents, Characteristic Functions and Values, *Proc. Nat. Acad. Sci. U.S.A.* 44 (1958), 905—907.

§§ 10—11. Содержание этих параграфов заимствовано из более общего построения в работе

Беллман и Леман (R. Bellman, S. Lehman), Functional Equations in the Theory of Dynamic Programming, IX: Variational Analysis, Analytic Continuation and Imbedding of Operator, *Duke Math. J.*, 1960.

Обширное обсуждение материала этого плана имеется в работах

Долф и Сэдл (C. L. Dolph, A. Saddle), Point Characterization of the Schwinger Stationary Points in Exterior Scattering Problems, *J. Soc. Ind. Appl. Math.* 5 (1957), 89—104;

Долф, Мак-Лафлин и Маркс (C. L. Dolph, J. E. McLaughlin, I. Marx), Symmetric Linear Transformations and Complex Quadratic Forms, Comm. Pure and Appl. Math. 7 (1954), 621—632.

§§ 12—14. Эти результаты заимствованы из статьи

Беллман (R. Bellman), On Some Applications of Dynamic Programming to Matrix Theory, III. J. Math. 1 (1957), 297—301.

§§ 15—16. Эти результаты следуют из работы

Беллман (R. Bellman), Dynamic Programming and Mean Square Deviation, The RAND Corporation, Paper P-1147, Sept. 13, 1957.

§ 17. Этот результат получен в статье

Стилтьес (T. J. Stieltjes), Sur la racines de $X_n=0$, Acta Math. 9 (1886—1887).

Его обобщение будет дано в гл. 16. Другие приложения предложенного способа см.

Беллман (R. Bellman), On the Non-negativity of Green's Functions, Bull. Unione Matematico 12 (1957), 411—413.

Дальнейшие приложения аппарата функциональных уравнений к квадратичным формам см.

Калмани Кёлке (R. E. Kalman, R. W. Koepke), Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indexes, Amer. Soc. Mech. Eng., Instrument and Regulators Conferences, Newark, April 2—4 1958;

Беллман и Ричардсон (R. Bellman, J. M. Richardson), On the Application of Dynamic Programming to a Class of Implicit Variational Problems, Quart. Appl. Math., 1959;

Фреймер (M. Freimer), Dynamic Programming and Adaptive Control Processes, Lincoln Laboratory Report, 1959;

Саймон (H. A. Simon), Dynamic Programming under Uncertainty with a Quadratic Criterion Functions, Econometrica 24 (1956), 74—81.

МАТРИЦЫ И ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

§ 1. Обоснование. В этой главе, открывающей вторую часть книги, мы обсудим применение теории матриц к решению линейных систем уравнений вида

$$\frac{dx_i}{dt} = \sum_{j=1}^N a_{ij}x_j, \quad x_i(0) = c_i, \quad i = 1, 2, \dots, N, \quad (1)$$

где a_{ij} —постоянные коэффициенты.

Приведем некоторые соображения, объясняющие, почему уравнения этого на первый взгляд частного вида играют центральную роль. Рассмотрим физическую систему S , предполагая, что ее положение в некоторый момент времени t полностью описывается N функциями $x_1(t), x_2(t), \dots, x_N(t)$. Сделаем дальнейшее допущение о том, что скорость изменения всех этих функций при любом t зависит только от величины x_i в этот же момент времени.

Это, несомненно, всегда лишь приближение к действительному положению дел, но приближение весьма полезное и удобное.

Аналитическим выражением этих допущений является система дифференциальных уравнений вида

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \dots, x_N), \quad i = 1, 2, \dots, N, \quad (2)$$

с соответствующими начальными условиями

$$x_i(0) = c_i, \quad i = 1, 2, \dots, N. \quad (3)$$

Вектор $\mathbf{c} = (c_1, c_2, \dots, c_N)$ представляет начальное состояние системы. Набор постоянных $\{c_i\}$, для которых

$$f_i(c_1, c_2, \dots, c_N) = 0, \quad i = 1, 2, \dots, N, \quad (4)$$

играет особенно важную роль. Они определяют, очевидно, положение равновесия, поскольку система S не может отклониться от этого положения при отсутствии внешних сил. Если такие положения равновесия существуют, то чрезвычайно интересно исследовать поведение системы в их окрестности. Другими сло-

вами, мы хотим изучить устойчивость системы при малых возмущениях. Если возмущенная система возвратится в конечном счете в положение равновесия, то мы говорим, что положение равновесия устойчиво. В противном случае положение равновесия будет неустойчивым*). Рассмотрение этих ситуаций имеет большое практическое значение.

Для того чтобы привести это исследование, положим

$$x_i = c_i + y_i, \quad (5)$$

где через y_i обозначены малые отклонения. Подставляя (5) в (2), мы получим уравнения

$$\begin{aligned} \frac{dy_i}{dt} = \frac{dx_i}{dt} &= f_i(c_1 + y_1, c_2 + y_2, \dots, c_N + y_N) = \\ &= f_i(c_1, c_2, \dots, c_N) + \sum_{j=1}^N a_{ij} y_j + \dots, \quad i = 1, 2, \dots, N, \end{aligned} \quad (6)$$

где

$$a_{ij} = \left. \frac{\partial f_i}{\partial x_j} \right|_{x_1=c_1, x_2=c_2, \dots, x_N=c_N}, \quad (7)$$

а многоточием обозначены члены высшего порядка малости. Поведение системы S в окрестности положения равновесия $\{c_i\}$ приближенно описывается линейной системой вида (1) с постоянными коэффициентами. Разумеется, законность аппроксимации должна быть тщательно исследована.

Мы имеем, таким образом, веское обоснование для изучения линейных систем вида (1). Нашей целью является получение аналитического представления решения, что позволит нам установить его асимптотическое поведение при $t \rightarrow \infty$.

К вопросам устойчивости мы снова вернемся в гл. 13.

§ 2. Векторно-матричные обозначения. Для исследования системы (1.1) мы введем векторы y и c с компонентами соответственно y_i и c_i и матрицу $A = \|a_{ij}\|$. Из определения разности двух векторов очевидно, что производная от вектора должна быть определена как

$$\frac{dy}{dt} = \begin{bmatrix} \frac{dy_1}{dt} \\ \frac{dy_2}{dt} \\ \vdots \\ \frac{dy_N}{dt} \end{bmatrix} \quad (1)$$

*) Речь идет об асимптотической устойчивости. (Прим. перев.)

Подобным образом интеграл от $y(t)$ определяется как

$$\int_0^t y(s) ds = \begin{pmatrix} \int_0^t y_1(s) ds \\ \int_0^t y_2(s) ds \\ \vdots \\ \int_0^t y_N(s) ds \end{pmatrix}. \quad (2)$$

Аналогично вводятся производные и интегралы от матриц. Отсюда следует, что уравнения (1.1) могут быть записаны как

$$\frac{dx}{dt} = Ax, \quad x(0) = c. \quad (3)$$

Матрица A в общем случае не симметрична. Поэтому можно ожидать, что результаты и методы первой части этой книги не будут играть главной роли при рассмотрении системы (1.1). Мы должны развить новые методы для рассмотрения квадратных матриц общего вида.

Вектор, составляющие которого являются функциями от t , будет называться векторной функцией, или, коротко, функцией от t . Векторная функция *непрерывна*, если все ее составляющие в рассматриваемом интервале являются непрерывными функциями аргумента t . Эту же терминологию мы будем использовать при описании матричных функций.

Упражнения

1. Показать, что

$$(a) \frac{d}{dt}(x, y) = \left(\frac{dx}{dt}, y\right) + \left(x, \frac{dy}{dt}\right),$$

$$(б) \frac{d}{dt}(Ax) = \left(\frac{dA}{dt}\right)x + A \frac{dx}{dt},$$

$$(в) \frac{d}{dt}(AB) = \left(\frac{dA}{dt}\right)B + A \frac{dB}{dt},$$

$$(г) \frac{d}{dt}(X^{-1}) = -X^{-1} \left(\frac{dX}{dt}\right) X^{-1},$$

$$(д) \frac{d}{dt}(X^n) = \left(\frac{dX}{dt}\right) X^{n-1} + X \left(\frac{dX}{dt}\right) X^{n-2} + \dots + X^{n-1} \left(\frac{dX}{dt}\right).$$

2. Получить уравнение для производной от $X^{1/2}$.

§ 3. Нормы векторов и матриц. Мы могли бы при желании использовать скалярную функцию (x, x) как меру вектора x .

Однако более удобно использовать не эту евклидову норму, а более простое выражение

$$\|\mathbf{x}\| = \sum_{i=1}^N |x_i| \quad (1)$$

для вектора и

$$\|A\| = \sum_{i,j=1}^N |a_{ij}| \quad (2)$$

для матрицы. Легко проверяется, что

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\| &\leq \|\mathbf{x}\| + \|\mathbf{y}\|, & \|A + B\| &\leq \|A\| + \|B\|, \\ \|A\mathbf{x}\| &\leq \|A\| \cdot \|\mathbf{x}\|, & \|AB\| &\leq \|A\| \cdot \|B\|, \\ \|c_1\mathbf{x}\| &= |c_1| \cdot \|\mathbf{x}\|, & \|c_1A\| &= |c_1| \cdot \|A\|. \end{aligned} \quad (3)$$

Причиной, по которой мы выбрали предыдущие выражения в качестве норм для вектора и матрицы, является то, что проверка всех результатов (3) чрезвычайно проста. Как мы увидим в упражнениях, приведенных в конце параграфа, имеется большой выбор в определении нормы. Все нормы в равной степени применимы при рассмотрении конечномерных векторов и матриц. Выбор нормы становится делом более сложным лишь в случае, когда мы обращаемся к бесконечномерным векторам и матрицам.

Упражнения

1. Показать, что в качестве норм, удовлетворяющих соотношениям (3), можно выбрать выражения

$$\begin{aligned} \|\mathbf{x}\| &= \left(\sum_{i=1}^N |x_i|^2 \right)^{1/2} = (\mathbf{x}, \mathbf{x})^{1/2}, \\ \|A\| &= \left(\sum_{i,j=1}^N |a_{ij}|^2 \right)^{1/2} = [\text{Sp}(AA^*)]^{1/2}. \end{aligned}$$

2. Если мы положим $\|\mathbf{x}\| = \max |x_i|$, какое определение необходимо принять для $\|A\|$, чтобы все неравенства (3) выполнялись?

3. Пусть $\|\mathbf{x}\|$ есть норма вектора \mathbf{x} , удовлетворяющая условиям $\|\mathbf{x}\| + \|\mathbf{y}\| \geq \|\mathbf{x} + \mathbf{y}\|$, $\|c_1\mathbf{x}\| = |c_1| \|\mathbf{x}\|$, $\|\mathbf{x}\| \geq 0$ и $\|\mathbf{x}\| = 0$ тогда и только тогда, когда $\mathbf{x} = 0$. Показать, что выражение $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$ определяет

норму матрицы, удовлетворяющую остальным неравенствам (3). (Это стандартный путь введения нормы операторов.)

4. Если для вектора \mathbf{x} мы используем норму, введенную в упражнении 2, то какую норму мы получим для A , используя прием упражнения 3?

5. Показать, что сходимость последовательности векторов $\{\mathbf{x}^n\}$ к вектору \mathbf{x} по норме влечет покоординатную сходимость, и обратно, сходимость по всем координатам влечет сходимость по норме.

6. Показать, что сходимость последовательности векторов $\{\mathbf{x}^n\}$ по одной из норм, удовлетворяющей условиям упражнения 3, влечет сходимость по

любой другой норме, удовлетворяющей этим же условиям.

$$7. \text{ Показать, что } \left\| \int_a^b \mathbf{x}(t) dt \right\| < \int_a^b \|\mathbf{x}(t)\| dt,$$

$$\left\| \int_a^b A(t) dt \right\| < \int_a^b \|A(t)\| dt.$$

8. Показать, что $\|A^n\| < \|A\|^n$ для любой нормы, удовлетворяющей соотношениям (3).

9. Существует ли норма, удовлетворяющая неравенствам (3), для которой $\|AB\| = \|A\| \|B\|$?

§ 4. Бесконечные ряды векторов и матриц. В ходе доказательства существования решений линейного векторного уравнения, введенного выше, нам понадобятся бесконечные ряды векторов и матриц. Под вектором $\sum_{n=0}^{\infty} \mathbf{x}^n$ мы будем понимать вектор, i -я составляющая которого есть сумма ряда $\sum_{n=0}^{\infty} x_i^n$. Поэтому сходимость векторного ряда эквивалентна одномерной сходимости N рядов $\sum_{n=0}^{\infty} x_i^n$. Отсюда следует, что достаточным условием сходимости векторного ряда $\sum_{n=0}^{\infty} \mathbf{x}^n$ является сходимость скалярного ряда $\sum_{n=0}^{\infty} \|\mathbf{x}^n\|$.

Подобно этому *матричный ряд* вида $\sum_{n=0}^{\infty} A_n$ представляет собой N^2 рядов, а для сходимости матричного ряда достаточно, чтобы сходилась ряд $\sum_{n=0}^{\infty} \|A_n\|$.

§ 5. Существование и единственность решений линейной системы уравнений. Благодаря этим предварительным замечаниям мы можем доказать следующий основной результат.

Теорема 1. Если матрица $A(t)$ непрерывна при $t \geq 0$, то решение векторного дифференциального уравнения

$$\frac{d\mathbf{x}}{dt} = A(t)\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{c}, \quad (1)$$

существует для всех $t \geq 0$, является единственным и может быть записано в виде

$$\mathbf{x} = X(t)\mathbf{c}, \quad (2)$$

где $X(t)$ — матрица, определенная единственным образом и удовлетворяющая матричному дифференциальному уравнению

$$\frac{dX}{dt} = A(t)X, \quad X(0) = I. \quad (3)$$

Доказательство. Для установления существования решения (3) мы используем метод последовательных приближений. Рассмотрим вместо (3) эквивалентное интегральное уравнение

$$X = I + \int_0^t A(s)X ds. \quad (4)$$

Определим последовательность матриц $\{X_n\}$ следующим образом:

$$X_0 = I, \\ X_{n+1} = I + \int_0^t A(s)X_n ds, \quad n = 0, 1, 2, \dots \quad (5)$$

Далее имеем

$$X_{n+1} - X_n = \int_0^t A(s)(X_n - X_{n-1}) ds, \quad n = 1, 2, \dots \quad (6)$$

Положим

$$m = \max_{0 \leq t \leq t_1} \|A(t)\|. \quad (7)$$

Здесь и далее мы используем определение нормы (3.1) и (3.2). Используя (6), получим при $0 \leq t \leq t_1$

$$\|X_{n+1} - X_n\| = \left\| \int_0^t A(s)(X_n - X_{n-1}) ds \right\| < \\ < \int_0^t \|A(s)\| \|X_n - X_{n-1}\| ds \leq m \int_0^t \|X_n - X_{n-1}\| ds. \quad (8)$$

Поскольку в этом же интервале

$$\|X_1 - X_0\| \leq \int_0^t \|A(s)\| ds \leq mt, \quad (9)$$

то, используя (8), по индукции получим

$$\|X_{n+1} - X_n\| \leq \frac{m^{n+1}t^{n+1}}{(n+1)!} \quad \text{при } 0 \leq t \leq t_1. \quad (10)$$

Следовательно, ряд $\sum_{n=0}^{\infty} (X_{n+1} - X_n)$ сходится равномерно в интервале $[0, t_1]$. Поэтому матрица X_n равномерно сходится к матрице

$X(t)$, которая удовлетворяет уравнению (4), а следовательно, и (3).

В предположении, что $A(t)$ непрерывна при $t \geq 0$, мы можем выбрать t_1 сколь угодно большим и, таким образом, получим решение, существующее при всех $t \geq 0$.

Легко проверяется, что $x = X(t)c$ есть решение уравнения (1), удовлетворяющее требуемому начальному условию. Установим далее единственность решения уравнения (3). Пусть имеется другое решение, которое мы обозначим через Y . Тогда Y удовлетворяет (4), и, следовательно, мы имеем соотношение

$$X - Y = \int_0^t A(s) [X(s) - Y(s)] ds. \quad (11)$$

Поэтому

$$\|X - Y\| \leq \int_0^t \|A(s)\| \|X(s) - Y(s)\| ds. \quad (12)$$

Так как матрицы Y и X дифференцируемы, а следовательно, непрерывны, то существует максимум

$$m_1 = \max_{0 < t < t_1} \|X - Y\|. \quad (13)$$

Из (12) получим

$$\|X - Y\| \leq m_1 \int_0^t \|A(s)\| ds, \quad 0 \leq t \leq t_1. \quad (14)$$

Используя неравенство (12), мы получим

$$\|X - Y\| \leq m_1 \int_0^t \|A(s)\| \left(\int_0^{s_1} \|A(s_1)\| ds_1 \right) ds \leq \frac{m_1 \left(\int_0^t \|A(s)\| ds \right)^2}{2}. \quad (15)$$

Продолжая итеративную процедуру, мы получим

$$\|X - Y\| \leq \frac{m_1 \left(\int_0^t \|A(s)\| ds \right)^{n+1}}{(n+1)!}. \quad (16)$$

Устремляя n к бесконечности, мы видим, что $\|X - Y\| \leq 0$. Поэтому $X = Y$.

Имея матрицу X , легко построить решение уравнения (1). Оно равно $X(t)c$. Единственность решения уравнения (1) легко установить, используя те же рассуждения, что и выше.

Упражнение

1. Доказать существование решения уравнения (1) при условии интегрируемости матрицы $A(t)$ в смысле интеграла Римана в любом конечном интервале. Должно ли дифференциальное уравнение в этом случае удовлетворяться всюду? Рассмотреть, в частности, случай $A(t) = A$ при $0 \leq t \leq t_0$ и $A(t) = B$ при $t > t_0$.

§ 6. Матричная экспонента. Рассмотрим теперь частный случай, когда $A(t)$ — постоянная матрица. В скалярном случае уравнение

$$\frac{du}{dt} = au, \quad u(0) = c, \quad (1)$$

имеет решение $u = e^{at}c$. Было бы очень удобно найти аналогичное решение для матричного уравнения

$$\frac{dX}{dt} = AX, \quad X(0) = C, \quad (2)$$

имеющее форму $X = e^{At}C$.

По аналогии со скалярным случаем и имея в виду метод последовательных приближений, использованный в § 5, мы попытаемся определить матричную экспоненциальную функцию посредством ряда

$$e^{At} = I + At + \dots + \frac{A^n t^n}{n!} \dots \quad (3)$$

Докажем следующий результат:

Теорема 2. Матричный ряд, определенный выше, существует для всех матриц A при любом фиксированном t , и для фиксированной матрицы A он равномерно сходится в любой конечной области комплексной плоскости t .

Доказательство. Мы имеем

$$\frac{\|A^n t^n\|}{n!} \leq \frac{\|A\|^n |t|^n}{n!}. \quad (4)$$

Учитывая, что $\|A\|^n |t|^n / n!$ является общим членом разложения в ряд экспоненты $e^{\|A\| |t|}$, мы видим, что ряд (3) мажорируется равномерно сходящимся рядом и, следовательно, сам равномерно сходится в любой конечной области плоскости t .

Упражнение

1. Используя представление e^{At} в виде ряда, показать, что $\frac{d}{dt}(e^{At}) = Ae^{At} = e^{At}A$.

§ 7. Функциональные уравнения — I. Скалярная экспоненциальная функция удовлетворяет основному функциональному тождеству:

$$e^{as+bt} = e^{as}e^{bt}. \quad (1)$$

До тех пор, пока аналогичное равенство не доказано для матричной экспоненты, мы не имеем права использовать обозначение (6.3). Покажем далее, что

$$e^{A(s+t)} = e^{As}e^{At}. \quad (2)$$

Используя разложение в ряд для трех экспоненциальных функций и тот факт, что члены абсолютно сходящегося ряда можно группировать произвольным образом, можно записать

$$e^{As}e^{At} = \sum_{k=0}^{\infty} \frac{A^k s^k}{k!} \left(\sum_{l=0}^{\infty} \frac{A^l t^l}{l!} \right) = \sum_{n=0}^{\infty} A^n \left(\sum_{k+l=n} \frac{s^k t^l}{k! l!} \right) = \sum_{n=0}^{\infty} A^n \frac{(s+t)^n}{n!} = e^{A(s+t)}. \quad (3)$$

Полагая в (2) $s = -t$, мы получим важный результат:

$$e^{A(-t+t)} = I = e^{-At}e^{At}. \quad (4)$$

Следовательно, матрица e^{At} всегда невырождена и ее обратная равна e^{-At} . Это матричный аналог того факта, что скалярная экспонента никогда не обращается в нуль.

§ 8. Функциональные уравнения — II. Доказательство функционального равенства в § 7 является скорее проверкой результата, чем выводом. Для понимания результата обратимся к дифференциальному уравнению

$$\frac{dX}{dt} = AX. \quad (1)$$

Заметим, что e^{At} есть решение этого уравнения при начальном условии $X(0) = I$, а $e^{A(s+t)}$ — решение при $X(0) = e^{As}$. Поэтому из теоремы единственности мы можем заключить, что $e^{A(s+t)} = e^{At}e^{As}$.

§ 9. Функциональные уравнения — III. После вывода функционального равенства, обсуждавшегося выше, возникает естественный вопрос о связи между $e^{(A+B)t}$ и $e^{At}e^{Bt}$. Поскольку

$$e^{(A+B)t} = I + (A+B)t + \frac{(A+B)^2}{2} t^2 + \dots, \quad (1)$$

а

$$e^{At}e^{Bt} = \left(I + At + \frac{A^2 t^2}{2} + \dots \right) \left(I + Bt + \frac{B^2 t^2}{2} + \dots \right),$$

то

$$e^{(A+B)t} - e^{At}e^{Bt} = (BA - AB) \frac{t^2}{2} + \dots \quad (2)$$

Следовательно, равенство $e^{(A+B)t} = e^{At}e^{Bt}$ справедливо для всех t только в случае, когда $AB = BA$, т. е. когда матрицы A и B перестановочны. Легко видеть, что это условие является и достаточным.

§ 10. **Невырожденность решения.** В § 7 мы установили, что матрица e^{At} всегда невырождена. Докажем теперь, что этот факт вытекает из общего результата, заключающегося в том, что решение уравнения

$$\frac{dX}{dt} = A(t)X, \quad X(0) = I, \quad (1)$$

является невырожденным в любом интервале $0 \leq t \leq t_1$, в котором существует интеграл $\int_0^{t_1} |A(t)| dt$.

Имеется несколько путей доказательства. Наиболее интересный из них изложен ниже, в то время как два других метода даны далее в качестве упражнений. Рассматриваемый метод базируется на *тождестве Якоби*:

$$|X(t)| = e^{\int_0^t \text{Sp } A(s) ds}. \quad (2)$$

Для вывода этого результата рассмотрим производную от скалярной функции $|X(t)|$. Для упрощения обозначений рассмотрим двумерный случай.

Мы имеем

$$|X(t)| = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}, \quad (3)$$

причем

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2; & \frac{dy_1}{dt} &= a_{11}y_1 + a_{12}y_2, \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2; & \frac{dy_2}{dt} &= a_{21}y_1 + a_{22}y_2. \end{aligned} \quad (4)$$

и

$$\begin{aligned} x_1(0) &= 1, & y_1(0) &= 0, \\ x_2(0) &= 0, & y_2(0) &= 1. \end{aligned} \quad (5)$$

Далее имеем

$$\begin{aligned} \frac{d}{dt} |X(t)| &= \begin{vmatrix} \frac{dx_1}{dt} & \frac{dy_1}{dt} \\ x_2 & y_2 \end{vmatrix} + \begin{vmatrix} x_1 & y_1 \\ \frac{dx_2}{dt} & \frac{dy_2}{dt} \end{vmatrix} = \\ &= \begin{vmatrix} a_{11}x_1 + a_{12}x_2 & a_{11}y_1 + a_{12}y_2 \\ x_2 & y_2 \end{vmatrix} + \begin{vmatrix} x_1 & y_1 \\ a_{21}x_1 + a_{22}x_2 & a_{21}y_1 + a_{22}y_2 \end{vmatrix} = \\ &= a_{11} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} + a_{22} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = \text{Sp } (A(t)) |X(t)|. \end{aligned} \quad (6)$$

Поскольку, кроме того, $|X(0)| = I$, то

$$|X(t)| = e^{\int_0^t \text{Sp } A(s) ds} \quad (7)$$

Упражнения

1. Рассмотреть уравнения $\frac{dX}{dt} = A(t)X$, $X(0) = I$, и $\frac{dY}{dt} = -YA(t)$, $Y(0) = I$. Показать, что $Y = X^{-1}$ и что матрица $X(t)$ является невырожденной в любом интервале, где интегрируема $\|A(t)\|$.

2. Рассмотреть дифференциальное уравнение второго порядка $u'' + p(t)u' + q(t)u = 0$. Обозначив $u' = v$, показать, что данное уравнение эквивалентно следующей системе уравнений первого порядка:

$$\frac{du}{dt} = v,$$

$$\frac{dv}{dt} = -p(t)v - q(t)u.$$

3. Рассмотрим интеграл $J = \int_0^{t_1} \omega(u'' + p(t)u' + q(t)u) dt$. Интегрируя по частям, получим

$$J = [\dots] + \int_0^{t_1} u(\omega'' + p_1(t)\omega' + q_1(t)\omega) dt.$$

Какова связь между векторно-матричной системой для уравнения $\omega'' + p_1(t)\omega' + q_1(t)\omega = 0$ и векторно-матричной системой для уравнения $u'' + p(t)u' + q(t)u = 0$?

4. Рассмотреть следующее доказательство невырожденности матрицы $X(t)$. Если $|X(t)| = 0$ в некоторой точке t_2 , $0 < t_2 < t_1$, то существует линейная зависимость между столбцами матрицы $X(t_2)$, т. е. $c_1x^1 + c_2x^2 + \dots + c_Nx^N = 0$.

Поскольку $c_1x^1 + c_2x^2 + \dots + c_Nx^N$ есть решение уравнения $\frac{dx}{dt} = A(t)x$, то из равенства его нулю в одной точке следует тождественное равенство нулю в интервале $[0, t_1]$. Однако это противоречит условию $|X(t)| = 1$ при $t = 0$.

§ 11. Решение неоднородного уравнения. Постоянные коэффициенты. Рассмотрим теперь задачу решения неоднородной системы

$$\frac{dx}{dt} = Ax + f(t), \quad x(0) = c. \quad (1)$$

Здесь выявляется преимущество матричного экспоненциального обозначения. Мы имеем

$$e^{-At} \left(\frac{dx}{dt} - Ax \right) = \frac{d}{dt} (e^{-At}x) = e^{-At}f, \quad (2)$$

следовательно,

$$e^{-At}x = c + \int_0^t e^{-As}f(s) ds \quad (3)$$

или

$$\mathbf{x} = e^{At} \mathbf{c} + \int_0^t e^{A(t-s)} \mathbf{f}(s) ds. \quad (4)$$

Заметим, что использование матричной экспоненциальной функции позволяет нам записать решение системы (1) точно в такой же форме, что и в случае скалярного уравнения.

§ 12. Неоднородное уравнение. Переменные коэффициенты. Рассмотрим случай, когда матрица A зависит от времени. Наша цель состоит в решении системы

$$\frac{d\mathbf{x}}{dt} = A(t) \mathbf{x} + \mathbf{f}(t), \quad \mathbf{x}(0) = \mathbf{c}. \quad (1)$$

Воспользуемся методом вариации постоянных интегрирования, принадлежащим Лагранжу, и попытаемся найти решение в форме $\mathbf{x} = X\mathbf{y}$, где $X = X(t)$ — решение уравнения $dX/dt = A(t)X$, $X(0) = I$. Подстановка в (1) даст уравнение

$$X' \mathbf{y} + X \mathbf{y}' = A(t) X \mathbf{y} + \mathbf{f}(t), \quad X' \mathbf{y} = A(t) X \mathbf{y}. \quad (2)$$

Поэтому

$$X \mathbf{y}' = \mathbf{f}(t) \quad (3)$$

и, следовательно,

$$\mathbf{y}' = X^{-1}(t) \mathbf{f}(t). \quad (4)$$

Интегрирование (4) даст

$$\mathbf{y} = \mathbf{c} + \int_0^t X^{-1}(s) \mathbf{f}(s) ds. \quad (5)$$

Мы приходим к соотношению

$$\mathbf{x} = X(t) \mathbf{c} + \int_0^t X(t) X^{-1}(s) \mathbf{f}(s) ds, \quad (6)$$

которое является обобщением результата (11.4).

§ 13. Неоднородное уравнение. Сопряженная система. Продемонстрируем далее другой подход, являющийся одним из наиболее важных в общей теории линейных функциональных уравнений. Пусть $Y(t)$ — переменная неопределенная пока матрица. Проинтегрируем в пределах от 0 до t уравнение

$$Y(t) \frac{d\mathbf{x}}{dt} = Y(t) A(t) \mathbf{x} + Y(t) \mathbf{f}(t). \quad (1)$$

Интегрируя по частям, получаем следующий результат:

$$\begin{aligned}
 Y(t) \mathbf{x}(t) - Y(0) \mathbf{c} - \int_0^t \frac{dY}{ds} \mathbf{x}(s) ds = \\
 = \int_0^t Y(s) A(s) \mathbf{x}(s) ds + \int_0^t Y(s) \mathbf{f}(s) ds. \quad (2)
 \end{aligned}$$

Без потери общности мы можем принять $\mathbf{c} = \mathbf{0}$, поскольку мы всегда можем получить общее решение уравнения (12.1), прибавив к частному решению слагаемое $X(t)\mathbf{c}$. Поскольку нашей целью является получение решения $\mathbf{x}(t)$, мы примем наиболее удобное из возможных допущений относительно матрицы $Y(t)$, а именно, что

$$\frac{dY}{ds} = -Y(s) A(s), \quad 0 \leq s \leq t, \quad (3a)$$

$$Y(t) = I. \quad (3b)$$

Если мы сможем удовлетворить обоим этим уравнениям, то решение исходного уравнения (12.1) запишется в простой форме:

$$\mathbf{x} = \int_0^t Y(s) \mathbf{f}(s) ds. \quad (4)$$

Система (3) называется *сопряженной* по отношению к системе (12.1). Из общей теоремы существования и единственности, доказанной выше, вытекает, что решение системы (3) существует и является единственным. Матрица Y является при этом функцией от s и t .

Упражнение

1. Показать, что $Y(s, t) = X(t) X^{-1}(s)$.

§ 14. Теория возмущений. Интересные приложения получает формула для решения неоднородного уравнения в теории возмущений. Рассмотрим матричную экспоненциальную функцию $e^{A+\epsilon B}$. Мы хотим представить ее в виде ряда по степеням ϵ :

$$e^{A+\epsilon B} = e^A + \sum_{n=1}^{\infty} \epsilon^n Q_n(A, B). \quad (1)$$

Задача легко решается, если матрицы A и B перестановочны, поскольку тогда

$$e^{A+\epsilon B} = e^A e^{\epsilon B}.$$

Рассмотрим поэтому более интересный случай, когда $AB \neq BA$. Если мы запишем

$$e^{A+\varepsilon B} = I + \sum_{n=1}^{\infty} \frac{(A + \varepsilon B)^n}{n!} \quad (2)$$

и попытаемся сгруппировать члены с одинаковой степенью ε , то мы вскоре увидим, что сделать это достаточно трудно. Вместо этой прямой процедуры мы последуем другим путем. Матрица $e^{A+\varepsilon B}$ есть решение дифференциального уравнения

$$\frac{dX}{dt} = (A + \varepsilon B) X, \quad X(0) = I, \quad (3)$$

вычисленного в точке $t = 1$.

Запишем это уравнение в форме

$$\frac{dX}{dt} = AX + \varepsilon BX, \quad X(0) = I. \quad (4)$$

Из (11.4) следует, что X удовлетворяет линейному интегральному уравнению

$$X = e^{A^t} + \varepsilon \int_0^t e^{A(t-s)} BX(s) ds. \quad (5)$$

Решая это интегральное уравнение Вольтерра методом последовательных приближений, мы получаем бесконечный ряд вида

$$X = e^{A^t} + \varepsilon \int_0^t e^{A(t-s)} B e^{A^s} ds + \dots \quad (6)$$

Поэтому разложение в ряд матрицы $e^{A+\varepsilon B}$ содержит следующие два первых члена:

$$e^{A+\varepsilon B} = e^{A^1} + \varepsilon \int_0^1 e^{A(1-s)} B e^{A^s} ds + \dots \quad (7)$$

Упражнения

1. Положить $X(t) = e^{At} + \sum_{n=1}^{\infty} \varepsilon^n P_n(t)$ и использовать (5) для вывода рекуррентного соотношения, связывающего $P_n(t)$ и $P_{n-1}(t)$.

2. Приняв допущение, что $e^{At}e^{Bt}$ может быть записано в форме e^C (этот результат будет установлен ниже), где $C = C_1 t + C_2 t^2 + C_3 t^3 + \dots$, определить матрицы C_1, C_2, C_3 .

3. Приняв, что $e^{A+\varepsilon B}$ может быть записана в форме

$$e^{A+\varepsilon B} = e^{A^1} e^{C_1 \varepsilon} e^{C_2 \varepsilon^2} e^{C_3 \varepsilon^3} \dots,$$

определить матрицы C_1, C_2, C_3 ¹⁾.

¹⁾ См. также Фер (F. Fer), Acad. Roy. Belg. Cl. Sci. 44, № 5 (1958), 818—829.

(Разложение в ряд (7) обладает тем серьезным недостатком, что для матрицы $e^{A+\varepsilon B}$, являющейся унитарной в случае, если A и B косозермитовы, оно дает приближение, не являющееся унитарным. Разложение, приведенное выше, лишено этого недостатка.)

§ 15. Неотрицательность решения. Следующий вопрос возникает в математической экономике. Рассмотрим уравнение

$$\frac{dx}{dt} = Ax + f(t), \quad x(0) = c, \quad (1)$$

где A — матрица коэффициентов. Каковы необходимые и достаточные условия, налагаемые на матрицу A , чтобы все компоненты вектора x были неотрицательны при $t \geq 0$ при условии, что компоненты c неотрицательны и компоненты $f(t)$ неотрицательны при $t \geq 0$.

Обращаясь к (11.4), мы видим, что достаточное условие сводится к неотрицательности элементов e^{At} при $t \geq 0$. Легко видеть, что это условие является также необходимым.

Неожиданным является то, что имеется очень простой критерий выполнения этого условия.

Теорема 3. Для неотрицательности всех элементов матрицы e^{At} при $t \geq 0$ необходимо и достаточно, чтобы

$$a_{ij} \geq 0, \quad i \neq j, \quad i, j = 1, 2, \dots, N. \quad (2)$$

Доказательство. Поскольку

$$e^{At} = I + At + \dots, \quad (3)$$

то ясно, что условие (2) необходимо для того, чтобы e^{At} было неотрицательным при малых t . Чтобы установить достаточность, покажем, что условие $a_{ij} > 0$ при $i \neq j$ приводит к положительности элементов e^{At} при всех t . Ясно, что это условие гарантирует положительность e^{At} при малых t . Для любого целого n имеем тождество

$$e^{At} = (e^{At/n})^n. \quad (4)$$

Тот факт, что произведение положительных матриц само является положительным, доказывает требуемый результат.

Так как элементы e^{At} являются непрерывными функциями a_{ij} , то мы видим, что положительность элементов e^{At} при положительных a_{ij} ($i \neq j$) означает неотрицательность элементов e^{At} при $a_{ij} \geq 0$ ($i \neq j$).

Другой, более прямой путь доказательства достаточности условия (2) сводится к следующему. Пусть c_1 — скаляр такой, что все элементы матрицы $A + c_1 I$ неотрицательны. Тогда ясно, что все элементы матрицы $e^{(A+c_1 I)t}$ неотрицательны. Элементы диагональной матрицы $e^{-c_1 t}$ также неотрицательны в силу неот-

рицательности экспоненциальных функций. Заметив, что матрицы $A + c_1 I$ и $-c_1 I$ перестановочны и что поэтому

$$e^{At} = e^{(A+c_1 I)t - c_1 t I} = e^{(A+c_1 I)t} e^{-c_1 t I}, \quad (5)$$

получим желаемый результат.

Упражнения

1. Доказать предыдущий результат, используя систему дифференциальных уравнений

$$\frac{dx_i}{dt} = \sum_{j=1}^N a_{ij} x_j, \quad x_i(0) = c_i, \quad i = 1, 2, \dots, N.$$

2. Показать, что условие $a_{ij}(t) \geq 0$ ($i \neq j$) достаточно для того, чтобы обеспечить неотрицательность решения соответствующей системы дифференциальных уравнений с переменными коэффициентами, если начальные условия неотрицательны.

§ 16. Функциональное уравнение Пойа. Мы видели, что матрица $Y(t) = e^{At}$ удовлетворяет функциональному уравнению

$$Y(s+t) = Y(s)Y(t), \quad -\infty < s, t < \infty. \quad (1)$$

Возникает интересный и важный вопрос о том, существуют или нет другие типы решений этого фундаментального матричного уравнения. Ответ является простым, если считать, что $Y(t)$ имеет первую производную при всех конечных t . Дифференцируя (1) один раз по t , а другой раз по s , мы получаем два уравнения:

$$\begin{aligned} Y'(s+t) &= Y(s)Y'(t), \\ Y'(s+t) &= Y'(s)Y(t). \end{aligned} \quad (2)$$

Поэтому

$$Y(s)Y'(t) = Y'(s)Y(t). \quad (3)$$

Из (1) мы видим, что $Y(0) = Y(-t)Y(t)$, а отсюда видно, что матрица $Y(t)$ не может быть вырожденной ни при одном t^*). Поэтому согласно (3)

$$Y^{-1}(s)Y'(s) = Y'(t)Y^{-1}(t) \quad (4)$$

для всех s и t . Следовательно, $Y'(t)Y^{-1}(t) = A$, где A — постоянная матрица. Уравнение

$$Y'(t) = AY(t) \quad (5)$$

приводит нас к формуле $Y(t) = e^{At}Y(0)$.

Получим теперь более общий результат.

* Предполагается, что $Y(0)$ — невырожденная матрица. Полагая в (1) $s = t = 0$, получаем $Y(0) = Y^2(0)$; отсюда $Y(0) = I$. (Прим. ред.)

Теорема 4. Пусть $Y(t)$ — непрерывная невырожденная матричная функция t , удовлетворяющая функциональному уравнению (1) при $0 \leq s, t \leq s+t \leq t_0$. Тогда в интервале $[0, t_0]$ матрица $Y(t)$ записывается в виде $Y(t) = e^{At}$, где A — некоторая постоянная матрица.

Доказательство. Рассмотрим риманов интеграл

$$\int_0^t Y(s) ds = \lim_{\delta \rightarrow 0} \sum_{k=0}^N Y(k\delta) \delta, \quad (6)$$

где N определяется из условия $N\delta = t$.

Поскольку из (1) при $\delta > 0$ следует, что

$$Y(k\delta) = Y((k-1)\delta)Y(\delta) = Y^k(\delta), \quad (7)$$

мы получаем

$$\int_0^t Y(s) ds = \lim_{\delta \rightarrow 0} \sum_{k=0}^N Y^k(\delta). \quad (8)$$

Мы также имеем

$$\frac{[Y(\delta) - I]}{\delta} \sum_{k=0}^N Y(k\delta) \delta = Y((N+1)\delta) - I. \quad (9)$$

Это приводит к соотношению

$$\lim_{\delta \rightarrow 0} \left\{ \left(\frac{Y(\delta) - I}{\delta} \right) \left(\sum_{k=0}^N Y(k\delta) \delta \right) \right\} = Y(t) - I. \quad (10)$$

Поскольку $Y(0) = I$, то в силу непрерывности $Y(t)$

$$\int_0^t Y(s) ds = tI + o(t) \quad (11)$$

и, следовательно, матрица $\int_0^t Y(s) ds$ является невырожденной при малых t . Отсюда следует, что для фиксированного малого t и достаточно малого δ матрица $\sum_{k=0}^N Y(k\delta) \delta$ является невырожденной и

$$\lim_{\delta \rightarrow 0} \left(\sum_{k=0}^N Y(k\delta) \delta \right)^{-1} = \left(\int_0^t Y(s) ds \right)^{-1}. \quad (12)$$

Поэтому из (10) следует, что $[Y(\delta) - I]/\delta$ имеет предел при $\delta \rightarrow 0$. Обозначим этот предел через A . Тогда для малых t

$$A = \lim_{\delta \rightarrow 0} \left(\frac{Y(\delta) - I}{\delta} \right) = [Y(t) - I] \left[\int_0^t Y(s) ds \right]^{-1}. \quad (13)$$

Отсюда получаем

$$A \int_0^t Y(s) ds = Y(t) - I \quad (14)$$

и дифференцированием найдем

$$Y'(t) = AY(t), \quad Y(0) = I, \quad (15)$$

во всяком случае для малых t . Из функционального уравнения уже нетрудно заключить, что $Y(t) = e^{At}$ для всех t в интервале $[0, t_0]$.

Упражнения

1. Найти ошибку в следующем выводе, основанном на соотношении (10): Согласно (10) имеем

$$\lim_{\delta \rightarrow 0} \left[\frac{Y(\delta) - I}{\delta} \right] \lim_{\delta \rightarrow 0} \left[\sum_{k=0}^N Y(k\delta) \delta \right] = Y(t) - I,$$

$$\lim_{\delta \rightarrow 0} \left[\frac{Y(\delta) - I}{\delta} \right] \int_0^t Y(s) ds = Y(t) - I$$

и т. д.

2. Пусть Y и Z являются решениями уравнений

$$\frac{dY}{dt} = A(t)Y, \quad Y(0) = I, \quad \frac{dZ}{dt} = ZB(t), \quad Z(0) = I.$$

Тогда решение уравнения

$$\frac{dX}{dt} = A(t)X + XB(t), \quad X(0) = C,$$

дается равенством $X = YCZ$.

3. Запишем линейное уравнение $u'' + p(t)u' + q(t)u = 0$ в форме

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -q(t) & -p(t) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.$$

Пусть u_1 и u_2 — линейно независимые решения уравнения второго порядка. Используя предыдущий результат, показать, что функция $u = a_1 u_1^2 + a_2 u_1 u_2 + a_3 u_2^2$ является решением уравнения

$$u''' + 3p(t)u'' + [2p^2(t) + p'(t) + 4q(t)]u' + [4p(t)q(t) + 2q'(t)]u = 0.$$

§ 17. Уравнение $\frac{dX}{dt} = AX + XB$. Докажем теперь следующую теорему.

Теорема 5. *Решение уравнения*

$$\frac{dX}{dt} = AX + XB, \quad X(0) = C, \quad (1)$$

дается равенством

$$X = e^{At} C e^{Bt}. \quad (2)$$

Доказательство проводится непосредственно подстановкой. Этот результат (см. также упражнение 2 § 16), несмотря на свою простоту, играет важную роль в различных областях математической физики.

Упражнения

1. Получить решение, данное выше, разыскивая его в форме $X = e^{At}Y$.
2. Пусть $Y(t)$ является квадратным корнем из $X(t)$. Показать, что

$$\left(\frac{dY}{dt}\right)Y + Y\left(\frac{dY}{dt}\right) = \frac{dX}{dt} \text{ *).$$

§ 18. Уравнение $AX + XB = C$. Используя результат, установленный выше, мы можем доказать следующую теорему:

Теорема 6. Если матрица

$$X = -\int_0^{\infty} e^{At} C e^{Bt} dt \quad (1)$$

существует для всех C , то она является единственным решением уравнения

$$AX + XB = C. \quad (2)$$

Доказательство. Рассмотрим уравнение

$$\frac{dZ}{dt} = AZ + ZB, \quad Z(0) = C. \quad (3)$$

Согласно теореме 5 его решение $Z(t)$ определяется формулой (17.2).

Проинтегрируем обе части (3) в пределах от 0 до ∞ в предположении, что $\lim_{t \rightarrow \infty} Z(t) = 0$ **). В результате получаем

$$-C = A\left(\int_0^{\infty} Z ds\right) + \left(\int_0^{\infty} Z ds\right)B. \quad (4)$$

Мы видим, что матрица X , равная

$$-\int_0^{\infty} Z ds = -\int_0^{\infty} e^{As} C e^{Bs} ds,$$

удовлетворяет уравнению (2). В гл. 12 мы детально обсудим вопрос существования интеграла (1) и решения уравнения (2).

Единственность решения следует из линейности уравнения (2). Действительно, рассматривая матричное уравнение (2) как N^2

* См. упражнение 2, стр. 189. (Прим. ред.)

** Этот факт следует из существования интеграла (1) и не является дополнительным предположением. (Прим. ред.)

скалярных линейных уравнений для элементов x_{ij} матрицы X , мы видим, что из существования решения для всех матриц C следует неравенство нулю определителя коэффициентов системы. Следовательно, полученное решение единственно.

Упражнения к гл. 10

1. Рассмотрим последовательность матриц $\{X_n\}$, определенную рекуррентным соотношением

$$X_{n+1} = X_n (2I - AX_n), \quad X_0 = B.$$

Каким условиям должны удовлетворять матрицы A и B , чтобы эта последовательность сходилась, и чему равен ее предел?

2. Если матрица A является положительно определенной, а $B(t)$ положительно определена для всех $t \geq 0$, причем $|A| < |B(t)|$ при $t \geq 0$, то для любой неотрицательной функции $g(t)$ такой, что $\int_0^{\infty} g(t) dt = 1$, имеет место неравенство

$$|A| < \left| \int_0^{\infty} B(t) g(t) dt \right|.$$

3. Если A — положительно определенной матрица, причем такая, что $A < I$, то последовательность матриц, определенная посредством рекуррентного соотношения

$$X_{n+1} = X_n + \frac{1}{2}(A - X_n^2), \quad X_0 = 0,$$

сходится к положительно определенному квадратному корню из A ¹⁾.

4. Задача определения величин λ , удовлетворяющих детерминантному уравнению $|I - \lambda F_1 - \lambda^2 F_2 - \dots - \lambda^k F_k| = 0$, где F_i — квадратные матрицы N -го порядка, эквивалентна задаче определения собственных значений матрицы

$$A = \begin{vmatrix} F_1 & F_2 & \dots & F_k \\ I & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & I & 0 \end{vmatrix}.$$

См. Гадерли (K. G. Guderley), On Nonlinear Eigenvalue Problems for Matrices, J. Ind. and Appl. Math. 6 (1958), 335—353.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 1. Вопросы определения равновесия и теория устойчивости систем дифференциальных уравнений начали непосредственно и почти одновременно изучаться Ляпуновым и Пуанкаре. Подробные ссылки и обсуждение результатов могут быть найдены в гл. 13 и в книге

¹⁾ Виссер (C. Visser), Notes on Linear Operators, Proc. Acad. Sci. Amsterdam 40 (1937), 270—272.

Беллман (R. Bellman), *Stability Theory of Differential Equations*. McGraw-Hill Book Company, Inc., New York, 1954. [Русский перевод: Теория устойчивости дифференциальных уравнений, ИЛ, 1954.]

Некоторые интересные преобразования дифференциальных матричных уравнений содержатся в следующих статьях:

Баррет (J. H. Barrett), *Matrix Systems of Second Order Differential Systems*, *Portugaliae Math.* 14 (1955), 79—89;

Баррет (J. H. Barrett), *A Prufer Transformation for Matrix Differential Equations*, *Proc. Amer. Math. Soc.* 8 (1957), 510—518;

Левин (J. J. Levin), *On the Matrix Riccati Equation*, *Proc. Amer. Math. Soc.* 10 (1959), 519—524.

§ 2. Понятие матричной экспоненты является центральным во всех работах в области теории линейных функциональных уравнений. Его обобщение на класс произвольных операторов является краеугольным камнем теории полугрупп. См.

Хилле (E. Hille), *Functional Analysis and Semi-groups*, *Amer. Math. Soc. Publ.*, 1942. [Русский перевод: Функциональный анализ и полугруппы, ИЛ, 1951];

Хилле и Филлипс (E. Hille, R. Phillips), *Functional Analysis and Semi-groups*, *Amer. Math. Soc. Colloq. Publ.* 31 (1958). [Русский перевод: Функциональный анализ и полугруппы, ИЛ, 1961.]

Проблема представления $e^{At}e^{Bt}$ в виде e^{Ct} , где A и B не коммутируют, является весьма важной не только в теории групп и алгебр Ли, но и в современной квантовой теории. Этот материал послужит объектом специального исследования в следующем томе этой серии. Читатель, интересующийся деталями, может обратиться к статьям:

Магнус (W. Magnus), *Algebraic Aspects of the Theory of Systems of Linear Differential Equations*, *Research Rept. BR-3*, New York University, Institute of Mathematical Sciences, June 1953, см. также *Comm. Pure Appl. Math.* 7, № 4 (1954);

Бейкер (H. F. Baker), *On the Integration of Linear Differential Equations*, *Proc. London Math. Soc.* 34 (1902), 347—360; 35 (1903), 333—374; second series 3 (1904), 293—296;

Бейкер (H. F. Baker), *Alternants and Continuous Groups*, *Proc. London Math. Soc.*, second series 3 (1904), 24—47;

Келлер Г. и Келлер Дж. (H. B. Keller and J. B. Keller), *On Systems of Linear Ordinary Differential Equations*, *Research Rept. EM-33*, New York University, Washington Square College, Mathematics Research Group, 1957;

Хаусдорф (F. Hausdorff), *Die Symbolische exponential Formel in der Gruppentheorie*, *Saechsische Akademie der Wissenschaften zu Leipzig, Math. Phys. Klasse* 58 (1906), 19—48;

Голдберг (K. Goldberg), *The Formal Power series for $\log(e^x e^y)$* , *Duke Math. J.* 23 (1956), 13—21 *);

*) См. также Е. Б. Дынкин, *О представлении ряда $\log(e^x e^y)$ от некоммутирующих x и y через коммутаторы*, *Матем. сб.* 25 (67) (1949), 155—162.

Чен Кво-цай (Kuo-Tsai Chen), Integration of Paths, Geometric Invariants and a Generalized Baker-Hausdorff Formula, *Ann. Math.* 65 (1957), 163—178.

Другое направление в проблеме отыскания решения уравнения вида $\frac{dX}{dt} = A(t)X$, $X(0) = I$, приводит к понятию экспоненциала. Изучение этого вопроса связано с теорией мультипликативного интеграла. См. работы

Шлезингер (L. Schlesinger), Neue Grundlagen für einen Infinitesimal-kalkül der Matrizen, *Math. Z.* 33 (1931), 33—61;

Шлезингер (L. Schlesinger), Weitere Beiträge zum Infinitesimal-kalkül der Matrizen, *Math. Z.* 35 (1932), 485—501;

Раш (G Rasch), Zur Theorie and Anwendung des Produktintegrals, *J. reine angew. Math.* 171 (1934), 65—119.

Результаты более поздних исследований содержатся в работе

Магнус (W. Magnus), On the Experimental Solution of Differential Equations for a Linear Operator, *Comm. Pure Appl. Math.* 7 (1954), 649—673.

Для понимания связи с физической стороной этой проблемы см.

Фейнман (R. P. Feynman), An Operator Calculus Having Applications in Quantum Electrodynamics, *Phys. Rev.* 84 (1951), 108—128.

Наконец, упомянем, что большие усилия были сделаны в области обобщения этих результатов и методов применительно к бесконечной системе линейных дифференциальных уравнений с постоянными коэффициентами. Системы подобного типа естественным образом появляются в теории вероятностей и, в частности, в ее приложениях к физике и биологии при изучении процессов «размножения и гибели». Эти вопросы вновь будут обсуждаться в гл. 16. См. также

Арли и Борксениус (N. Arley and A. Borchsenius), On the Theory of Infinite Systems of Differential Equations and Their Applications to the Theory of Stochastic Processes and the Perturbation Theory of Quantum Mechanics, *Acta Math.* 78 (1945), 261—322;

Беллман (R. Bellman), The Boundedness of Solutions of Infinite Systems of Linear Differential Equations, *Duke Math. J.* 14 (1947), 695—706.

§ 3. Дальнейшие результаты, касающиеся норм матриц, и дальнейшие ссылки могут быть найдены в работах

Хаусхолдер (A. S. Householder), The Approximate Solution of Matrix Problems, *J. Assoc. Comp. Mach.* 5 (1958), 205—243;

Нейман и Голдстейн (J. von Neumann and H. Goldstine), Numerical Inverting of Matrices of High Order, *Bull. Amer. Math. Soc.* 53 (1947), 1021—1099;

Островский (A. Ostrowski), Über Normen von Matrizen, *Math. Z.* 63 (1955), 2—18;

Фань Цзы и Гоффман (Ku Fan and A. J. Hoffman), Some Metric Inequalities in the Space of Matrices, *Proc. Amer. Math. Soc.* 6 (1955), 111—116;

Истерфилд (T. E. Easterfield), *Matrix Norms and Vector Measures*, *Duke Math. J.* 24 (1957), 663—671.

§ 5. Читатель, знакомый с теорией интеграла Лебега, может заметить, что результат теоремы 1 может быть получен при гораздо более слабых ограничениях на матрицу $A(t)$. Мы, однако, не стремились к формулировке более общего результата, поскольку в этом не было необходимости.

§ 14. Изменение собственных значений A , вызванное изменением размерности, может быть рассмотрено с привлечением методов, изложенных в работе

Беллман и Леман (R. Bellman and S. Lehman), *Functional Equations in the Theory of Dynamic Programming, X: Resolvents, Characteristic Values and Functions*, *Duke Math. J.*, 1960.

§ 15. Вопросы неотрицательности решений функциональных уравнений играют большую роль в теории вероятностей и в математической экономике, где физические модели делают результаты интуитивно ясными. Мы вернемся к этим вопросам в гл. 16. Их рассмотрение существенно также при изучении различных классов нелинейных уравнений. См. работы

Беллман (R. Bellman), *Functional Equations in the Theory of Dynamic Programming, V: Positivity and Quasi-linearity*, *Proc. Nat. Acad. Sci. U. S. A.* 41 (1955), 743—746;

Калаба (R. Kalaba), *On Nonlinear Differential Equations, the Maximum Operation, and Monotone Convergence*, Ph. D. Thesis, New York, University, February, 1958.

Обсуждение вопроса о положительности операторов содержится в гл. 5 монографии

Бекенбах и Беллман (E. F. Beckenbach and R. Bellman), *Inequalities*, *Ergeb. Math.*, 1960. [Русский перевод: *Неравенства*, «Мир», 1965.]

Первое доказательство дано С. Карлиным; второе получено О. Таусски. Впервые этот результат опубликован в статье

Беллман, Гликсберг, Гросс (R. Bellman, I. Glicksberg and O. Gross), *On some Variational Problems Occurring in the Theory of Dynamic Programming*, *Rend. Circ. Palermo, Serie II*, 1954, 1—35.

§ 16. Результат и доказательство см. в работе

Поля (G. Pólya), *Über die Funktionalgleichung in Matrizkalkul*, *Sitzber. Akad. Berlin*, 1928, 96—99.

§ 17. Мы не можем точно указать источник этого результата, который встречается в работах многих авторов. Развернутое обсуждение этого и связанных вопросов содержится в работе

И. А. Лаппо-Данилевский, *Memoires sur la théorie des systèmes des equations differentielles lineaires*, *Тр. физ.-матем. ин-та АН СССР* 5 (1934), 7 (1935), 8 (1936).

Тот факт, что решение уравнения $\frac{dX}{dt} = AX + XB$ имеет указанную форму, становится весьма прозрачным, если рассмотреть его под углом зрения квантовой механики. См.

Хаар (D. ter Haar), *Elements of Statistical Mechanics*, Rinehart & Company, Inc., New York, 1954, 149;

Намбу (Y. Nambu), *Progr. Theoret. Phys. (Kyoto)* 4 (1949), 331;

Примас и Гунтхард (H. Primas and H. Gunthard), *Eine Methode zur direkten Berechnung ...*, *Helv. Phys. Acta* 31 (1958), 413—434.

§ 18. Рассуждения, связанные с оператором T , определенным посредством равенства $TX = AX - XB$, см. в работе

Розенблюм (M. Rosenbloom). *On the Operator Equations $BX - XA = Q$* , *Duke Math. J.* 23 (1956), 263—269.

Обобщение на оператор S , определенный как $SX = \sum A_j X B_j$, см. в работе

Люмер и Розенблюм (G. Lumer and M. Rosenbloom), *Linear Operator Equations*, *Proc. Amer. Math. Soc.* 10 (1959), 32—41.

Упражнения 2 и 3 взяты из статьи

Беллман (R. Bellman), *On the Linear Differential Equation Whose Solutions Are the Products of Solutions of Two Given Linear Differential Equations*, *Bol. Unione mat. Italiana, Ser. III, anno XII, 1957, 12—15*

Глубокая связь теории матриц с теорией поведения решений линейных дифференциальных уравнений и, следовательно, с изучением линейных колебательных систем приводит к прекрасной теории, развитой в статье

Ф. Р. Гантмахер и М. Г. Крейн, *Sur les matrices complètement non négatives et oscillatoires*, *Comp. math.* 4 (1937), 445—476 *).

Как будет отмечено в конце замечаний в гл. 16, эти результаты имеют важную область приложения в теории вероятностей. Рассмотрение уравнения

$$A \frac{d^2 x}{dt^2} + B \frac{dx}{dt} + Cx = f,$$

где A , B и C — симметрические матрицы, при определенных дополнительных допущениях относительно вида этих матриц содержится в работе

Даффин (R. J. Duffin), *A Minimax Theory for overdamped Networks*, *J. Rat. Mech and Analysis* 4 (1955), 221—233.

Исследование линейных электрических цепей приводит другим путем к интересной области матричного анализа, а именно

* См. книгу Ф. Р. Гантмахера и М. Г. Крейна, *Осцилляционные матрицы и ядра и малые колебания механических систем*, Гостехиздат, 1950.

к изучению вопроса о том, когда заданная матрица рациональных функций может рассматриваться как матричный импеданс разомкнутого n -полюсника. Здесь в качестве фундаментального понятия возникает концепция *положительной вещественной* матрицы. Последние результаты в этой области и ссылки см. в книге

Вайнберг и Следиан (L. Weinberg and P. Slepian), *Positive Real Matrices*, Hughes Research Laboratories, Culver, City, Calif., 1958.

Некоторые интересные результаты относительно связи линейных систем и уравнений Риккати см. в статье

Рейд (W. T. Reid), *Solutions of a Riccati Matrix Differential Equation as Functions of Initial Values*, *J. Math. and Mech.* 8 (1959), 221—230.

где имеются также ссылки на более ранние результаты Редхеффера.

ЯВНЫЕ РЕШЕНИЯ И КАНОНИЧЕСКИЕ ФОРМЫ МАТРИЦ

§ 1. Введение. Хотя результаты, полученные в гл. 10, весьма элегантны, с сожалением приходится отметить, что они не могут нас удовлетворить в тех случаях, когда необходим явный вид компонент решения.

В настоящей главе мы исследуем этот вопрос и покажем, что он приводит к задаче определения собственных значений и собственных векторов матриц, не являющихся, вообще говоря, симметрическими. Так же, как и в случае симметрических матриц, мы придем к изучению различных канонических форм.

§ 2. Метод Эйлера. В гл. 10 мы установили тот факт, что векторно-матричное уравнение

$$\frac{dx}{dt} = Ax, \quad x(0) = c, \quad (1)$$

обладает единственным решением, которое можно записать в виде

$$x = e^{At}c. \quad (2)$$

Здесь мы выберем другой путь. Следуя Эйлеру, начнем с рассмотрения частного решения вида $x = e^{\lambda t}c^1$, где λ — скаляр, а c^1 — вектор. При этом забудем на время о начальном условии $x(0) = c$. Подставляя это решение в уравнение (1), мы видим, что λ и c^1 связаны равенством

$$\lambda c^1 = Ac^1. \quad (3)$$

Поскольку по предположению вектор c^1 отличен от нуля, число λ должно удовлетворять характеристическому уравнению

$$|A - \lambda I| = 0. \quad (4)$$

При этом c^1 является соответствующим собственным вектором. В результате мы с совершенно другой стороны пришли к основной проблеме, рассматривавшейся в первой части книги, — к определению характеристических чисел и собственных векторов

матриц. В данном случае, однако, анализ будет одновременно более сложным и не столь завершенным, поскольку рассматриваемые матрицы не являются симметрическими.

Упражнение

1. Найти характеристические числа и собственные векторы матрицы

$$A = \begin{vmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & \ddots \\ 0 & & & \lambda \end{vmatrix}.$$

§ 3. Построение решения. Посмотрим, каким образом, следуя намеченной линии, можно построить решение, удовлетворяющее заданным начальным условиям. Обозначим, как и ранее, характеристические числа через $\lambda_1, \lambda_2, \dots, \lambda_N$. Поскольку в общем случае эти корни могут быть комплексными, они не поддаются упорядочению.

Для упрощения рассуждений будем первоначально считать, что все характеристические числа различны. Пусть c^1, c^2, \dots, c^N — набор соответствующих собственных векторов, в общем случае также комплексных.

Используем *принцип суперпозиции*, чтобы получить требуемое решение. Поскольку вектор $e^{\lambda_k t} c^k$ является решением системы при любом $k, k = 1, 2, \dots$, то при произвольных скалярах a_k линейная комбинация

$$x = \sum_{k=1}^N a_k e^{\lambda_k t} c^k \quad (1)$$

также окажется решением уравнения $\frac{dx}{dt} = Ax$.

Теперь вопрос заключается лишь в том, могут ли коэффициенты a_k быть выбраны таким образом, чтобы удовлетворить условию $x(0) = c$. Это последнее условие приводит к векторному уравнению

$$c = \sum_{k=1}^N a_k c^k. \quad (2)$$

Система (2) имеет единственное решение, если матрица

$$C = [c^1, c^2, \dots, c^N] \quad (3)$$

невырождена. При этом, как и ранее, столбцами матрицы C служат собственные векторы c^i .

§ 4. Невырожденность матрицы C . Имеется несколько путей доказательства того факта, что матрица C является невырожденной. Следуя одному из них, поступим следующим образом. Предположим, что $|C|=0$. Тогда существует нетривиальный набор скаляров b_1, b_2, \dots, b_N такой, что

$$0 = \sum_{k=1}^N b_k c^k. \quad (1)$$

Отсюда следует, что функция

$$z = \sum_{k=1}^N b_k c^k e^{\lambda_k t} \quad (2)$$

является решением уравнения $\frac{dx}{dt} = Ax$, удовлетворяющим начальному условию $x(0)=0$. Из теоремы единственности вытекает, что $z=0$ при всех t .

Мы должны далее показать, что при различных λ_i соотношение

$$0 = \sum_{k=1}^N b_k c^k e^{\lambda_k t} \quad (3)$$

не может быть выполнено при любых t , если только скаляры b_k и векторы c^k все не равны нулю.

Без ограничения общности примем, что $b_N \neq 0$. Деление (3) на $e^{\lambda_1 t}$ и дифференцирование по t дают

$$\sum_{k=2}^N b_k c^k (\lambda_k - \lambda_1) e^{(\lambda_k - \lambda_1)t} = 0. \quad (4)$$

Поскольку λ_k различны, то и все разности $\lambda_k - \lambda_1$ также различны. Теперь мы можем провести доказательство по индукции или просто повторить предыдущую операцию несколько раз.

Ясно, что, следуя последним путем, мы придем к соотношению вида

$$b_N c^N (\lambda_N - \lambda_1) (\lambda_N - \lambda_2) \dots (\lambda_N - \lambda_{N-1}) e^{(\lambda_N - \lambda_1)t} = 0. \quad (5)$$

Поскольку вектор c^N не равен нулю, мы должны иметь $b_N = 0$, что противоречит допущению.

§ 5. Другой метод. Начав с соотношения (4.1), умножим обе части на A . В результате получим

$$0 = \sum_{k=1}^N b_k \lambda_k c^k. \quad (1)$$

Повторяя эту операцию $N-1$ раз, приходим к системе линейных уравнений

$$0 = \sum_{k=1}^N b_k \lambda_k^r c^k, \quad r=0, 1, 2, \dots, N-1. \quad (2)$$

Рассмотрим лишь i -е компоненты векторов c^k , т. е. c_i^k . Это дает систему скалярных уравнений

$$0 = \sum_{k=1}^N b_k \lambda_k^r c_i^k, \quad r=0, 1, \dots, N-1. \quad (3)$$

Поскольку не все b_k равны нулю и, кроме того, все c_i^k не могут быть равны нулю, когда индексы i и k пробегают все значения от 1 до N , то система (3) обязательно имеет нетривиальные решения.

Отсюда следует, что должен быть равен нулю определитель коэффициентов системы

$$|\lambda_k^r|, \quad k=1, 2, \dots, N; \quad r=0, 1, \dots, N-1. \quad (4)$$

Этого, однако, не может быть, в чем мы убедимся в следующем параграфе.

§ 6. Определитель Вандермонда. Определитель вида

$$|\lambda_k^r| = \begin{vmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_N \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_N^2 \\ \dots & \dots & \dots & \dots \\ \lambda_1^{N-1} & \lambda_2^{N-1} & \dots & \lambda_N^{N-1} \end{vmatrix} \quad (1)$$

носит название *определителя Вандермонда*.

Мы хотим показать, что $|\lambda_k^r| \neq 0$, если $\lambda_i \neq \lambda_j$ для всех $i \neq j$. Простейшим путем доказательства является вычисление определителя. Рассматривая этот определитель как полином от λ_1 степени $N-1$, мы видим, что он имеет следующие корни:

$$\lambda_1 = \lambda_2, \quad \lambda_1 = \lambda_3, \quad \dots, \quad \lambda_1 = \lambda_N.$$

Действительно, определитель равен нулю, если два его столбца совпадают. Отсюда следует, что

$$|\lambda_k^r| = (\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3) \dots (\lambda_1 - \lambda_N) q(\lambda_2, \lambda_3, \dots, \lambda_N), \quad (2)$$

где функция q не зависит от λ_1 .

Аналогично, если считать определитель полиномом степени $N-1$ от λ_2 , то он должен иметь сомножителем выражение

$$(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3) \dots (\lambda_2 - \lambda_N).$$

Продолжая действовать таким образом, мы убедимся, что

$$|\lambda_k| = \prod_{1 < i < j < N} (\lambda_j - \lambda_i) \varphi(\lambda_1, \lambda_2, \dots, \lambda_N), \quad (3)$$

где функция φ является полиномом от λ_i . Однако сравнивая наибольшие степени, в которых λ_j входит в правую и левую части, легко увидеть, что φ должна быть константой. Исследование коэффициентов при члене $\lambda_1 \lambda_2^2 \dots \lambda_N^{N-1}$ в обеих частях равенства дает значение $\varphi = 1$.

Из соотношения (3) получаем вывод о том, что $|\lambda_k| \neq 0$, если $\lambda_i \neq \lambda_j$ при $i \neq j$.

Упражнение

1. Используя аналогичные рассуждения, вычислить определитель Коши

$$\left| \frac{1}{\lambda_i + \mu_j} \right|, \quad i, j = 1, 2, \dots, N.$$

§ 7. Явная форма решения линейного дифференциального уравнения. Диагональные матрицы. Рассмотрим подход к решению дифференциального уравнения, отличный от предыдущего.

В уравнении

$$\frac{dx}{dt} = Ax, \quad x(0) = c, \quad (1)$$

произведем замену переменных $x = Ty$, где T — постоянная невырожденная матрица, которую мы определим в дальнейшем. Уравнение для y имеет вид

$$\frac{dy}{dt} = T^{-1}ATy, \quad y(0) = T^{-1}c. \quad (2)$$

Можно ли выбором матрицы T настолько упростить систему, чтобы она допускала непосредственное интегрирование?

Предположим, что мы нашли такую матрицу T , что матрица $T^{-1}AT$ является диагональной:

$$T^{-1}AT = \begin{bmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ 0 & & & \mu_N \end{bmatrix}. \quad (3)$$

Если это выполнено, то уравнения (2) распадаются на N независимых уравнений вида

$$\frac{dy_i}{dt} = \mu_i y_i, \quad y_i(0) = c_i', \quad i = 1, 2, \dots, N. \quad (4)$$

Эти последние имеют простейшие решения: $y_i = e^{\mu_i t} c_i'$. После этого исходный вектор x легко определяется по известному вектору y .

§ 8. **Диагонализация матрицы.** Рассмотрим весьма интересную, но трудную проблему диагонализации матрицы A . Как нам известно из первой части книги, если матрица A является симметрической, то матрица T , обладающая требуемыми свойствами, всегда может быть найдена. При этом $T^{-1} = T'$. Как мы увидим сейчас, существуют другие важные классы матриц, которые могут быть приведены к диагональной форме. Не менее важен тот факт, что, помимо диагонального, имеются также другие полезные канонические представления матриц.

Рассмотрим, однако, общий случай. Сразу же отметим, что набор величин $\{\mu_i\}$ должен быть точно таким же, как и $\{\lambda_i\}$, поскольку характеристические числа матрицы $T^{-1}AT$ те же, что и у матрицы A .

Как и в предшествующих главах, отсюда следует, что столбцами T являются собственные векторы матрицы A . Обратное, если все λ_i различны, а T — матрица, столбцами которой являются соответствующие собственные векторы A , то

$$AT = T \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix}. \quad (1)$$

Мы получили, таким образом, следующий важный результат.

Теорема 1. Если характеристические числа матрицы A различны, то существует матрица T такая, что

$$T^{-1}AT = \begin{vmatrix} & & & 0 \\ \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix}. \quad (2)$$

Как мы увидим, предположение об отсутствии кратных характеристических чисел является весьма важным. В § 10 мы покажем, что такое простое представление, как (2), не может быть получено в общем случае.

Упражнения

1. Показать, что теорема Гамильтона—Кэли справедлива для матриц с различными характеристическими числами.

2. Показать, что предположение о существовании N различных характеристических чисел может быть заменено требованием существования N линейно независимых собственных векторов.

§ 9. Связь между двумя подходами. Как мы установили, один из методов решения линейного уравнения

$$\frac{dx}{dt} = Ax, \quad x(0) = c, \quad (1)$$

приводит к выражению $x = e^{At}c$, в то время как другой, основанный на введении характеристических чисел и векторов, приводит к скалярным экспоненциальным функциям.

Установим связь между этими подходами (пока еще в предположении $\lambda_i \neq \lambda_j$), которая должна существовать вследствие единственности решения. В уравнении

$$\frac{dX}{dt} = AX, \quad X(0) = I, \quad (2)$$

произведем замену переменных $X = TY$, где T выбрана в соответствии с § 8. Тогда для Y получим уравнение

$$\frac{dY}{dt} = T^{-1}ATY, \quad Y(0) = T^{-1}, \quad (3)$$

или

$$\frac{dY}{dt} = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix} Y, \quad Y(0) = T^{-1}. \quad (4)$$

Отсюда следует, что

$$Y = \begin{vmatrix} e^{\lambda_1 t} & & & 0 \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_N t} \end{vmatrix} T^{-1} \quad (5)$$

и, следовательно,

$$X = e^{At} = T \begin{vmatrix} e^{\lambda_1 t} & & & 0 \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_N t} \end{vmatrix} T^{-1}. \quad (6)$$

Подчеркнем еще раз, что это представление получено в предположении, что характеристические числа матрицы A различны.

Упражнения

1. Вывести формулу (6) непосредственно из представления e^{At} в виде ряда и диагонального представления матрицы A .

2. Использовать (6) для доказательства того, что $|e^{At}| = e^{\text{Sp } At}$.

3. Используя метод продолжения, показать, что этот результат справедлив для любой квадратной матрицы A .

§ 10. Кратные характеристические числа. Обратимся к исследованию случая кратных характеристических чисел матрицы A . Чтобы оценить некоторые из трудностей, с которыми мы можем встретиться в этом случае, начнем с доказательства того факта, что не всегда возможно получить каноническое представление вида (8.2).

Теорема 2. Существуют матрицы, которые не сводятся к диагональной форме (8.2) посредством невырожденного преобразования. Иначе говоря, существуют матрицы N -го порядка, не имеющие N линейно независимых собственных векторов.

Доказательство. Рассмотрим матрицу частного вида

$$A = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}. \quad (1)$$

Если существует такая матрица T , что

$$T^{-1}AT = \begin{vmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{vmatrix}, \quad (2)$$

то столбцы T должны быть собственными векторами A . Определим характеристические числа и собственные векторы.

Характеристическое уравнение

$$\begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 = 0 \quad (3)$$

имеет корень $\lambda = 1$ двойной кратности. Собственные векторы определяются из уравнений

$$\begin{aligned} x_1 + x_2 &= x_1, \\ x_2 &= x_2. \end{aligned} \quad (4)$$

Отсюда следует, что $x_2 = 0$, а x_1 произвольно. Следовательно, все собственные векторы с точностью до скалярного множителя равны

$$c' = \begin{vmatrix} 1 \\ 0 \end{vmatrix}. \quad (5)$$

Это означает, что матрица T является вырожденной.

Неожиданным является то, что произвольная матрица не обязательно имеет полный набор собственных векторов. Это позволяет нам оценить, сколь существенным ограничением является требование симметрии матриц.

Тот факт, что диагонализация матриц не всегда возможна, в высшей степени затрудняет исследование матриц общего вида, но в равной мере и повышает интерес к этому исследованию. Как мы увидим ниже, имеется много способов, чтобы обойти некоторые из возникающих препятствий. Используемые методы базируются на различных канонических представлениях и на теоремах об аппроксимации.

Упражнения

1. Почему нельзя воспользоваться методом продолжения для вывода диагонального представления матриц общего вида, основываясь на результатах, полученных для матриц с различными характеристическими числами?

2. Найдя решение системы дифференциальных уравнений

$$\frac{dx_1}{dt} = x_1 + x_2, \quad x_1(0) = c_1,$$

$$\frac{dx_2}{dt} = x_2, \quad x_2(0) = c_2,$$

показать, что матрица

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

не может быть приведена к диагональной форме.

§ 11. Каноническая форма Жордана. Введем теперь каноническую форму, к которой может быть приведена произвольная квадратная матрица. Применение этой формы облегчает рассмотрение ряда проблем. Поскольку соответствующий вывод является достаточно громоздким и излагается в большинстве учебников по линейной алгебре и, кроме того, поскольку в наших рассуждениях можно так или иначе обойтись без использования этого вывода, то мы просто выпишем конечный результат без доказательства.

Теорема 3. Пусть $L_k(\lambda)$ означают матрицы размера $k \times k$ вида

$$L_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \lambda & 1 \\ 0 & 0 & \dots & \dots & \lambda \end{bmatrix}. \quad (1)$$

причем $L_1(\lambda) = \lambda$. Существует матрица T такая, что

$$T^{-1}AT = \begin{pmatrix} L_{k_1}(\lambda_1) & & & 0 \\ & L_{k_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & L_{k_r}(\lambda_r) \end{pmatrix}, \quad (2)$$

где $k_1 + k_2 + \dots + k_r = N$. Величины λ_i являются характеристическими числами матрицы A , не обязательно различными.

Представление матрицы A в виде (2) называется канонической формой Жордана.

Например, если A — матрица третьего порядка, имеющая корень λ_1 кратности три, то она может быть приведена к одной из трех форм:

$$A_1 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}, \quad (3)$$

Упражнения

1. Используя жорданову каноническую форму, определить для произвольной A матрицу e^{At} .

2. Используя этот результат, получить необходимые и достаточные условия стремления e^{At} к нулю при $t \rightarrow \infty$.

3. Доказать, что все A_i , представленные соотношениями (3), различны в том смысле, что не существует матрицы T такой, что $T^{-1}A_iT = A_j$ при $i \neq j$. Дать два доказательства — алгебраическое и основанное на решении соответствующих систем дифференциальных уравнений.

4. Показать, что матрица $L_k(\lambda) - \lambda I$, возведенная в k -ю степень, дает нулевую матрицу.

§ 12. Кратные характеристические числа (другой метод). Используя жорданову каноническую форму, мы обладаем систематическим методом получения точной аналитической структуры решений линейных дифференциальных уравнений с постоянными коэффициентами. Так, если уравнение сведено к виду

$$\begin{aligned} \frac{dy_1}{dt} &= \lambda y_1 + y_2, & y_1(0) &= c_1, \\ \frac{dy_2}{dt} &= \lambda y_2, & y_2(0) &= c_2, \end{aligned} \quad (1)$$

мы находим $y_2 = e^{\lambda t} c_2$, после чего y_1 определяется как решение неоднородного уравнения

$$\frac{dy_1}{dt} = \lambda y_1 + e^{\lambda t} c_2, \quad y_1(0) = c_1. \quad (2)$$

Введение интегрирующего множителя $e^{-\lambda t}$ дает

$$\frac{d}{dt}(y_1 e^{-\lambda t}) = c_1 \quad (3)$$

и, следовательно,

$$y_1 = c_1 e^{\lambda t} + c_2 t e^{\lambda t}. \quad (4)$$

Это поясняет, каким путем члены $t e^{\lambda t}$, или в общем случае $t^k e^{\lambda t}$, появляются в решении уравнения в случае кратных корней. Обсудим теперь другой подход.

Если A имеет различные характеристические числа $\lambda_1, \lambda_2, \dots, \lambda_N$, то мы можем записать N частных решений уравнения $\frac{dy}{dt} = Ay$ вида $y = c(\lambda) e^{\lambda t}$, $\lambda = \lambda_1, \lambda_2, \dots, \lambda_N$, где $c(\lambda)$ — полином от λ . Мы можем показать это следующим образом.

Как нам известно, для получения решения мы подставляем в уравнение функцию $y = e^{\lambda t}$ и находим λ и c из уравнения

$$Ac = \lambda c, \quad \text{или} \quad (A - \lambda I)c = 0. \quad (5)$$

После того как параметр λ определен из характеристического уравнения

$$|A - \lambda I| = 0,$$

мы должны найти решение линейной системы (5). Пусть

$$b_{ij} = |A - \lambda I|_{ij} \quad (6)$$

означают алгебраические дополнения элемента $a_{ij} - \lambda \delta_{ij}$ определителя $|A - \lambda I|$.

Тогда решением системы (5), обладающим тем свойством, что c_i являются полиномами от λ , будет

$$c_i = b_{ii}, \quad i = 1, 2, \dots, N. \quad (7)$$

Может оказаться, что это решение является тривиальным в том смысле, что все c_i равны нулю. В этом случае мы можем попытаться определить другой набор c_i , а именно:

$$c_i = b_{ki}, \quad i = 1, 2, \dots, N, \quad (8)$$

для какого-либо $k \neq 1$.

Покажем, что какой-либо из этих наборов будет нетривиальным, если характеристические числа A различны. В частности, покажем, что для произвольного характеристического числа $\lambda = \lambda_i$ все b_{ii} не могут быть равны нулю, если λ_1 — простой корень характеристического уравнения.

Рассмотрим характеристическое уравнение

$$f(\lambda) = |A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} - \lambda & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} - \lambda \end{vmatrix} = 0. \quad (9)$$

Используя правило дифференцирования определителя, найдем

$$\begin{aligned} f'(\lambda) &= \begin{vmatrix} -1 & 0 & \dots & 0 \\ a_{21} & a_{22} - \lambda & \dots & a_{2N} \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & \dots & a_{NN} - \lambda \end{vmatrix} + \\ &+ \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1N} \\ 0 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & \dots & a_{NN} - \lambda \end{vmatrix} + \dots \\ &\dots + \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} - \lambda & \dots & a_{2N} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 \end{vmatrix} = -b_{11} - b_{22} - \dots - b_{NN}. \quad (10) \end{aligned}$$

Если для $\lambda = \lambda_1$ все $b_{ii}(\lambda_1) = 0$ ($i = 1, 2, \dots, N$), то и $f'(\lambda_1) = 0$. В противоречие с допущением это означает, что корень λ_1 является кратным.

Таким образом, решения желаемого вида могут быть найдены этим путем.

Если имеются два различных корня λ_1 и λ_2 , то функция

$$z = \frac{c(\lambda_1) e^{\lambda_1 t} - c(\lambda_2) e^{\lambda_2 t}}{\lambda_1 - \lambda_2} \quad (11)$$

также является решением. Случай, когда λ_1 становится кратным, может рассматриваться как предельный при $\lambda_2 \rightarrow \lambda_1$.

Вычисляя предел (11) при $\lambda_2 \rightarrow \lambda_1$, мы находим выражение, которое может претендовать на роль решения:

$$z = \left[\frac{dc(\lambda)}{d\lambda} e^{\lambda t} + t e^{\lambda t} c(\lambda) \right] \Big|_{\lambda=\lambda_1} \quad (12)$$

Мы оставляем читателю в качестве упражнения строгое обоснование этого факта. Укажем лишь, что возможны два пути доказательства: непосредственная проверка и вывод из общих теорем о непрерывной зависимости решения от матрицы A .

Упражнение

1. Как получить решение, используя непосредственный алгебраический подход, если λ_1 — двукратный корень?

§ 13. Треугольная форма матрицы. Теорема Шура. Рассмотрим теперь одну из наиболее полезных теорем в теории матриц.

Теорема 4. Если дана матрица A , то существует унитарная матрица T такая, что

$$T^*AT = \begin{vmatrix} \lambda_1 & b_{12} & \dots & b_{1N} \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \\ 0 & & & & \end{vmatrix}, \quad (1)$$

где, как указывает обозначение, элементы под главной диагональю равны нулю.

Доказательство. Будем действовать по индукции, начав с двумерного случая. Пусть λ_1 — характеристическое число матрицы A , а c^1 — соответствующий собственный вектор, нормированный условием $(c^1, \bar{c}^1) = 1$. Рассмотрим матрицу T , первым столбцом которой является c^1 , а второй выбран так, чтобы T была унитарной. Тогда, вычисляя $T^{-1}AT$ как произведение $T^{-1}(AT)$, получим, что

$$T^{-1}AT = \begin{vmatrix} \lambda_1 & b_{12} \\ 0 & b_{22} \end{vmatrix}. \quad (2)$$

Элемент b_{22} должен быть равен λ_2 , поскольку матрица $T^{-1}AT$ имеет те же характеристические числа, что и A .

Покажем теперь, что мы можем использовать приведение матрицы порядка N к треугольной форме для доказательства возможности получения треугольной формы в случае матрицы порядка $N+1$. Пусть, как и ранее, c^1 — нормированный собственный вектор матрицы порядка $N+1$, соответствующий λ_1 . Выберем векторы a^1, a^2, \dots, a^N так, чтобы матрица T_1 , столбцами которой являются $c^1, a^1, a^2, \dots, a^N$, была унитарной. Тогда так же, как и для случая $N=2$, получим

$$T_1^{-1}AT_1 = \begin{vmatrix} \lambda_1 & b'_{12} & \dots & b'_{1, N+1} \\ 0 & & & \\ & B_N & & \\ & & \ddots & \\ & & & 0 \end{vmatrix}, \quad (3)$$

где B_N — матрица N -го порядка. Поскольку характеристическое уравнение правой части имеет вид

$$(\lambda_1 - \lambda) |B_N - \lambda I| = 0, \quad (4)$$

то характеристическими числами B_N являются величины $\lambda_1, \lambda_2, \dots, \lambda_{N+1}$, т. е. N остальных характеристических чисел матрицы A . По предположению индукции существует унитарная матрица T_N такая, что

$$T_N^{-1} B_N T_N = \begin{vmatrix} \lambda_1 & c_{12} & \dots & c_{1N} \\ & \lambda_2 & & c_{2N} \\ & & \ddots & \vdots \\ & & & \lambda_{N+1} \\ 0 & & & & \end{vmatrix}. \quad (5)$$

Пусть T_{N+1} — унитарная матрица $(N+1)$ -го порядка, построенная следующим образом:

$$T_{N+1} = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & T_N & \\ \vdots & & & \\ 0 & & & \end{vmatrix}. \quad (6)$$

Рассмотрим равенство

$$(T_1 T_{N+1})^{-1} A (T_1 T_{N+1}) = T_{N+1}^{-1} (T_1^{-1} A T_1) T_{N+1} = \begin{vmatrix} \lambda_1 & b_{12} & \dots & b_{1, N+1} \\ & \lambda_2 & & \vdots \\ & & \ddots & \vdots \\ & & & \lambda_{N+1} \\ 0 & & & & \end{vmatrix}. \quad (7)$$

Матрица $T_1 T_{N+1}$ и является, таким образом, той унитарной матрицей, которая приводит A к треугольной форме.

Упражнения

1. Показать, что если матрица A имеет k простых корней, то существует такая матрица T , что

$$T^{-1} A T = \begin{vmatrix} \lambda_1 & 0 & \dots & 0 & b_{1, k+1} & \dots & b_{1N} \\ 0 & \lambda_2 & \dots & 0 & b_{2, k+1} & \dots & b_{2N} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_k & b_{k, k+1} & \dots & b_{kN} \\ 0 & 0 & \dots & 0 & \lambda_{k+1} & \dots & b_{k+1, N} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & \lambda_N \end{vmatrix}.$$

2. Используя введенную выше треугольную форму, получить решение уравнения $\frac{dx}{dt} = Ax$, $x(0) = c$.

3. Получить необходимое и достаточное условие отсутствия в решении членов $te^{\lambda t}$.

4. Найти общее решение уравнения $x' + Ax = 0$, где A — положительно определенная матрица.

5. Использовать треугольную форму для получения необходимых и достаточных условий того, что $e^{At} \rightarrow 0$ при $t \rightarrow \infty$.

6. Рассмотреть разностное уравнение $x(n+1) = Ax(n)$, $n=0, 1, \dots$, где $x(0) = c$. Показать, что его решение имеет вид: $x(n) = A^n c$.

7. Найти частные решения, полагая $x(n) = \lambda^n c$, и определить λ и c .

8. Показать, что общее решение разностного уравнения имеет форму

$$x(n) = \sum_{i=1}^N p_i(n) \lambda_i^n,$$

где $p_i(n)$ — векторы с компонентами, являющимися полиномами от n степени не выше $N-1$.

9. Каково необходимое и достаточное условие того, что каждое решение уравнения $x(n+1) = Ax(n)$ стремится к нулю при $n \rightarrow \infty$?

10. Рассмотреть разностное уравнение $x[(n+1)\Delta] = x[n\Delta] + A\Delta x[n\Delta]$, $n=0, 1, 2, \dots$, $x(0) = c$, где Δ — положительная скалярная величина. Показать, что $x[n\Delta] \rightarrow x(t)$ при $\Delta \rightarrow 0$, где $x(t)$ — решение дифференциального уравнения

$\frac{dx}{dt} = Ax$, $x(0) = c$, при условии, что $n\Delta \rightarrow t$. Дать два доказательства: одно, использующее явную форму решения, и другое — без нее.

§ 14. Нормальные матрицы. Вещественная матрица, перестановочная со своей транспонированной или, в более общем случае, комплексная матрица, перестановочная со своей транспонированной и комплексно сопряженной, называется *нормальной*. Для нормальной матрицы мы имеем

$$AA^* = A^*A. \quad (1)$$

Польза введения этого понятия выясняется следующей теоремой.

Теорема 5. Если матрица A нормальна, то она может быть приведена к диагональной форме посредством унитарного преобразования.

Доказательство. Как нам уже известно, мы можем найти унитарную матрицу T , приводящую A к треугольному виду

$$A = T \begin{vmatrix} \lambda_1 & b_{12} & \dots & b_{1N} \\ & \lambda_2 & \dots & b_{2N} \\ & & \ddots & \\ & & & \lambda_N \end{vmatrix} T^{-1} = TST^{-1}, \quad (2)$$

Вследствие унитарности T имеем

$$A^* = T \begin{vmatrix} \bar{\lambda}_1 & & & 0 \\ \bar{b}_{12} & \bar{\lambda}_2 & & \\ \vdots & \ddots & \ddots & \\ \bar{b}_{1N} & & & \bar{\lambda}_N \end{vmatrix} T^{-1}. \quad (3)$$

Поэтому

$$AA^* = TSS^*T^{-1} = A^*A = TS^*ST^{-1}. \quad (4)$$

Отсюда

$$SS^* = S^*S. \quad (5)$$

Приравнивая соответствующие элементы справа и слева, получим

$$b_{ij} = 0, \quad 1 \leq i < j \leq N.$$

Это означает, что треугольная форма (2) в действительности является диагональной.

Упражнения

1. Если все характеристические числа нормальной матрицы A вещественны, то она является эрмитовой и приводится к диагональной форме с помощью унитарного преобразования.

2. Показать, что $(Ax, \bar{A}x) = (A^*x, \bar{A}^*x)$ в случае, если A нормальна.

3. Показать, что из нормальности A вытекает нормальность матрицы $A - \lambda I$.

4. Показать, что если A нормальна, то x является собственным вектором матрицы A тогда и только тогда, когда он является собственным вектором матрицы A^* . Для доказательства можно использовать результат предыдущего упражнения.

5. Если A нормальна, то собственные векторы x и y , принадлежащие различным собственным значениям, ортогональны в том смысле, что $(x, y) = 0$.

6. Используя этот факт, доказать теорему 5.

7. Рассмотреть теорему 5 и доказать обратное утверждение.

8. Если матрица B нормальна и существует угол θ такой, что $Ae^{i\theta} + A^*e^{-i\theta} \geq 0$, где $A^2 = B$, то A нормальна; см. Путнам (Putnam), Proc. Amer. Math. Soc. 8 (1957), 768—769.

9. Если A и B перестановочны и A нормальна, то A^* и B перестановочны.

§ 15. Теорема об аппроксимации. Установим очень полезный результат, который часто используется при рассмотрении вопросов, связанных с квадратными матрицами общего вида.

ϵ к нулю приведет либо к тому, что матрица T станет вырожденной, либо к тому, что соответствующий предел перестанет существовать.

§ 16. Другая теорема об аппроксимации. Следующая теорема, как и предыдущая, может быть использована для сведения доказательств в случае матриц общего вида к доказательствам, использующим диагональную форму.

Теорема 7. Для заданной матрицы A всегда можно найти матрицу B с различными собственными значениями, причем такую, что $\|A - B\| \leq \epsilon$, где ϵ — любая наперед заданная константа.

Доказательство. Рассмотрим матрицу $A + G$, где элементы g_{ij} матрицы G — независимые комплексные переменные. Если $A + G$ имеет кратные характеристические числа, то $f(\lambda) = \det(A + G - \lambda I)$ и $f'(\lambda)$ имеют общий корень. При этом результат этих двух полиномов $R(G)$, являющийся полиномом от g_{ij} , должен быть равен нулю. Мы хотим показать, что можно найти набор величин g_{ij} , для которых сумма $\sum_{i,j} |g_{ij}|$ произвольно мала, причем $R(G) \neq 0$. Альтернативой этому является утверждение, что результат $R(G)$, рассматриваемый как полином от g_{ij} , тождественно равен нулю в окрестности начала координат пространства g_{ij} (*). Это означает, что $f(\lambda)$ всегда имеет кратный корень. Рассмотрим, однако, следующий набор g_{ij} :

$$\begin{aligned} g_{ij} &= -a_{ij} \quad \text{при } i \neq j, \\ g_{ii} &= i - a_{ii}, \quad i = 1, 2, \dots, N. \end{aligned} \quad (1)$$

Очевидно, что матрица $A + G$ не имеет при этом кратных корней. Следовательно, $R(G)$ не равняется нулю тождественно, и мы можем найти матрицу $B = A + G$, обладающую требуемыми свойствами.

Упражнение

1. Построить доказательство этого утверждения на основе того факта, что матрица A может быть приведена к треугольной форме (это желательно сделать потому, что предыдущее доказательство, несмотря на свою строгость и краткость, использует концепцию и свойства результата двух полиномов, строго доказать которые не так легко, как это может показаться).

§ 17. Теорема Гамильтона — Кэли. Используя теорему об аппроксимации, установленную в § 16, мы можем, наконец, доказать во всей полноте знаменитую теорему Гамильтона — Кэли.

Теорема 8. Каждая матрица удовлетворяет своему характеристическому уравнению.

*) И, следовательно, при всех g_{ij} . (Прим. ред.)

Доказательство. Пусть $A + G$ — матрица, имеющая различные собственные значения, причем $\|G\| \leq \epsilon$. Тогда, как нам известно, $A + G$ удовлетворяет своему характеристическому уравнению. Характеристический полином $f(\lambda, G) = |A + G - \lambda I|$ матрицы $A + G$, рассматриваемый как полином от λ , имеет коэффициенты, являющиеся полиномами от элементов g_{ij} матрицы G . Эти коэффициенты зависят от g_{ij} непрерывным образом; следовательно,

$$\lim f(A + G, G) = f(A).$$

Поскольку $f(A + G, G) = 0$, то и $f(A) = 0$.

Упражнения

1. Доказать теорему Гамильтона—Кэли, используя жорданову каноническую форму.

2. Доказать теорему Гамильтона—Кэли в предположении, что A имеет полный набор линейно независимых собственных векторов.

3. Использовать теорему Гамильтона—Кэли для представления невырожденной матрицы A^{-1} как полинома от A .

§ 18. Другое доказательство теоремы Гамильтона—Кэли. Зная результат, легко предложить серию коротких и элегантных доказательств. Укажем одно из них, имеющее чисто алгебраическую природу и не опирающееся на свойство непрерывности.

Рассмотрим матрицу $(A - \lambda I)^{-1}$, обратную по отношению к $A - \lambda I$ при значениях λ , не совпадающих с характеристическими числами. Мы видим, что

$$(A - \lambda I)^{-1} = B(\lambda)/f(\lambda), \quad (1)$$

причем элементы матрицы $B(\lambda)$ имеют степень не выше $N-1$, а $f(\lambda) = |A - \lambda I|$. Следовательно,

$$(A - \lambda I) B(\lambda) = f(\lambda) I. \quad (2)$$

Запишем

$$\begin{aligned} B(\lambda) &= \lambda^{N-1} B_{N-1} + \lambda^{N-2} B_{N-2} + \dots + B_0, \\ f(\lambda) &= (-1)^N \lambda^N + c_1 \lambda^{N-1} + \dots + c_N, \end{aligned} \quad (3)$$

где матрицы B_i не зависят от λ .

Приравнивая коэффициенты в (2), получим серию соотношений:

$$\begin{aligned} -B_{N-1} &= (-1)^N I, \\ AB_{N-1} - B_{N-2} &= c_1 I, \\ AB_{N-2} - B_{N-3} &= c_2 I, \\ &\dots \end{aligned} \quad (4)$$

и т. д. Мы видим, что каждая из матриц B_i является полиномом от A со скалярными коэффициентами и, следовательно, $B_i A = A B_i$ для всех i . Это означает, что тождество (2) справед-

ливо не только для скаляра λ , но также и для любой матрицы, перестановочной с A .

В частности, оно сохраняется при замене λ на A . Отсюда и следует желаемый результат, $f(A) = 0$.

§ 19. Линейные уравнения с периодическими коэффициентами. Изучим теперь проблему решения линейного дифференциального уравнения с периодическими коэффициентами:

$$\frac{dx}{dt} = P(t)x, \quad x(0) = c, \quad (1)$$

где $P(t)$ — непрерывная функция от t , удовлетворяющая для всех t условию

$$P(t+1) = P(t). \quad (2)$$

Эта проблема неожиданно является одной из чрезвычайно сложных. Исследование даже относительно простого скалярного уравнения (уравнения Матье)

$$\frac{d^2u}{dt^2} + (a + b \cos t)u = 0 \quad (3)$$

наталкивается на многие трудности, так что имеется специальная теория, посвященная этому уравнению.

Задача получения канонического представления решения (1) приводит нас к обсуждению проблемы, представляющей самостоятельный интерес, а именно к представлению невырожденной матрицы в виде матричной экспоненты.

Как обычно, удобно рассмотреть первоначально дифференциальное матричное уравнение.

Мы докажем следующую теорему:

Теорема 9. Решение уравнения

$$\frac{dX}{dt} = P(t)X, \quad X(0) = I, \quad (4)$$

где $P(t)$ — периодическая функция с периодом, равным единице, может быть записано в виде

$$X(t) = Q(t)e^{Ct}, \quad (5)$$

где $Q(t)$ — периодическая функция с тем же периодом, что и $P(t)$, а C — постоянная матрица.

Доказательство. Из периодичности $P(t)$ ясно, что если $X(t)$ — решение уравнения (4), то и $X(t+1)$ также является решением. Поскольку матрица $X(1)$ невырожденная, то мы видим, что $X(t+1)X(1)^{-1}$ является решением уравнения (4), удовлетворяющим начальному условию. По теореме единственности

имеем

$$X(t+1)X(1)^{-1} = X(t). \quad (6)$$

Предположим, что $X(1)$ можно записать в виде

$$X(1) = e^C. \quad (7)$$

Рассмотрим тогда матрицу $Q(t) = X(t)e^{-Ct}$. Мы имеем

$$\begin{aligned} Q(t+1) &= X(t+1)e^{-C(t+1)} = X(t+1)e^{-t}e^{-C} = \\ &= X(t+1)X(1)^{-1}e^{-Ct} = X(t)e^{-Ct} = Q(t). \end{aligned} \quad (8)$$

Это доказывает утверждение (5). Таким образом, осталось доказать законность представления (7).

§ 20. Представление невырожденной матрицы в виде экспоненты.

Докажем теперь следующую теорему:

Теорема 10. *Невырожденная матрица может быть представлена в виде матричной экспоненты.*

Доказательство. Пусть матрица A является невырожденной и имеет различные собственные значения. В этом случае для нее справедливо соотношение

$$A = T \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \\ 0 & & & & \end{vmatrix} T^{-1}. \quad (1)$$

Тогда, очевидно, что $A = e^C$, где

$$C = T \begin{vmatrix} \log \lambda_1 & & & 0 \\ & \log \lambda_2 & & \\ & & \ddots & \\ & & & \log \lambda_N \\ 0 & & & & \end{vmatrix} T^{-1}, \quad (2)$$

что и доказывает результат для этого случая. Если A имеет кратные корни, то мы можем с равным эффектом использовать каноническую жорданову форму.

Пусть

$$A = T \begin{vmatrix} L_{k_1}(\lambda_1) & & & 0 \\ & L_{k_2}(\lambda_2) & & \\ & & \ddots & \\ & & & L_{k_r}(\lambda_r) \\ 0 & & & & \end{vmatrix} T^{-1}. \quad (3)$$

Достаточно показать, что каждая из матриц $L_{k_i}(\lambda_i)$ имеет логарифм. Действительно, если $B_i = \log L_{k_i}(\lambda_i)$, то матрица

$$B = T \begin{vmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_r \end{vmatrix} T^{-1} \quad (4)$$

является логарифмом A .

В упражнении 4 к § 11 мы заметили, что

$$[L_k(\lambda) - \lambda I]^k = 0. \quad (5)$$

Отсюда следует, что формальный логарифм

$$\begin{aligned} B = \log L_k(\lambda) &= \log [\lambda I + L_k(\lambda) - \lambda I] = \\ &= I \log \lambda + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n \lambda^n} [L_k(\lambda) - \lambda I]^n = \\ &= I \log \lambda + \sum_{n=1}^{k-1} \frac{(-1)^{n+1}}{n \lambda^n} [L_k(\lambda) - \lambda I]^n \end{aligned}$$

существует и в действительности является логарифмом в том смысле, что $e^B = A$.

Упражнения

1. Для невырожденной матрицы существует корень k -й степени, где $k = 2, 3, \dots$

2. Решение уравнения $\frac{dX}{dt} = Q(t)X$, $X'(0) = I$, можно искать в виде

$$e^P e^{P_1} \dots e^{P_n} \dots,$$

где

$$P = \int_0^t Q(s) ds, \quad P_n = \int_0^t Q_n ds$$

и

$$Q_n = e^{-P_{n-1}} Q_{n-1} e^{P_{n-1}} + \int_0^{-1} e^{s P_{n-1}} Q_{n-1} e^{-s P_{n-1}} ds,$$

$$Q_0 = Q.$$

Бесконечное произведение сходится, если t достаточно мало; см Фер (F. Fer), Bull. classe sci. Acad. roy. Belg. 44, № 5 (1958), 818—829.

§ 21. Другое доказательство. Поскольку нами не доказана возможность представления матрицы в жордановой форме, дадим другое доказательство теоремы 10, воспользовавшись методом индукции. Предположим, что представление $A = e^B$ справедливо для матрицы N -го порядка и что мы привели матрицу $N+1$ -го порядка к треугольной форме

$$A_{N+1} = \begin{vmatrix} A_N & a_N \\ 0 & \lambda_{N+1} \end{vmatrix}, \quad (1)$$

где A_N — матрица N -го порядка, приведенная к треугольной форме, a_N — N -мерный вектор.

Пусть B_N — логарифм матрицы A_N , являющейся, очевидно, невырожденной, поскольку матрица A_{N+1} невырождена. Запишем

$$B_{N+1} = \begin{vmatrix} B_N & x \\ 0 & l \end{vmatrix}, \quad (2)$$

где $l = \log \lambda_{N+1}$, а x — неизвестный N -мерный вектор-столбец. Остается показать, что x может быть выбран таким, что $e^{B_{N+1}} = A_{N+1}$.

Легко по индукции показать, что для $k=1, 2, \dots$ имеем

$$B_{N+1}^k = \begin{vmatrix} B_N^k (B_N^{k-1} + l B_N^{k-2} + \dots + l^{k-1} I) & x \\ 0 & l^k \end{vmatrix}. \quad (3)$$

Следовательно,

$$e^{B_{N+1}} = \begin{vmatrix} e^{B_N} \sum_{k=0}^{\infty} \frac{1}{k!} (B_N^{k-1} + B_N^{k-2} l + \dots + l^{k-1} I) & x \\ 0 & \lambda_{N+1} \end{vmatrix}, \quad (4)$$

где два первых члена ряда взяты равными 0 и 1.

Если l не является собственным значением B_N , то

$$\begin{aligned} C(l) &= \sum_{k=0}^{\infty} \frac{1}{k!} (B_N^{k-1} + B_N^{k-2} l + \dots + l^{k-1} I) = \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} (B_N^k - l^k I) (B_N - lI)^{-1} = (e^{B_N} - e^{lI}) (B_N - lI)^{-1} \end{aligned} \quad (5)$$

и, следовательно,

$$|C(l)| = \frac{|e^{B_N} - e^{lI}|}{|B_N - lI|} = \prod_{k=1}^N \frac{e^{r_k} - e^l}{r_k - l}, \quad (6)$$

где r_1, r_2, \dots, r_N — собственные значения B_N . Поскольку, как видно из (5), $|C(l)|$ является для всех l непрерывной функцией l ,

то (6) справедливо для всех l . Если l не равно r_k или любому другому значению логарифма от λ_k , то ясно, что $|C(l)| \neq 0$. Если $l = r_k$, то множитель $(e^{l k} - e^l)/(r_k - l)$ становится равным $e^{l k}$. Поэтому, ограничиваясь главным значением логарифма, мы видим, что матрица $C(l)$ всегда является невырожденной. Следовательно, вектор x может быть всегда определен так, чтобы $C(l)x = a_N$.

Это завершает доказательство.

§ 22. Некоторые интересные преобразования. Приведение к канонической форме дифференциальных уравнений с переменной матрицей

$$\frac{dx}{dt} = A(t)x, \quad x(0) = c, \quad (1)$$

является задачей весьма трудной. Поскольку эта проблема более тесно связана с теорией дифференциальных уравнений, чем с теорией матриц, мы не будем здесь ею заниматься в деталях.

Некоторые интересные преобразования встречаются, однако, в самом начале исследования. Положим $x = Ty$, где T — функция t . Тогда y удовлетворяет уравнению

$$\frac{dy}{dt} = \left(T^{-1}AT - T^{-1}\frac{dT}{dt} \right) y. \quad (2)$$

Положим $f(A, T) = T^{-1}AT - T^{-1}\frac{dT}{dt}$. Выполняя дифференцирование, получим, что $f(A, T)$ удовлетворяет функциональному уравнению

$$f(A, TS) = f\left(T^{-1}AT - T^{-1}\frac{dT}{dt}, S \right). \quad (3)$$

Упражнение

1. Рассмотреть скалярное уравнение

$$\frac{d^n u}{dt^n} + a_1(t) \frac{d^{n-1}u}{dt^{n-1}} + \dots + a_n(t) u = 0.$$

Вводя преобразование $u = sv$, $t = \varphi(s)$, получить аналогичный класс функциональных уравнений.

§ 23. Биортогональность. В первой части книги при рассмотрении симметрических матриц мы установили, что произвольный N -мерный вектор может быть записан как линейная комбинация N ортогональных собственных векторов матрицы A , соответствующих N собственным значениям этой матрицы.

Таким образом, если

$$x = \sum_{i=1}^N a_i x^i, \quad (1)$$

то коэффициенты a_i определяются весьма просто по формуле

$$a_i = (x, x^i). \quad (2)$$

Если матрица A не является симметрической, то мы встречаемся с двумя трудностями при получении результата, аналогичного этому. Во-первых, матрица A может не иметь полного набора собственных векторов и, во-вторых, они не обязаны быть взаимно ортогональными.

Чтобы обойти первую трудность, мы просто сделаем предположение, что рассматриваемая матрица имеет N линейно независимых собственных векторов. Вторая трудность преодолевается введением в рассмотрение сопряженных матриц. При этом понятие ортогональности заменяется понятием биортогональности. Пусть y^1, y^2, \dots, y^N означают линейно независимые собственные векторы, соответствующие характеристическим числам $\lambda_1, \lambda_2, \dots, \lambda_N$ матрицы A . Тогда матрица

$$T = \|y^1, y^2, \dots, y^N\| \quad (3)$$

является невырожденной и преобразует A к диагональной форме

$$T^{-1}AT = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \\ 0 & & & & \end{vmatrix}. \quad (4)$$

Отсюда следует, что

$$T^* A^* (T^*)^{-1} = \begin{vmatrix} \bar{\lambda}_1 & & & 0 \\ & \bar{\lambda}_2 & & \\ & & \ddots & \\ & & & \bar{\lambda}_N \\ 0 & & & & \end{vmatrix}. \quad (5)$$

Следовательно, A^* также обладает N линейно независимыми собственными векторами, равными столбцам матрицы $(T^*)^{-1}$ и соответствующими характеристическим числам $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_N$. Обозначим эти векторы соответственно через z^1, z^2, \dots, z^N . Для определения коэффициентов разложения

$$x = \sum_{i=1}^N a_i y^i \quad (6)$$

поступим следующим образом. Из равенств

$$\begin{aligned} Ay' &= \lambda_i y', \\ A^* z' &= \bar{\lambda}_j z' \end{aligned} \quad (7)$$

получаем следующие соотношения:

$$(Ay', \bar{z}^j) = (\lambda_i y', \bar{z}^j) \quad (8)$$

и далее

$$(y', \overline{A^* z^j}) = (y', \lambda_j \bar{z}^j) = (\lambda_i y', \bar{z}^j). \quad (9)$$

Следовательно,

$$(\lambda_i - \lambda_j)(y', \bar{z}^j) = 0. \quad (10)$$

Поэтому при $\lambda_i \neq \lambda_j$ мы имеем

$$(y', \bar{z}^j) = 0.$$

Очевидно, что при этом $(y', \bar{z}^i) \neq 0$, поскольку вектор \bar{z}^i , будучи ненулевым, не может быть ортогонален ко всем y' .

Таким образом, в частном случае, когда все λ_i различны, мы можем записать для коэффициентов разложения (6) формулу

$$a_i = \frac{(x, \bar{z}^i)}{(y', \bar{z}^i)}. \quad (11)$$

Два набора векторов $\{y^i\}$ и $\{z^i\}$, удовлетворяющих равенству

$$(y^i, \bar{z}^j) = 0 \quad \text{при } i \neq j, \quad (12)$$

называются *биортогональными*. Очевидно, что мы можем нормировать векторы так, чтобы получить дополнительное условие

$$(y^i, \bar{z}^i) = 1. \quad (13)$$

Этот метод представления x как линейной комбинации y' подтверждает сделанное ранее замечание о том, что свойства матрицы проще всего изучить, рассматривая наряду с ней транспонированную.

Упражнения

1. Исследовать случай кратных характеристических чисел.
2. Какие упрощения следуют из допущения нормальности A ?

§ 24. Преобразование Лапласа. Поскольку нам известно, что решение линейного дифференциального уравнения с постоянными коэффициентами имеет вид

$$x = \sum_{k=1}^N e^{\lambda_k t} p_k(t), \quad (1)$$

где $p_k(t)$ — вектор, компоненты которого являются полиномами от t степени не выше $N-1$, то мы можем применить к решению уравнения преобразование Лапласа.

Интеграл

$$g(s) = \int_0^{\infty} e^{-st} f(t) dt \quad (2)$$

в случае его сходимости называется *преобразованием Лапласа* от функции $f(t)$. Поскольку мы интересуемся только функциями $f(t)$ вида (1), очевидно, что $g(s)$ всегда существует, если вещественная часть s достаточно велика.

Основная формула имеет вид

$$\int_0^{\infty} e^{-st} e^{at} dt = \frac{1}{s-a} \quad (3)$$

при $\operatorname{Re} s > \operatorname{Re} a$. Очевидно, если $f(t)$ имеет вид

$$f(t) = c_1 e^{\lambda_1 t}, \quad (4)$$

то равенство

$$\int_0^{\infty} e^{-st} f(t) dt = \frac{c_1}{s-a} \quad (5)$$

влечет $c_1 = c_1$ и $\lambda_1 = a$.

Подобно этому, из аддитивности интеграла (2) легко заключить, что если $f(t)$ принадлежит классу функций (1), а $g(s)$ имеет вид

$$g(s) = \frac{c_1}{s-a_1} + \frac{c_2}{s-a_2}, \quad (6)$$

то

$$f(t) = c_1 e^{a_1 t} + c_2 e^{a_2 t}.$$

§ 25. Пример. Рассмотрим линейное дифференциальное уравнение

$$u'' - 3u' + 2u = 0 \quad (1)$$

и будем искать решение, удовлетворяющее условиям

$$u(0) = 1, \quad u'(0) = 0.$$

Поскольку характеристическое уравнение

$$\lambda^2 - 3\lambda + 2 = 0 \quad (2)$$

имеет корни $\lambda_1 = 1$ и $\lambda_2 = 2$, то решение (1) должно иметь вид

$$u = c_1 e^{t} + c_2 e^{2t}. \quad (3)$$

Постоянные c_1 и c_2 определяются из условий

$$\begin{aligned} c_1 + c_2 &= 1, \\ 2c_1 + c_2 &= 0, \end{aligned} \quad (4)$$

что дает

$$c_1 = -1, \quad c_2 = 2. \quad (5)$$

Окончательно решение записывается в форме

$$u = 2e^t - e^{2t}. \quad (6)$$

Воспользуемся вместо этой процедуры преобразованием Лапласа. Из (1) для $\operatorname{Re}(s) > 2$ получим

$$\int_0^{\infty} u' e^{-st} dt - 3 \int_0^{\infty} u' e^{-st} dt + 2 \int_0^{\infty} u e^{-st} dt = 0. \quad (7)$$

Интегрируя два первых члена по частям, получим

$$\begin{aligned} \int_0^{\infty} u' e^{-st} dt &= u e^{-st} \Big|_0^{\infty} + s \int_0^{\infty} u e^{-st} dt = -1 + s \int_0^{\infty} u e^{-st} dt, \\ \int_0^{\infty} u' e^{-st} dt &= u' e^{-st} \Big|_0^{\infty} + s \int_0^{\infty} u' e^{-st} dt = \\ &= 0 + s \int_0^{\infty} u' e^{-st} dt = -s + s^2 \int_0^{\infty} u e^{-st} dt. \end{aligned} \quad (8)$$

Таким образом, (7) преобразуется к виду

$$(s^2 - 3s + 2) \int_0^{\infty} u e^{-st} dt = s - 3, \quad (9)$$

или

$$\int_0^{\infty} u e^{-st} dt = \frac{s-3}{s^2-3s+2}. \quad (10)$$

Рациональная функция в правой части может быть разложена на простейшие дроби:

$$\frac{s-3}{s^2-3s+2} = \frac{a_1}{s-1} + \frac{a_2}{s-2}, \quad (11)$$

где

$$\begin{aligned} a_1 &= \lim_{s \rightarrow 1} \frac{(s-1)(s-3)}{s^2-3s+2} = \frac{s-3}{s-2} \Big|_{s=1} = 2, \\ a_2 &= \lim_{s \rightarrow 2} \frac{(s-2)(s-3)}{s^2-3s+2} = \frac{s-3}{s-1} \Big|_{s=2} = -1. \end{aligned} \quad (12)$$

Следовательно,

$$\int_0^{\infty} u e^{-st} dt = \frac{2}{s-1} - \frac{1}{s-2},$$

откуда получаем соотношение (6) для $u(t)$.

§ 26. Обсуждение результата. Если вместо уравнения второго порядка, рассмотренного в § 25, нам необходимо исследовать уравнение десятого порядка, то обычный метод определения постоянных интегрирования потребует решения системы десяти совместных уравнений. Это труднопреодолимая задача. Использование преобразования Лапласа устраняет эти трудности. Это лишь малая часть тех преимуществ, которые присущи методу Лапласа и которые приводят к тому, что преобразование Лапласа играет фундаментальную роль в анализе.

Упражнения

1. Показать, что

$$\int_0^{\infty} t^n e^{-st} e^{at} dt = \frac{n!}{(s-a)^{n+1}},$$

используя:

(а) интегрирование по частям,

(б) дифференцирование по s или по a .

2. Используя эту формулу, решить дифференциальное уравнение

$$u'' - 2u' + u = 0, \quad u(0) = 1, \quad u'(0) = 0,$$

характеристическое уравнение для которого имеет кратный корень.

3. Использовать преобразования Лапласа для решения системы линейных уравнений:

$$b_k = \sum_{i=1}^N e^{\lambda_i t} x_i, \quad k = 1, 2, \dots, N,$$

с неизвестными x_i . Для этого составить и решить соответствующее дифференциальное уравнение.

4. Как подойти к решению соответствующей задачи, если неизвестными являются x_i и N ?

§ 27. Матричный случай. Рассмотрим уравнение

$$\frac{dx}{dt} = Ax, \quad x(0) = c, \quad (1)$$

где A — постоянная матрица. Производя преобразование Лапласа, получим

$$\int_0^{\infty} \frac{dx}{dt} e^{-st} dt = A \int_0^{\infty} x e^{-st} dt. \quad (2)$$

Интегрируя по частям, найдем

$$(A - sI) \int_0^{\infty} x e^{-st} dt = -c. \quad (3)$$

Таким образом, преобразование Лапласа от решения имеет вид

$$\int_0^{\infty} x e^{-st} dt = (Is - A)^{-1} c. \quad (4)$$

Далее мы можем использовать тот же путь, что и ранее, а именно разложить правую часть на простейшие дроби. В результате получим

$$(Is - A)^{-1} = \sum_{k=1}^N A_k (s - \lambda_k)^{-1}. \quad (5)$$

Для простоты, как и в одномерном случае, предположим, что собственные значения A различны. Тогда

$$A_k = \lim_{s \rightarrow \lambda_k} (s - \lambda_k) (Is - A)^{-1}. \quad (6)$$

Упражнения

1. Найти точное выражение для A_k , используя результат упражнения 34 к гл. 6.
2. Что произойдет, если матрица A имеет кратные корни?

Упражнения к гл. 11

1. При каких условиях справедливо разложение

$$A^{-1} = I - B + B^2 - \dots,$$

если $A = I + B$?

2. Рассмотрев решение уравнения $x'' + Ax = 0$, где A — положительно определенная матрица, показать, что характеристические числа положительно определенной матрицы положительны.

3. Пусть комплексная матрица A является произведением положительно определенной эрмитовой матрицы H и унитарной матрицы U , $A = HU$. Показать, что A нормальна тогда и только тогда, когда H и U перестановочны.

4. Доказать, что если A , B и AB нормальны, то BA нормальна. (Вигман.)

5. Пусть для любой матрицы A H -матрица означает неотрицательно определенный квадратный корень из AA^* . Показать, что необходимым условием того, чтобы произведение двух нормальных матриц было нормальным, является перестановочность каждой из матриц с H -матрицей другой матрицы. (Вигман.)

6. $A^n = I$ для некоторого n в том и только в том случае, когда собственные значения матрицы A являются корнями из единицы и A приводится к диагональной форме.

7. Если $B^k = 0$ для некоторого k и $AB = BA$, то $|A + B| = |A|$.

8. Пусть P , Q , R и X — матрицы второго порядка. Тогда каждое из характеристических чисел матрицы X , являющейся решением уравнения $PX^2 + QX + R = 0$, является одновременно корнем уравнения $|P\lambda^2 + Q\lambda + R| = 0$. (Силь вестр.)

Следовательно,

$$\int_0^{\infty} u e^{-st} dt = \frac{2}{s-1} - \frac{1}{s-2},$$

откуда получаем соотношение (6) для $u(t)$.

§ 26. Обсуждение результата. Если вместо уравнения второго порядка, рассмотренного в § 25, нам необходимо исследовать уравнение десятого порядка, то обычный метод определения постоянных интегрирования потребует решения системы десяти совместных уравнений. Это труднопреодолимая задача. Использование преобразования Лапласа устраняет эти трудности. Это лишь малая часть тех преимуществ, которые присущи методу Лапласа и которые приводят к тому, что преобразование Лапласа играет фундаментальную роль в анализе.

Упражнения

1. Показать, что

$$\int_0^{\infty} t^n e^{-st} e^{at} dt = \frac{n!}{(s-a)^{n+1}},$$

используя:

(а) интегрирование по частям,

(б) дифференцирование по s или по a .

2. Используя эту формулу, решить дифференциальное уравнение

$$u'' - 2u' + u = 0, \quad u(0) = 1, \quad u'(0) = 0,$$

характеристическое уравнение для которого имеет кратный корень.

3. Использовать преобразования Лапласа для решения системы линейных уравнений:

$$b_k = \sum_{i=1}^{N_i} e^{\lambda_i k} x_i, \quad k = 1, 2, \dots, N,$$

с неизвестными x_i . Для этого составить и решить соответствующее дифференциальное уравнение.

4. Как подойти к решению соответствующей задачи, если неизвестными являются x_i и N ?

§ 27. Матричный случай. Рассмотрим уравнение

$$\frac{dx}{dt} = Ax, \quad x(0) = c, \quad (1)$$

где A — постоянная матрица. Производя преобразование Лапласа, получим

$$\int_0^{\infty} \frac{dx}{dt} e^{-st} dt = A \int_0^{\infty} x e^{-st} dt. \quad (2)$$

Интегрируя по частям, найдем

$$(A - sI) \int_0^{\infty} x e^{-st} dt = -c. \quad (3)$$

Таким образом, преобразование Лапласа от решения имеет вид

$$\int_0^{\infty} x e^{-st} dt = (Is - A)^{-1} c. \quad (4)$$

Далее мы можем использовать тот же путь, что и ранее, а именно разложить правую часть на простейшие дроби. В результате получим

$$(Is - A)^{-1} = \sum_{k=1}^N A_k (s - \lambda_k)^{-1}. \quad (5)$$

Для простоты, как и в одномерном случае, предположим, что собственные значения A различны. Тогда

$$A_k = \lim_{s \rightarrow \lambda_k} (s - \lambda_k) (Is - A)^{-1}. \quad (6)$$

Упражнения

1. Найти точное выражение для A_k , используя результат упражнения 34 к гл. 6.
2. Что произойдет, если матрица A имеет кратные корни?

Упражнения к гл. 11

1. При каких условиях справедливо разложение

$$A^{-1} = I - B + B^2 - \dots,$$

если $A = I + B$?

2. Рассмотрев решение уравнения $x' + Ax = 0$, где A — положительно определенная матрица, показать, что характеристические числа положительно определенной матрицы положительны.

3. Пусть комплексная матрица A является произведением положительно определенной эрмитовой матрицы H и унитарной матрицы U , $A = HU$. Показать, что A нормальна тогда и только тогда, когда H и U перестановочны.

4. Доказать, что если A , B и AB нормальны, то BA нормальна. (Вигман.)

5. Пусть для любой матрицы A H -матрица означает неотрицательно определенный квадратный корень из AA^* . Показать, что необходимым условием того, чтобы произведение двух нормальных матриц было нормальным, является перестановочность каждой из матриц с H -матрицей другой матрицы. (Вигман.)

6. $A^n = I$ для некоторого n в том и только в том случае, когда собственные значения матрицы A являются корнями из единицы и A приводится к диагональной форме.

7. Если $B^k = 0$ для некоторого k и $AB = BA$, то $|A + B| = |A|$.

8. Пусть P , Q , R и X — матрицы второго порядка. Тогда каждое из характеристических чисел матрицы X , являющейся решением уравнения $PX^2 + QX + R = 0$, является одновременно корнем уравнения $|P\lambda^2 + Q\lambda + R| = 0$. (Силь вестр.)

17. Используя этот результат, показать, что

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0,$$

где $[A, B]$ — скобки Якоби: $[A, B] = AB - BA$.

18. Необходимое и достаточное условие того, что $\lim_{n \rightarrow \infty} B^n = 0$ для вещественной или комплексной матрицы B , сводится к требованию существования такой положительно определенной матрицы H , что матрица $H - B^*HB$ положительно определена. (Стейн.)

19. Матрица A приводится к диагональной форме в том и только в том случае, если существует положительно определенная эрмитова матрица H такая, что HAH^{-1} нормальна. (Митчелл.)

20. Существует ортогональная матрица $B(t)$ такая, что $y = B(t)z$ преобразует уравнение $dy/dt = A(t)y$ в $dz/dt = A_1(t)z$, где $A_1(t)$ — треугольная матрица. (Дилиберто^{*)}.)

21. Может быть найдена такая ограниченная и невырожденная матрица $B(t)$, что $A_1(t)$ — диагональная матрица. (Дилиберто.)

22. Матрица A имеет форму eS , где S — кососимметрическая матрица, тогда и только тогда, когда A — вещественная ортогональная матрица с определителем, равным единице. (Табер.)

23. Каждая невырожденная матрица может быть выражена как произведение:

(а) симметрической и ортогональной матриц, если она вещественна,

(б) эрмитовой и унитарной матриц, если она комплексна.

Если первый сомножитель выбран положительно определенным, то оба сомножителя определяются единственным образом. Эти и многие другие результаты установлены в работе Де Брёйна и Секереша¹⁾ с использованием логарифмов от матриц.

24. Вывести результат упражнения 22 из канонического представления для ортогональных матриц, данного выше.

25. Пусть u_1, u_2, \dots, u_N — набор линейно независимых решений дифференциального уравнения порядка N :

$$u^{(N)} + p_1(t) u^{(N-1)} + \dots + p_N(t) u = 0.$$

Определитель

$$W(u_1, u_2, \dots, u_N) = \begin{vmatrix} u_1 & u_2 & \dots & u_N \\ u_1' & u_2' & \dots & u_N' \\ \vdots & \vdots & \ddots & \vdots \\ u_1^{(N-1)} & u_2^{(N-1)} & \dots & u_N^{(N-1)} \end{vmatrix}$$

называется *определителем Вронского* функций $u_1(t), u_2(t), \dots, u_N(t)$. Показать, что

$$W(t) = W(u_1, u_2, \dots, u_N) = W(t_0) e^{-\int_{t_0}^t p_1(s) ds}.$$

^{*}) Это теорема Перрона. См. Math. Z. 32 (1930), см. также Р. Э. Винograd, УМН 9, № 2 (60) (1954), 129—136. (Прим. ред.)

¹⁾ N. G. De Bruijn and G. Szekeres, Nieuw. Arch. Wisk (3), 111 (1955), 20—32.

Обобщение понятия определителя Вронского содержится в статье Островского¹⁾.

Соответствующий определитель, связанный с решением линейного разностного уравнения вида

$$u(x+N) + p_1(x)u(x+N-1) + \dots + p_N(x)u(x) = 0, \quad x=0, 1, 2, \dots,$$

$$C(u_1, \dots, u_N) = \begin{vmatrix} u_1(x) & u_2(x) & \dots & u_N(x) \\ u_1(x+1) & u_2(x+1) & \dots & u_N(x+1) \\ \vdots & \vdots & \ddots & \vdots \\ u_1(x+N-1) & u_2(x+N-1) & \dots & u_N(x+N-1) \end{vmatrix}.$$

называется *определителем Касорати*²⁾.

Многие дальнейшие результаты, связанные с определителем Вронского, могут быть найдены в книге Поля и Сеге³⁾.

26. Унитарная матрица U может быть выражена в форме $U = V^{-1}W^{-1}VW$, где V и W унитарны, тогда и только тогда, когда $|U| = 1$.

27. Унитарные матрицы U и V могут быть выражены в форме $U = W_1W_2W_3$, $V = W_3W_2W_1$, где W_1, W_2, W_3 унитарны, тогда и только тогда, когда $|U| = |V|$.

28. Если все элементы матрицы $AA^* - A^*A$ неотрицательны, то A нормальна.

29. Если матрицы H_1 и H_2 эрмитовы и хотя бы одна из них является положительно определенной, то матрица H_1H_2 может быть приведена к диагональной форме. (Таусски.)

30. Если $\lambda A + \mu B$ может быть приведена к диагональной форме для произвольных скаляров λ и μ , то $AB = BA$. (Мощкин и Таусски.)

31. Показать, что матрица A , подобная вещественной диагональной матрице D в том смысле, что $A = TDT^{-1}$, является произведением двух эрмитовых матриц. (Таусски.)

32. Характеристические числа матрицы

$$A = \begin{vmatrix} a & b & . & . & . & . \\ c & a & b & . & . & . \\ \dots & c & a & b & . & . \\ . & . & . & . & c & a \end{vmatrix}$$

равны $\lambda_k = a - 2\sqrt{bc} \cos k\theta$, $k=1, 2, \dots, N$, где $\theta = \frac{\pi}{N+1}$. Показать, что полином $\Delta_N(\lambda) = |A - \lambda I|$ удовлетворяет разностному уравнению

$$\Delta_N(\lambda) = (a - \lambda)\Delta_{N-1}(\lambda) - bc\Delta_{N-2}(\lambda).$$

¹⁾ A. Ostrowski, Über ein Analogon der Wronskischen Determinante bei Funktionen mehrerer veränderlicher, Math. Z. 4 (1919), 223—230.

²⁾ Монте́ль (P. Montel), Leçons sur les recurrences et leurs applications, Gauthier-Villars, Paris, 1957; Кре́билл (D. M. Krabill), On Extension of Wronskian Matrices, Bull. Amer. Math. Soc. 49 (1943), 593—601.

³⁾ G. Polya and G. Szegő, Aufgaben und Lehrsätze aus der Analysis, Zweiter Band, Dover Publications, New York, 1945. [Русский перевод: Задачи и теоремы из анализа, в двух томах, Гостехиздат, 1956.]

См. Като и Таусски¹⁾.

36. Если собственные значения матрицы A равны $\lambda_1, \lambda_2, \dots, \lambda_N$, c_1 и c_2 — два ненулевых комплексных числа, а собственные значения матрицы $c_1 A + c_2 A^*$ равны $c_1 \lambda_i + c_2 \bar{\lambda}_j$ (1), где $j(i)$ — перестановка из индексов $1, 2, \dots, N$, то A нормальна.

37. Пусть матрицы A и B перестановочны. Тогда характеристические числа матрицы $f(A, B)$ равны $f(\lambda_i, \mu_i)$; λ_i и μ_i — собственные значения A и B , расположенные в порядке, не зависящем от f . (Фробениус.)

(О парах матриц A и B , для которых собственные значения матрицы $c_1 A + c_2 B$ равны $c_1 \lambda^{(1)} + c_2 \lambda^{(2)}$, говорят²⁾, что они обладают свойством L .)

38. Доказать, что A нормальна в том и только в том случае, когда

$$\operatorname{Sp}(A^* A) = \sum_{i=1}^N |\lambda_i|^2. \quad (\text{Шур.})$$

39. Использовать это свойство для доказательства результата, установленного Вигманом, что BA нормальна, если A , B и AB нормальны. (Мирский.)

40. В случае, если матрица A эрмитова, показать, что для функции $s(A)$, введенной в упражнении 34, имеет место формула

$$s(A) = 2 \sup_{u, \varphi} |(\bar{u}, A \varphi)|,$$

где верхняя грань берется по всем парам взаимно ортогональных нормированных векторов u и φ . (Мирский.)

41. Если матрица A эрмитова, то $s(A) \geq 2 \max_{r \neq s} |a_{rs}|$. (Паркер, Мирский.)

42. Если A нормальна, то

$$s(A) \leq \sup_{|z|=1} s\left(\frac{zA + \bar{z}A^*}{2}\right). \quad (\text{Мирский}^3))$$

43. Если A и B — нормальные матрицы с характеристическими числами λ_i и μ_i , то существует упорядочение этих чисел такое, что

$$\sum_i |\lambda_i - \mu_i|^2 \leq \|A - B\|^2, \quad \text{где} \quad \sum_{i,j} |a_{ij}|^2 = \|A\|^2.$$

(Гоффман, Виландт.)

44. Как известно, любая матрица A может быть записана в виде $B + iC$, где матрицы B и C эрмитовы. Пусть характеристические числа B и C расположены соответственно между b_1 и b_2 и между c_1 и c_2 . Тогда все характеристические числа A расположены внутри квадрата с вершинами $b_1 + ic_1$, $b_1 + ic_2$, $b_2 + ic_1$, $b_2 + ic_2$. (Бендиксон, Гирш.)

45. Под областью D_A матрицы A будем понимать комплексную область, ограниченную значениями, принимаемыми квадратичной формой (x, Ax) при $(x, x) = 1$. Показать, что характеристические числа принадлежат D_A ⁴⁾.

46. Пусть K — наименьшая выпуклая комплексная область, включающая все характеристические числа матрицы A . Если A нормальна, то K и D_A совпадают.

¹⁾ T. Kato and O. Tausski, Commutators of A and A^* , J. Washington Acad. Sci. 46 (1956), 38—40.

²⁾ Мотзкин и Таусски (T. S. Motzkin and O. Tausski), Pairs of Matrices with Property L , Trans. Amer. Math. Soc. 73 (1952), 108—114.

³⁾ L. Mirsky, Inequalities for Normal and Hermitian Matrices, Duke Math. J. 24 (1957), 592—600.

⁴⁾ Гёплиц (O. Toeplitz), Das algebraische Analogon zu einem Satze von Fejer, Math. Z. 2 (1919), 187—197.

47. Рассмотрев случай матриц размеров 2×2 , показать, что этот результат не всегда справедлив, если A не является нормальной матрицей.

48. Область D_A является выпуклой независимо от того, нормальна или нет матрица A . Библиографию, касающуюся последних результатов в этой области, см. у Таусски¹⁾.

49. Уравнению $x' = Ax$, где x — N -мерный вектор, а A — матрица размеров $N \times N$, соответствует линейное однородное дифференциальное уравнение N -го порядка, которому удовлетворяет каждая компонента x . Пусть каждое решение дифференциального уравнения, соответствующего $y' = By$, совпадает с решением уравнения, соответствующего $x' = Ax$. Что можно сказать о связи A и B ?

50. Рассмотрим уравнение $\frac{dx}{dt} = Bx$, $x(0) = c$, где постоянная матрица B выбрана так, что функция (x, x) не зависит от времени (при любом c). Показать, что матрица B обладает этим свойством в том и только в том случае, когда она косимметрическая. Заключить из этого, что e^{Bt} ортогональна.

51. Используя тот факт, что векторное дифференциальное уравнение $A \frac{d^2x}{dt^2} + B \frac{dx}{dt} + Cx = 0$ эквивалентно системе $\frac{dx}{dt} = y$, $A \frac{dy}{dt} + By + Cx = 0$, получить детерминантное равенство, эквивалентное $|\lambda I^2 + B\lambda + C| = 0$, с элементами, линейными по λ .

52. Пусть x_1, x_2, \dots, x_m — скаляры и через $F(x_1, x_2, \dots, x_m)$ обозначена матрица $A_1x_1 + A_2x_2 + \dots + A_mx_m$, где A_i попарно перестановочны. Пусть $f(x_1, x_2, \dots, x_m) = |F(x_1, x_2, \dots, x_m)|$; тогда $f(X_1, X_2, \dots, X_m) = 0$ при условии, что $A_1X_1 + A_2X_2 + \dots + A_mX_m = 0$. (Обобщение теоремы Гамильтона — Кэли, данное Г. Филлипсом.)

53. Показать, что этот результат сохраняет силу, если линейная форма $\sum A_i x_i$ заменяется полиномом от x_i с матричными коэффициентами. См. Ко и Ли²⁾, где имеются также ссылки на более ранние работы Островского и Филлипса.

54. Если E_i ($i = 1, 2, \dots, n$) — матрицы размеров 4×4 и при этом $E_i^2 = -I$, $E_i E_j = -E_j E_i$ ($i \neq j$), то $n \leq 5$. (Эддингтон.)

55. Если матрицы E_i являются либо вещественными, либо чисто мнимыми (см. предыдущее упражнение), то из пяти матриц две вещественны, а три мнимы. Обобщение этого случая см. у Ньюмена³⁾.

56. Если B перестановочна с любой из матриц, с которой перестановочна A , то B — скалярный полином от A . (Веддерберн⁴⁾.)

57. Пусть комплексная матрица A имеет спектр $\lambda_1, \lambda_2, \dots, \lambda_N$. Положим $\|X\|^2 = \sum_{i,j} |x_{ij}|^2$. Тогда

$$\inf_s \|S^{-1}AS\|^2 = \sum_{i=1}^N |\lambda_i|^2,$$

где нижняя грань определяется по всем невырожденным матрицам S . Нижняя грань достигается в том и только в том случае, когда матрица A приводима к диагональной форме. (Мирский.)

¹⁾ O. Tausk, Bibliography on Bounds for Characteristic Roots of Finite Matrices, National Bureau of Standards, September 1951.

²⁾ Chao Ko and H. C. Lee, A Further Generalization of the Hamilton — Cayley Theorem, J. London Math. Soc. 15 (1940), 153—158.

³⁾ M. H. A. Newman, J. London Math. Soc. 7 (1932), 93—99.

⁴⁾ J. H. M. Wedderburn, Lectures on Matrices, Amer. Math. Soc. Colloq. Publ. 17 (1934), 106.

58. Пусть A , B и X означают матрицы размеров $N \times N$. Показать, что для существования по крайней мере одного решения матричного уравнения $X^2 - 2AX + B = 0$ достаточно, чтобы характеристические числа матрицы

$$R = \begin{bmatrix} A & I \\ A^2 - B & A \end{bmatrix}$$

были различны. (Тройенфельс.)

59. Пусть $\{X\}$ — набор матриц размеров $N \times N$, обладающих тем свойством, что каждая их вещественная линейная комбинация имеет только вещественные собственные значения. Тогда наибольшее собственное значение $\lambda_{\max}(X)$ является выпуклой функцией X , а наименьшее собственное значение $\lambda_{\min}(X)$ является вогнутой функцией X .

60. Пусть собственные значения любой вещественной линейной комбинации двух матриц X и Y вещественны, причем матрица X имеет отрицательные собственные значения. Тогда любое характеристическое число комбинации $X + iY$ имеет отрицательную вещественную часть.

61. Пусть $\{X\}$ — набор вещественных матриц размеров $N \times N$, обладающих тем свойством, что каждая их вещественная линейная комбинация имеет только вещественные собственные значения. Тогда если X_1 и X_2 — две матрицы из этого набора, обладающие тем свойством, что матрица $X_1 - X_2$ имеет неотрицательные характеристические числа, то $\lambda_i(X_1) \geq \lambda_i(X_2)$, где числа $\lambda_i(X)$ перенумерованы в порядке их возрастания. (Лакс.) Доказательство этих положений, базирующееся на результатах теории дифференциальных уравнений в частных производных гиперболического типа, можно найти в статье Лакса¹⁾. Более элементарный путь доказательства см. у А. Вайнберга²⁾.

62. Даны два набора комплексных чисел $\{\alpha_i\}$, $\{\beta_i\}$ ($1 \leq i \leq n$), каждое из которых по модулю равно единице. Доказать следующее утверждение. Две унитарные матрицы A и B порядка n с характеристическими числами соответственно $\{\alpha_i\}$ и $\{\beta_i\}$ и такие, что единица является характеристическим числом матрицы AB , существуют в том и только в том случае, когда выпуклая оболочка $\{\alpha_i\}$ и выпуклая оболочка $\{\beta_i\}$ имеют общую точку. (Фань Цзы³⁾.)

63. Мы знаем, что e^{At} может быть записана в виде полинома от A , а именно $e^{At} = u_0(t)I + u_1(t)A + \dots + u_{N-1}(t)A^{N-1}$. Найти дифференциальные уравнения для $u_i(t)$, используя тот факт, что $\frac{d}{dt} e^{At} = Ae^{At}$.

64. Пусть A — матрица размеров $N \times N$ и $A(+)$ образована из A путем замены нулями всех элементов, расположенных на главной диагонали и ниже. Пусть $A(-) = A - A(+)$ и предположим, что $A(+)$ и $A(-)$ перестановочны. Запишем $e^A = PQ$, где $P - I$ имеет ненулевые члены только над диагональю, а $Q - I$ имеет ненулевые элементы под диагональю. Тогда с необходимостью

$$P = e^{A(+)}, \quad Q = e^{A(-)}$$

и, следовательно, факторизация единственна. См. Бакстер (G. Baxter), An Operator Identity, Pacific J. Math. 8 (1958), 649—664.

Аналогия между этой факторизацией и факторизацией Винера — Хопфа глубже, чем это кажется с первого взгляда.

¹⁾ P. D. Lax, Differential Equations, Difference Equations and Matrix Theory, Comm. Pure Appl. Math. 11 (1958), 175—194.

²⁾ A. F. Weinberger, Remarks on the Preceding Paper of Lax, Comm. Pure Appl. Math. 11 (1958), 195—196.

³⁾ Этот результат аналогичен теореме Виландта о собственных значениях суммы двух нормальных матриц. См. Pacific J. Math. 5 (1955), 633—638.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§§ 1—10. Результаты и методы заимствованы из работ

Беллман (R. Bellman), *Stability Theory of Differential Equations*, McGraw-Hill Book Company, Inc., New York, 1954 [русский перевод: Теория устойчивости дифференциальных уравнений, ИЛ, 1954];

Лефшец (S. Lefschetz), *Lectures on Differential Equations*, Ann. Math. Studies, no 14, 1946.

§§ 2—3. Обсуждение проблемы определения характеристического полинома см.

Хаусхолдер и Бауэр (A. S. Householder and F. L. Bauer), *On Certain Methods For Expanding the Characteristic Polynomial*, Numerische Mathematik 1 (1959), 29—40.

§ 6. Некоторые более поздние результаты и дополнительные ссылки см. в

Гирш (K. A. Hirsch), *A Note on Vandermonde's Determinant*, J. London Math. Soc. 24 (1949), 144—145.

§ 11. Доказательство канонического представления Жордана можно найти в книгах:

Лефшец (S. Lefschetz), *Differential Equations, Geometric Theory*, Interscience Publishers, Inc., New York, 1957, Appendix I;

Веддерберн (J. H. M. Wedderburn), *Lectures on Matrices*, Amer. Math. Soc. Colloq. Publ. 17 (1934). (См. также: И. М. Гельфанд, *Лекции по линейной алгебре*, «Наука», 1971; А. И. Мальцев, *Основы линейной алгебры*, «Наука», 1975.)

§ 13. Результат получен Шуром в работе

Шур (I. Schur), *Über die charakteristischen Wurzeln einer linearen Substitution mit einer Anwendung auf die Theorie der Integral Gleichungen*, Math. Ann. 66 (1909), 488—510.

§ 14. Понятие нормальной матрицы, по-видимому, впервые введено в работе

Теплиц (O. Toeplitz), *Math. Z.* 2 (1919), 190.

§ 15. Эта приближенная диагонализация введена Перроном и обычно используется в теоремах об устойчивости. См.

Перрон (O. Perron), *Math. Z.* 29 (1929), 129—160.

§ 16. Обсуждение теорем, касающихся результатов, см. в книге

Ван-дер-Варден (B. L. Van der Waerden), *Moderne Algebra*, Springer-Verlag, Berlin, 1931. [Русский перевод: Современная алгебра, Гостехиздат, 1947.]

См. также А. Г. Курош, *Курс высшей алгебры*, «Наука», 1975.

§ 17. Этот результат установлен Гамильтоном для кватернионов и Кэли для матриц. На стр. 18 книги Мак-Даффи поме-

шено обобщение теоремы Гамильтона—Кэли, данное Г. Филлипсом. Дальнейшие обобщения даны Ко и Ли (см. ссылку к упражнению 53 в настоящей главе).

§ 19. Этот вывод следует монографии Лефшеца, ссылка на которую приведена выше. Определение матрицы C является проблемой большой трудности. Хороший обзор результатов в этой области содержится в статье

В. М. Старжинский, Обзор работ об условиях устойчивости тривиального решения системы линейных дифференциальных уравнений с периодическими коэффициентами, «Прикладная математика и механика», 18, № 4 (1954), 469—510.

§ 20. Посмотрим, что означает этот результат. Любое невырожденное точечное преобразование $y = Ax$ может быть выделено из непрерывного семейства преобразований, порождаемых дифференциальным уравнением $\frac{dZ}{dt} = CZ$, $Z(0) = I$, где $A = e^C$.

§ 22. Некоторые задачи, связанные с этими преобразованиями, могут быть найдены в Research Problems в журнале Bull. Amer. Math. Soc., 1958.

Существует большое количество интересных представлений решения дифференциального уравнения второго порядка $(pu')' + qu = 0$. Одно из них, названное преобразованием Прюфера, было обобщено Барретом на матричный случай $(P(t)X')' + Q(t)X = 0$ (Proc. Amer. Math. Soc. 8 (1957), 510—518). Дальнейшие обобщения и усовершенствования даны в статье

Рейд (W. T. Reid), A Prufer Transformation for Differential Systems, Pacific J. Math. 8 (1958), 575—584.

Приложение к теории устойчивости см. в статье

Конти (R. Conti), Sulla t_∞ -Similitudine tra matrici e l'equivalenza asintotica dei sistemi differenziali lineari, Rev. Mat. Univ. Parma 8 (1957), 43—47.

§ 23. Детальную сводку результатов об использовании биортогонализации в связи с основной проблемой обращения матриц и определением собственных значений и векторов см. в статье

Хестинс (M. R. Hestenes), Inversion of Matrices by Biorthogonalization and Related Results, J. Soc. Ind. Appl. Math. 6 (1958), 51—90.

Одной из наиболее важных задач теории линейных дифференциальных уравнений является определение структуры решений линейных систем уравнений вида

$$\sum_{j=1}^N p_{ij}(D)x_j = 0, \quad x_i(0) = c_i \quad (i = 1, 2, \dots, N),$$

где $D = \frac{d}{dt}$ — оператор дифференцирования. Соответствующая проблема в теории матриц связана с определением характери-

стических чисел и собственных значений матриц $A(\lambda)$, элементы которых являются полиномами от λ .

Поскольку подробное обсуждение этой проблемы, которую мы лишь кратко затронули при рассмотрении матриц $A - \lambda I$, увело бы нас далеко от главного направления этой книги, мы отсылаем читателя к книге

Фрезер, Дункан и Коллар (R. A. Fraser, W. J. Duncan and A. R. Collar), *Elementary Matrices*, The Macmillan Company, New York, 1946 [русский перевод: Теория матриц и ее приложения к дифференциальным уравнениям и динамике, ИЛ, 1950],

где рассмотрены также многие прикладные задачи, например проблема флаттера.

Рассмотрение аналитических свойств этих матриц приводит нас к общей теории матриц, элементы которых являются полиномами от комплексной переменной. Эти вопросы важны не только сами по себе, но имеют также важные ответвления в современную физику и теорию прогнозирования. По этому поводу см.

Масани (P. Masani), *The Laurent Factorization of Operator-valued Functions*, Proc. London Math. Soc. 6 (1956), 59—69;

Масани и Винер (P. Masani and N. Wiener), *The Prediction Theory of Multivariate Stochastic Processes*, Acta Math 1 98 (1957), 110—150;

Хелсон и Лауденслагер (H. Helson and D. Lowdenslager), *Prediction Theory and Fourier Series in Several Variables*, Acta Math-99 (1958)

Результаты, относящиеся к структуре матриц, элементами которых являются кватернионы, могут быть найдены в статье

Вигман (N. A. Wiegman), *The Structure of Unitary and Orthogonal Quaternion Matrices*, III. J. Math. 2 (1958), 402—407,

где имеются также дальнейшие ссылки.

СИММЕТРИЧЕСКИЕ ФУНКЦИИ, КРОНЕКЕРОВСКИЕ ПРОИЗВЕДЕНИЯ И ЦИРКУЛЯНТЫ

§ 1. Введение. В этой главе мы покажем, как некоторые задачи, связанные с различными симметрическими функциями собственных значений, приводят к одному важному понятию в теории матриц. Используя предельный переход, мы получим кронекеровскую сумму — матричную функцию, играющую важную роль при рассмотрении матричных уравнений вида

$$AX + XB = C. \quad (1)$$

Нам уже приходилось встречаться с таким уравнением в одной из предшествующих глав, и мы еще раз встретимся с ним в гл. 13, посвященной теории устойчивости. Кроме того, кронекеровское произведение понадобится нам в одной из последующих глав, посвященной теории стохастических матриц.

После этого мы построим еще один важный класс блочных матриц, который возникает при рассмотрении некоторых кососимметрических функций. Хотя эти вопросы тесно связаны с геометрией N -мерного евклидова пространства, мы их здесь подробно не рассматриваем.

Наконец, мы введем циркулянты, по-прежнему исходя из соображений симметрии.

Во всей этой главе мы будем пользоваться одним и тем же подходом к определению собственных значений и собственных векторов матриц, именно матрица будет рассматриваться как эквивалент преобразования одного множества величин в другое. Эта точка зрения иллюстрируется и в ряде упражнений.

§ 2. Степени собственных значений. Очень простой класс симметрических функций от собственных значений представляют собой степенные суммы

$$p_k = \sum_{i=1}^N \lambda_i^k, \quad k = 1, 2, \dots \quad (1)$$

Поставим себе задачу найти сумму $\sum_{i=1}^N \lambda_i^k$. Поскольку

$$\sum_{i=1}^N \lambda_i^2 = \left(\sum_{i=1}^N \lambda_i \right)^2 - 2 \sum_{i>j} \lambda_i \lambda_j, \quad (2)$$

мы можем выразить сумму квадратов собственных значений через два коэффициента характеристического полинома. Более того, так как $\sum_{i=1}^N \lambda_i^k$, $k=1, 2, \dots$, является симметрической функцией собственных значений, то мы знаем заранее, что она может быть записана как полином от элементарных симметрических функций.

Во многих случаях, однако, такое представление не является самым удобным. Докажем следующую теорему:

Теорема 1. При $k=1, 2, \dots$ имеет место формула

$$\sum_{i=1}^N \lambda_i^k = \text{Sp} (A^k). \quad (3)$$

Доказательство. Если все собственные значения матрицы A различны, то ее диагональное представление

$$A = T \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix} T^{-1}, \quad (4)$$

из которого, кстати, следует представление A^k

$$A^k = T \begin{vmatrix} \lambda_1^k & & & 0 \\ & \lambda_2^k & & \\ & & \ddots & \\ 0 & & & \lambda_N^k \end{vmatrix} T^{-1}, \quad (5)$$

делает утверждение теоремы очевидным. Легко также видеть, что из знакомых уже нам соображений непрерывности формула (3) следует и для матриц общего вида.

Можно воспользоваться также приведением матрицы к треугольной форме

$$A = T \begin{vmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{vmatrix} T^{-1}, \quad (6)$$

где элементы, расположенные под диагональю, равны нулю, а λ_i не обязательно различны.

Легко видеть, что при этом

$$A^k = T \begin{vmatrix} \lambda_1^k & & & & \\ & \lambda_2^k & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & \lambda_N^k \end{vmatrix} T^{-1}. \quad (7)$$

Это представление дает другое доказательство формулы (3).

§ 3. Полиномы и характеристические уравнения. Мы знаем, что каждой матрице соответствует характеристический полином. Не ясно, однако, можно ли каждому полиному отнести матрицу, для которой этот полином будет характеристическим.

Теорема 2. Для матрицы

$$A = \begin{vmatrix} -a_1 & -a_2 & \dots & -a_{n-1} & -a_n \\ 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{vmatrix} \quad (1)$$

уравнение

$$(-1)^n |A - \lambda I| = \lambda^n + a_1 \lambda^{n-1} + \dots + a_n = 0 \quad (2)$$

является характеристическим.

Доказательство мы оставляем читателю в качестве упражнения.

Упражнение

1. Используя приведенный выше результат, определить сумму кубов корней уравнения $\lambda^n + a_1 \lambda^{n-1} + \dots + a_n = 0$.

§ 4. Симметрические функции. Хотя мы уже умеем находить матрицы собственными значениями которых являются степени $\lambda_1^k, \lambda_2^k, \dots, \lambda_N^k$, не известно, как строить матрицы, собственные значения которых равны заданным функциям вида $\lambda_i \lambda_j$, $i, j = 1, 2, \dots$

Рассмотрим более сложную задачу: найти матрицу, собственные значения которой равны $\lambda_i \mu_j$, где λ_i — собственные значения матрицы A , а μ_j — собственные значения матрицы B . Очевидно, что, положив $\lambda_i = \mu_i$, мы приходим к исходной задаче.

Рассмотрим наиболее простой случай квадратных матриц второго порядка. Имеем

$$\begin{aligned} \lambda_1 x_1 &= a_{11} x_1 + a_{12} x_2, & \mu_1 y_1 &= b_{11} y_1 + b_{12} y_2, \\ \lambda_1 x_2 &= a_{21} x_1 + a_{22} x_2, & \mu_1 y_2 &= b_{21} y_1 + b_{22} y_2. \end{aligned} \quad (1)$$

Отсюда следует, что

$$\begin{aligned} \lambda_{1\mu_1} x_1 y_1 &= a_{11} b_{11} x_1 y_1 + a_{11} b_{12} x_1 y_2 + a_{12} b_{11} x_2 y_1 + a_{12} b_{12} x_2 y_2, \\ \lambda_{1\mu_1} x_1 y_2 &= a_{11} b_{21} x_1 y_1 + a_{11} b_{22} x_1 y_2 + a_{12} b_{21} x_2 y_1 + a_{12} b_{22} x_2 y_2, \\ \lambda_{1\mu_1} x_2 y_1 &= a_{21} b_{11} x_1 y_1 + a_{21} b_{12} x_1 y_2 + a_{22} b_{11} x_2 y_1 + a_{22} b_{12} x_2 y_2, \\ \lambda_{1\mu_1} x_2 y_2 &= a_{21} b_{21} x_1 y_1 + a_{21} b_{22} x_1 y_2 + a_{22} b_{21} x_2 y_1 + a_{22} b_{22} x_2 y_2. \end{aligned} \quad (2)$$

Заметим, что величины $x_1 y_1$, $x_1 y_2$, $x_2 y_1$, $x_2 y_2$ фигурируют как в правой, так и в левой частях уравнений (2). Введем четырехмерный вектор

$$z = \begin{pmatrix} x_1 y_1 \\ x_1 y_2 \\ x_2 y_1 \\ x_2 y_2 \end{pmatrix} \quad (3)$$

и матрицу размеров 4×4

$$C = \begin{pmatrix} a_{11} b_{11} & a_{11} b_{12} & a_{12} b_{11} & a_{12} b_{12} \\ a_{11} b_{21} & a_{11} b_{22} & a_{12} b_{21} & a_{12} b_{22} \\ a_{21} b_{11} & a_{21} b_{12} & a_{22} b_{11} & a_{22} b_{12} \\ a_{21} b_{21} & a_{21} b_{22} & a_{22} b_{21} & a_{22} b_{22} \end{pmatrix}. \quad (4)$$

Тогда уравнение (2) можно записать в виде

$$\lambda_{1\mu_1} z = Cz. \quad (5)$$

Отсюда следует, что z является собственным вектором матрицы C , принадлежащим собственному значению $\lambda_{1\mu_1}$. Точно так же легко убедиться, что $\lambda_{1\mu_2}$, $\lambda_{2\mu_1}$, $\lambda_{2\mu_2}$ будут собственными значениями матрицы C , а соответствующие им собственные векторы вычисляются по формуле (3).

Таким образом, мы решили задачу нахождения матрицы, собственными значениями которой служат произведения $\lambda_{i\mu_j}$, $i, j = 1, 2$.

§ 5. Кронекеровские произведения. Поскольку даже в двумерном случае необходимы довольно громоздкие вычисления, попробуем ввести более удобные обозначения. Возвращаясь к уравнению (4.4), мы видим, что в структуре матрицы C имеется определенная регулярность. При ближайшем рассмотрении оказывается, что C можно записать в виде следующей блочной матрицы:

$$C = \begin{pmatrix} a_{11} B & a_{12} B \\ a_{21} B & a_{22} B \end{pmatrix}, \quad (1)$$

или, еще проще,

$$C = \|a_{ij} B\|. \quad (2)$$

Легко видеть, что такое представление не зависит от размеров матриц A и B .

Определение. Пусть A и B — квадратные матрицы соответственно порядков M и N . Квадратная матрица порядка MN , определенная по формуле (2), называется *кронекеровским произведением* матриц A и B и обозначается

$$A \times B = \| a_{ij} B \| \quad (3)$$

Из приведенных соображений сразу же следует

Теорема 3. *Собственными значениями матрицы $A \times B$ являются произведения $\lambda_i \mu_j$, где λ_i — собственные значения матрицы A , а μ_j — собственные значения матрицы B .*

Собственные векторы имеют вид

$$z_{ij} = \begin{pmatrix} x'_1 y'_j \\ x'_1 y'_j \\ \vdots \\ x'_M y'_j \end{pmatrix} \quad (4)$$

Здесь x'_k , $k = 1, 2, \dots, M$, обозначают компоненты собственного вектора x' матрицы A , тогда как y'_j является собственным вектором матрицы B .

Упражнения

1. Показать, что $\text{Sp}(A \times B) = \text{Sp } A \text{ Sp } B$.
2. Найти определитель $|A \times B|$.
3. Показать, что

$$I \times B = \begin{pmatrix} B & & 0 \\ & B & \\ 0 & & B \end{pmatrix}$$

§ 6. Алгебра кронекеровских произведений. Для того чтобы оправдать название «кронекеровское произведение» и введенные нами обозначения, мы должны показать, что $A \times B$ обладает рядом свойств, естественных для произведения.

Мы оставляем читателю в качестве упражнения доказательство следующих результатов:

$$A \times B \times C = (A \times B) \times C = A \times (B \times C), \quad (1a)$$

$$(A + B) \times (C + D) = A \times C + A \times D + B \times C + B \times D, \quad (1b)$$

$$(A \times B)(C \times D) = (AC) \times (BD). \quad (1c)$$

§ 7. Кронекеровские степени — I. Рассмотрим теперь кронекеровские степени, которые будем обозначать

$$A^{[1]} = A \times A, \quad A^{[k+1]} = A \times A^{[k]}. \quad (1)$$

Матрица $A_{[k]}$ определяется следующей таблицей размеров $(k+1) \times (k+1)$:

$$A_{[k]} = \begin{vmatrix} a_{11}^k & ka_{11}^{k-1}a_{12} & \dots \\ a_{11}^{k-1}a_{21} & (k-1)a_{11}^{k-2}a_{12}a_{21} + a_{11}^{k-1}a_{22} & \dots \\ \vdots & \vdots & \ddots \end{vmatrix}. \quad (2)$$

Это представление получено для квадратных матриц второго порядка. Представление для $A_{[k]}$ в случае матрицы порядка N можно получить подобным же образом.

Упражнения

1. Показать, что $(AB)_{[k]} = A_{[k]} B_{[k]}$.
2. Показать, что $A_{[k+l]} = A_{[k]} A_{[l]}$.

§ 10. Кронекеровский логарифм. Введем обобщенную степень матрицы для нецелого k . Для этого рассмотрим бесконечную последовательность, общий член которой имеет вид

$$(a_{11}x_1 + a_{12}x_2)^{k-i} (a_{21}x_1 + a_{22}x_2)^i, \quad i=0, 1, \dots,$$

где число k , вообще говоря, не целое.

В этом случае вместо конечной таблицы вида (9.2) мы получаем бесконечную таблицу, поскольку биномиальный ряд для $(a_{11}x_1 + a_{12}x_2)^{k-i}$ бесконечен.

Производя замену переменных

$$\begin{vmatrix} x_1 \\ x_2 \end{vmatrix} = B \begin{vmatrix} x'_1 \\ x'_2 \end{vmatrix}, \quad (1)$$

убеждаемся, что для произвольных значений k

$$(AB)_{[k]} = A_{[k]} B_{[k]}. \quad (2)$$

Разложим теперь каждый элемент матрицы $A_{[k]}$ в окрестности $k=0$. Имеем

$$A_{[k]} = A_{[0]} + kL(A) + \dots \quad (3)$$

Это соотношение определяет бесконечную матрицу $L(A)$, о которой можно говорить как о кронекеровском логарифме. Подставляя разложения (3) в (2), мы видим, что

$$L(AB) = A_{[0]}L(B) + L(A)B_{[0]}. \quad (4)$$

Это соотношение аналогично функциональному уравнению для скалярного логарифма. Заметим, что здесь $A_{[0]}$ не является единичной матрицей.

Упражнение

1. Определить l -й элемент матрицы $L(A)$ для $l=0, 1$ и $l=0, 1$.

§ 11. Кронекеровская сумма—I. Обратимся теперь к задаче определения матриц, собственные значения которых равны $\lambda_i + \mu_j$. При этом мы воспользуемся применявшимся методом, позволяющим получать аддитивные свойства из известных мультипликативных. Рассмотрим соотношение

$$(I_M + \varepsilon A) \times (I_N + \varepsilon B) = I_M \times I_N + \varepsilon (I_M \times B + A \times I_N) + \varepsilon^2 (A \times B), \quad (1)$$

где ε — числовой параметр.

Поскольку собственные значения матрицы $(I_M + \varepsilon A) \times (I_N + \varepsilon B)$ имеют вид

$$(1 + \varepsilon \lambda_i)(1 + \varepsilon \mu_j) = 1 + \varepsilon(\lambda_i + \mu_j) + \varepsilon^2 \lambda_i \mu_j,$$

то собственные значения матрицы $I_M \times B + A \times I_N$ должны быть равны $\lambda_i + \mu_j$.

Упражнение

1. Определить собственные векторы матрицы $I_M \times B + A \times I_N$.

§ 12. Кронекеровская сумма—II. Матрицу $I_M \times B + A \times I_N$ можно получить также с помощью дифференциальных уравнений. Рассмотрим уравнения

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2, & \frac{dy_1}{dt} &= b_{11}y_1 + b_{12}y_2, \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2, & \frac{dy_2}{dt} &= b_{21}y_1 + b_{22}y_2. \end{aligned} \quad (1)$$

Вычислим производные произведений $x_1y_1, x_1y_2, x_2y_1, x_2y_2$. Имеем

$$\begin{aligned} \frac{d}{dt}(x_1y_1) &= (a_{11}x_1 + a_{12}x_2)y_1 + x_1(b_{11}y_1 + b_{12}y_2) = \\ &= (a_{11} + b_{11})x_1y_1 + b_{12}x_1y_2 + a_{12}x_2y_1 \end{aligned} \quad (2)$$

и т. д.

Легко видеть, что матрица, которую мы получаем, в точности равна $A \times I_N + I_M \times B$.

§ 13. Уравнение $AX + XB = C$. В гл. 11 было показано, что уравнение

$$AX + XB = C \quad (1)$$

имеет единственное решение

$$X = - \int_0^{\infty} e^{At} C e^{Bt} dt \quad (2)$$

при условии, что интеграл в правой части существует при любой матрице C .

Дополним этот результат.

Рассмотрим случай, когда A , B , C и X являются матрицами второго порядка. Уравнения для неизвестных элементов x_{ij} имеют вид

$$\begin{aligned} a_{11}x_{11} + a_{12}x_{21} + x_{11}b_{11} + x_{12}b_{21} &= c_{11}, \\ a_{11}x_{12} + a_{12}x_{22} + x_{11}b_{12} + x_{12}b_{22} &= c_{12}, \\ a_{21}x_{11} + a_{22}x_{21} + x_{21}b_{11} + x_{22}b_{21} &= c_{21}, \\ a_{21}x_{12} + a_{22}x_{22} + x_{21}b_{12} + x_{22}b_{22} &= c_{22}. \end{aligned} \quad (3)$$

Матрица коэффициентов

$$\begin{vmatrix} a_{11} + b_{11} & b_{21} & a_{12} & 0 \\ b_{12} & a_{11} + b_{22} & 0 & a_{12} \\ a_{21} & 0 & a_{22} + b_{11} & b_{21} \\ 0 & a_{21} & b_{12} & a_{22} + b_{22} \end{vmatrix} \quad (4)$$

есть не что иное, как матрица $A \times I + I \times B'$.

Собственные значения этой матрицы равны $\lambda_i + \mu_j$, поскольку собственные значения матриц B и B' совпадают. Нетрудно видеть, что этот результат может быть обобщен на случай матриц произвольного порядка. Поэтому имеет место

Теорема 4. Для того чтобы уравнение (1) имело решение при любой матрице C , необходимо и достаточно, чтобы $\lambda_i + \mu_j \neq 0$, где λ_i — собственные значения матрицы A , а μ_j — собственные значения матрицы B .

Упражнения

1. Доказать, что необходимым и достаточным условием того, что уравнение $AX + XA' = C$ имеет единственное решение при всех матрицах C , является неравенство $\lambda_i + \lambda_j \neq 0$ для всех i и j .

2. Доказать предыдущий результат следующим образом:

(а) Пусть T — матрица, приводящая A к треугольной форме

$$T^{-1}AT = B = \begin{vmatrix} b_{11} & b_{12} & \dots & b_{1N} \\ & b_{22} & \dots & b_{2N} \\ & & \dots & \dots \\ & & & b_{NN} \end{vmatrix} +$$

тогда $AX + XA' = C$ переходит в уравнение

$$B'(T'XT) + (T'XT)B = T'CT.$$

б) Пусть $Y = T'XT$. Рассмотреть линейное уравнение для элементов матрицы Y . Показать, что определитель этой системы равен $\prod_{i,j=1}^N (b_{ii} + b_{jj})$, и вывести отсюда сформулированный результат. (Хан.)

§ 14. Другое доказательство. В предыдущих параграфах мы отметили, что кронекеровские произведения возникают из рассмотрения симметрических функций собственных значений двух различных матриц A и B . Для случая одной матрицы A была определена кронекеровская степень частного вида.

Рассмотрим теперь другой тип степеней матриц, возникающих при рассмотрении некоторых кососимметрических функций, а именно определителей. Отметим, что формальным на первый взгляд выкладкам можно дать четкую геометрическую интерпретацию. Читателю, интересующемуся геометрической стороной затрагиваемых вопросов, мы рекомендуем обратиться к книге Бохера и Клейна, ссылка на которую имеется в библиографии к этой главе.

Для того чтобы подчеркнуть основную идею, не слишком углубляясь в арифметику и алгебру, рассмотрим матрицу третьего порядка и совокупность определителей второго порядка, образованных из собственных векторов этой матрицы. В принципе ясно, как обобщить предлагаемую процедуру на случай определителей порядка R , образованных из собственных векторов матриц порядка N . На практике это выливается в чрезвычайно трудоемкие выкладки, включающие вычисления определителей. Таким образом, мы оставляем ни в коем случае не тривиальные детали интересующемуся читателю. Рассмотрим матрицу

$$A = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}, \quad (1)$$

где мы на время отказались от наших обычных обозначений. Это делает более прозрачными наши выкладки и результаты.

Для любых двух собственных векторов x^i и x^j , $i \neq j$, образуем набор уравнений для определителей y_1, y_2, y_3 второго порядка, заданных соотношениями:

$$\begin{aligned} \lambda_i \lambda_j y_1 &= \begin{vmatrix} \lambda_i x_1^i & \lambda_j x_1^j \\ \lambda_i x_2^i & \lambda_j x_2^j \end{vmatrix}, & \lambda_i \lambda_j y_2 &= \begin{vmatrix} \lambda_i x_2^i & \lambda_j x_2^j \\ \lambda_i x_3^i & \lambda_j x_3^j \end{vmatrix}, \\ \lambda_i \lambda_j y_3 &= \begin{vmatrix} \lambda_i x_3^i & \lambda_j x_3^j \\ \lambda_i x_1^i & \lambda_j x_1^j \end{vmatrix}. \end{aligned} \quad (2)$$

Учитывая равенства

$$\begin{aligned} \lambda_i x_1^i &= a_1 x_1^i + a_2 x_2^i + a_3 x_3^i, \\ \lambda_i x_2^i &= b_1 x_1^i + b_2 x_2^i + b_3 x_3^i, \\ \lambda_i x_3^i &= c_1 x_1^i + c_2 x_2^i + c_3 x_3^i, \end{aligned} \quad (3)$$

а также аналогичные равенства с заменой i на j , легко заключить в силу (2), что

$$\lambda_i \lambda_j y_1 = \begin{vmatrix} a_1 x_1^i + a_2 x_2^i + a_3 x_3^i & a_1 x_1^j + a_2 x_2^j + a_3 x_3^j \\ b_1 x_1^i + b_2 x_2^i + b_3 x_3^i & b_1 x_1^j + b_2 x_2^j + b_3 x_3^j \end{vmatrix}. \quad (4)$$

Если \mathbf{a} — вектор с компонентами a_i , а \mathbf{b} — вектор с компонентами b_i , то (4) эквивалентно уравнению

$$\lambda_i \lambda_j y_1 = \begin{vmatrix} (\mathbf{a}, \mathbf{x}^i) & (\mathbf{a}, \mathbf{x}^j) \\ (\mathbf{b}, \mathbf{x}^i) & (\mathbf{b}, \mathbf{x}^j) \end{vmatrix}. \quad (5)$$

В этом месте нам потребуется обобщить разложение определителя Грама, данное в § 5, гл. 4. К счастью, это обобщение существует*). Легко проверить, что

$$\begin{vmatrix} (\mathbf{a}, \mathbf{x}^i) & (\mathbf{a}, \mathbf{x}^j) \\ (\mathbf{b}, \mathbf{x}^i) & (\mathbf{b}, \mathbf{x}^j) \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \begin{vmatrix} x_1^i & x_1^j \\ x_2^i & x_2^j \end{vmatrix} + \\ + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \begin{vmatrix} x_2^i & x_2^j \\ x_3^i & x_3^j \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ a_1 & b_1 \end{vmatrix} \begin{vmatrix} x_3^i & x_3^j \\ x_1^i & x_1^j \end{vmatrix}. \quad (6)$$

В силу симметрии отсюда непосредственно следует, что

$$\lambda_i \lambda_j \begin{vmatrix} y_1 \\ y_2 \\ y_3 \end{vmatrix} = \begin{vmatrix} g_{11}(\mathbf{a}, \mathbf{b}) & g_{12}(\mathbf{a}, \mathbf{b}) & g_{13}(\mathbf{a}, \mathbf{b}) \\ g_{12}(\mathbf{a}, \mathbf{c}) & g_{22}(\mathbf{a}, \mathbf{c}) & g_{23}(\mathbf{a}, \mathbf{c}) \\ g_{13}(\mathbf{b}, \mathbf{c}) & g_{23}(\mathbf{b}, \mathbf{c}) & g_{33}(\mathbf{b}, \mathbf{c}) \end{vmatrix} \begin{vmatrix} y_1 \\ y_2 \\ y_3 \end{vmatrix}, \quad (7)$$

где

$$g_{ij}(\mathbf{x}, \mathbf{y}) = \begin{vmatrix} x_i y_i \\ x_j y_j \end{vmatrix}. \quad (8)$$

Прежде всего отметим, что $\lambda_i \lambda_j$ является собственным значением, а \mathbf{y} — собственным вектором матрицы, фигурирующей в системе (7). Поскольку матрица не зависит от индексов i и j , мы видим, что собственными значениями этой матрицы могут быть только $\lambda_1 \lambda_1$, $\lambda_1 \lambda_2$ и $\lambda_2 \lambda_2$. Таким образом, мы получили способ определения матрицы, собственными значениями которой являются $\lambda_i \lambda_j$, $i \neq j$.

Упражнения

1. Возвращаясь к обычным обозначениям для элементов матрицы A , выписать выражения для i, j -го элемента матрицы, фигурирующей в формуле (7).

2. Пусть A — квадратная матрица порядка N . Воспользовавшись описанной процедурой, найти формулу для элементов матрицы порядка $N(N-1)/2$, собственными значениями которой являются $\lambda_i \lambda_j$, $i \neq j$.

*) См. Ф. Р. Гантмахер, Теория матриц, «Наука», 1967, стр. 20, формула Бине — Коши. (Прим. ред.)

3. На основании полученных результатов показать, что для симметрической матрицы A имеет место неравенство

$$\lambda_1 \lambda_2 \geq \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

а также ряд неравенств аналогичного вида.

4. Пусть A — квадратная матрица порядка N , а r — целое число, лежащее между 1 и N . Обозначим через S_r совокупность всех наборов из r различных целых чисел, выбранных из отрезка натурального ряда 1, 2, ..., N . Пусть s и t — два элемента, принадлежащие S_r , а A_{st} — матрица, образованная из матрицы A вычеркиванием всех строк с индексами, не входящими в s , и всех столбцов с индексами, не входящими в t .

Пусть все элементы S_r перенумерованы в определенном порядке, скажем s_1, s_2, \dots, s_M . Матрица порядка M , где

$$M = \frac{N!}{r!(N-r)!},$$

$$C_r(A) = \| |A_{s_i s_j}| \|.$$

называется r -й ассоциированной матрицей для матрицы A . Установить следующие результаты:

- (а) $C_r(AB) = C_r(A) \cdot C_r(B)$;
 (б) $C_r(A') = C_r(A)'$;
 (в) $C_r(A^{-1}) = C_r(A)^{-1}$;
 (г) $|C_r(A)| = |A|^k, \quad k = \frac{(N-1)!}{(r-1)!(N-r)!}$;

(д) собственными значениями матрицы $C_r(A)$ являются $\mu_1, \mu_2, \dots, \mu_M$, где $\mu_1 + \mu_2 + \dots + \mu_M - r$ -я элементарная симметрическая функция от $\lambda_1, \lambda_2, \dots, \lambda_N$, т. е. $\mu_1 = \lambda_1 \lambda_2 \dots \lambda_r, \mu_2 = \lambda_1 \lambda_2 \dots \lambda_{r-1} \lambda_{r+1}, \dots$ ¹⁾.

По поводу ассоциированных матриц см. Резерфорд (D. E. Rutherford), Compound Matrices, Koninkl. Ned. Acad. Wetenschap. Amsterdam, Proc. Sect. Sci., ser. A 54 (1951), 16—22^{*)}. В книге Вайценбука (R. Weitzenbuck, Invariantentheorie, Groningen, 1923) рассматривается связь теории ассоциированных матриц с теорией инвариантов.

§ 15. Циркулянты. При рассмотрении многих вопросов встречаются матрицы вида

$$C = \begin{vmatrix} c_0 & c_1 & c_2 & \dots & c_{N-1} \\ c_{N-1} & c_0 & c_1 & \dots & c_{N-2} \\ c_{N-2} & c_{N-1} & c_0 & \dots & c_{N-3} \\ \dots & \dots & \dots & \dots & \dots \\ c_1 & c_2 & c_3 & \dots & c_0 \end{vmatrix}, \quad (1)$$

¹⁾ Дальнейшие результаты и ссылки можно найти в работе Райзера (H. J. Ryser), Inequalities of Compound and Induced Matrices with Applications to Combinatorial Analysis, III, J. Math. 2 (1958), 240—253.

^{*)} См. также Ф. Р. Гантмахер, Теория матриц, «Наука», 1967, гл. 1, § 4. (Прим. ред.)

2. Верно ли, что

$$\epsilon^{\text{Sp}} X = \sum_{k=0}^{\infty} \frac{\text{Sp}(X^{1+k})}{k!} \gamma$$

3. Если A и B — положительно определенные матрицы, то матрица $A \times B$ положительно определенная.

4. Если A и B — симметрические матрицы, причем $A \geq B$, то $A^{(n)} \geq B^{(n)}$ для $n = 1, 2, \dots$

5. Пусть r удовлетворяет уравнению $r^N + a_1 r^{N-1} + \dots + a_N = 0$, а s — уравнению $s^M + b_1 s^{M-1} + \dots + b_M = 0$. Как найти уравнение, корнями которого являются $r_i s_j$?

6. Пусть $1, \alpha_1, \dots, \alpha_{N-1}$ — элементы конечной группы, и пусть $x = x_0 + \alpha_1 x_1 + \dots + \alpha_{N-1} x_{N-1}$, где x_i — скаляры. Рассмотрим произведения $\alpha_i x = \alpha_i(0) x_0 + \alpha_i(1) x_1 + \dots + \alpha_i(N-1) x_{N-1}$, где элементы $\alpha_i(j)$ представляют собой α_i , упорядоченные некоторым образом, или $\alpha_i x = x_0(i) + x_1(i) \alpha_1 + \dots + x_{N-1}(i) \alpha_{N-1}$, где $x_i(j)$ представляют собой x_i , упорядоченные некоторым образом. Введем матрицу $X = \|x_i(j)\|$ и определим соответствие $X \sim x$. Если $X \sim x$, а $Y \sim y$, то верно ли, что $XY \sim xy$?

7. Пусть $1, \alpha_1, \dots, \alpha_{N-1}$ — элементы конечной группы G , а $1, \beta_1, \dots, \beta_{M-1}$ — элементы конечной группы H . Рассмотрим прямое произведение G и H , которое определяется как группа порядка MN с элементами $\alpha_i \beta_j$. Используя процедуру, описанную в предыдущем упражнении, получить матрицу, соответствующую

$$\left(\sum_{i=0}^{N-1} x_i \alpha_i \right) \left(\sum_{j=0}^{M-1} y_j \beta_j \right).$$

Какова связь между этой матрицей и матрицами

$$X \sim \sum_{i=0}^{N-1} x_i \alpha_i, \quad Y \sim \sum_{j=0}^{M-1} y_j \beta_j?$$

В работе О. Таусски¹⁾ исследуется связь между понятием групповой матрицы и понятием нормальности.

8. Уравнение $AX - XA = \lambda X$ имеет нетривиальное решение X тогда и только тогда, когда $\lambda = \lambda_i - \lambda_j$, где $\lambda_1, \lambda_2, \dots, \lambda_N$ — собственные значения матрицы A . (Лаппо-Данилевский.)

9. Пусть F — матрица порядка N^2 , разбитая на N подматриц, так что каждая из подматриц имеет порядок N , причем каждая подматрица f_{ij} является рациональной функцией $f_{ij}(A)$ некоторой матрицы A порядка N . Если $\lambda_1, \lambda_2, \dots, \lambda_N$ — собственные значения матрицы A , то собственными значениями матрицы F являются собственные значения \forall матриц $\|f_{ij}(\lambda_k)\|$, $k = 1, 2, \dots, N$, каждая из которых имеет порядок N . См. Вильямсон²⁾ и Аффрайэт³⁾. В работах Тодда⁴⁾ и Лоуэн⁵⁾ эти вопросы излагаются

¹⁾ O. Taussky, A Note of Group Matrices, Proc. Amer. Math. Soc. 6 (1955), 984—986.

²⁾ J. D. W. Williamson, The Latent Roots of a Matrix of Special Type, Bull. Amer. Math. Soc. 37 (1931), 585—590.

³⁾ S. N. Afriat, Composite Matrices, Quart. J. Math., Oxford 2d, ser. 5, 1954, 81—98.

⁴⁾ J. Todd, The Condition of Certain Matrices, III J. Research Nat. Bur. Standards 60 (1958), 1—7.

⁵⁾ A. N. Lowan, The Operator Approach to Problems of Stability and Convergence of Solutions of Difference Equations, Scripta Math., no 8, 1957.

применительно к численному решению дифференциальных уравнений в частных производных.

10. Пусть $f(\theta)$ — действительная функция, определенная на отрезке $-\pi \leq \theta \leq \pi$. Рассмотрим коэффициенты Фурье функции $f(\theta)$:

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta, \quad n = 0, \pm 1, \pm 2, \dots$$

Конечные матрицы вида

$$T_N = \|c_{k-l}\|, \quad k, l = 0, 1, 2, \dots, N,$$

называются *тёплицевыми матрицами* порядка N . Показать, что матрица T_N является эрмитовой.

11. Обозначим через $\lambda_1^{(N)}, \lambda_2^{(N)}, \dots, \lambda_{N+1}^{(N)}$ собственные значения матрицы T_N . Показать, что

$$\lim_{N \rightarrow \infty} \frac{\lambda_1^{(N)} + \lambda_2^{(N)} + \dots + \lambda_{N+1}^{(N)}}{N+1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) d\theta.$$

12. Показать, что

$$\lim_{N \rightarrow \infty} \frac{(\lambda_1^{(N)})^2 + (\lambda_2^{(N)})^2 + \dots + (\lambda_{N+1}^{(N)})^2}{N+1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2(\theta) d\theta.$$

(Последние два результата являются частными случаями общего результата, установленного Сегё в 1917 г. Изложение более поздних результатов читатель найдет в работе Каца, Мердока и Сегё¹⁾).

Обобщение тёплицевых матриц было получено Гренандером²⁾, см. также работу Уидома³⁾.

13. С помощью кронекеровской суммы двух матриц $A \oplus B$ определить кронекеровскую производную от переменной матрицы $X(t)$ соотношением

$$\frac{\delta X}{\delta t} = \lim_{h \rightarrow 0} \frac{X(t+h) \oplus (-X(t))}{h}.$$

14. Рассмотреть дифференциальное уравнение

$$\frac{\delta Y}{\delta t} = A \times Y, \quad Y(0) = I.$$

Имеет ли оно решение? Если да, то найти его.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 1. Прямое произведение матриц тесно связано с понятием прямого произведения групп. Изложение этих вопросов имеется в монографиях

¹⁾ M. Kac, W. L. Murdoch and G. Szegő, On the Eigenvalues of Certain Hermitian Forms, J. Rat. Mech. Analysis 2 (1953), 767—800.

²⁾ U. Grenander, Trans. Amer. Math. Soc., 1958. [См. также У. Гренандер и Г. Сегё, Тёплицевы формы и их приложения, ИЛ, 1961. (Прим. ред.)]

³⁾ H. Widom, On the Eigenvalues of Certain Hermitian Operators, Trans. Amer. Math. Soc. 88 (1958), 491—522.

Мак-Даффи (С. С. MacDuffee), *The Theory of Matrices*, Chelsea Publishing Co., New York, 1946, chapter VII;

Мурнаган (F. D. Murphagan), *The Theory of Group Representation*, Johns Hopkins Press, Baltimore, 1938. [Русский перевод: Теория представлений групп, ИЛ, 1950].

Здесь мы хотели только показать, как это понятие возникает в связи с двумя различными задачами. Одна из них рассмотрена в этой главе, другая будет изложена в гл. 15. См. также

Картан (E. Cartan), *Oeuvres Completes*, part 1, vol. 2, 1045—1080,

где вводится понятие *зонной гармоник* положительно определенной матрицы.

§ 4. Эта процедура заимствована из работы

Беллман (R. Bellman), *Limit Theorems for Non-commutative Operations*, I, *Duke Math. J.* 21 (1954), 491—500.

§ 10. Кронекеровский логарифм введен в только что цитированной работе Беллмана.

§ 11. Теорема 4 приводится в цитированной выше книге Мак-Даффи. Эта теорема независимо была получена Ханом. См.

Хан (W. Hahn), *Eine Bemerkung zur zweiten Methode von Lyapunov*, *Math. Nachr.* 14, No 4/6 (1956), 349—354.

Этот же результат позднее был независимо получен Беллманом. См.

Беллман (R. Bellman), *Kronecker Products and the Second Method of Lyapunov*, *Math. Nachr.*, 1959.

§ 13. Обширные исследования были проведены для соответствующих операторных уравнений. Помимо самостоятельного интереса, который представляют эти уравнения, это вызвано еще и той ролью, которую они играют в квантовой механике. См.

Розенблюм (M. Rosenbloom), *On the Operator Equation $BX - XA = C$* , *Duke Math. J.* 23 (1956);

Хайнц (E. Heinz), *Math. Ann.* 123 (1951), 415—438;

Резерфорд (D. E. Rutherford), *Koninkl. Ned. Akad. Wetenschap., Proc., ser. A* 35 (1932), 54—59.

§ 14. Вопросы, затронутые в этом параграфе, относятся к теории инвариантов в линейных преобразованиях. См.

Бохер (M. Bocher), *Introduction to Higher Algebra*, The Macmillan Company, New York, reprinted, 1947;

Клейн (F. Klein), *Elementary Mathematics from an Advanced Standpoint, Geometry*, Dover Publication, New York [русский перевод: Элементарная математика с точки зрения высшей, т. I, ГТТИ, 1933; т. II, ГТТИ, 1934];

Вейль (H. Weyl), *The Classical Groups*, Princeton University Press, Princeton, N. J., 1946 [русский перевод: Классические группы, их инварианты и представления, ИЛ, 1947].

Определители второго порядка являются координатами Плюкера. Основная идея состоит в том, что каждое линейное преобразование имеет большое число связанных с ним индуцированных преобразований, которые позволяют очень просто вывести некоторые свойства исходного преобразования. Как мы уже упоминали, желая избежать громоздких вычислений, связанных с определителями, мы не разбирали эти вопросы подробно. См. статью

Швердфегер (H. Schwerdtfeger), *Skew-symmetric Matrices and Projective Geometry*, Amer. Math. Monthly 51 (1944), 137—148.

В упражнениях 15 и 16 к гл. 14 мы встретимся с вероятностной интерпретацией координат Плюкера, где также дается ссылка на работу Карлина и Мак-Грегора.

§ 15. Циркулянты играют важную роль во многих физико-математических теориях.

См. статью

Берлин и Кау (T. H. Berlin, M. Kac), *The Spherical Model of a Ferromagnet*, Phys. Rev. 88 (1952), 821—835,

где можно найти дальнейшие ссылки на работы Онзагера и др., связанные с прямыми произведениями матриц и групп.

Ассоциированные матрицы рассматриваются в работах

Райзер (H. J. Ryser), *Inequalities of Compound and Induced Matrices with Applications to Combinatorial Analysis*, III. J. Math. 2 (1958), 240—253;

Де-Брёйн (N. G. de Bruijn), *Inequalities Concerning Minors and Eigenvalues*, Nieuw. Arch. Wisk. (3) 4 (1956), 18—35;

Литлвуд (D. E. Littlewood), *The Theory of Group Characters and Matrix Representations of Group*, Oxford, New York, 1950;

Шур (I. Schur), *Über eine Klasse von Matrizen die sich einer gegebenen Matrix zuordnen lassen*, Dissertation Berlin, 1901;

Веддерберн (J. H. M. Wedderburn), *Lectures on Matrices*, Amer. Math. Soc. Colloq. Publ 17 (1934);

Мак-Даффи (C. C. MacDuffee), *The Theory of Matrices*, Chelsea Publishing Company, New York, 1946;

Тёплиц (O. Toeplitz), *Das algebraische Analogon zu einem Satze von Fejer*, Math. Z. 2 (1918), 187—197;

Маркус, Мойлс и Вествик (M. Marcus, B. N. Moys and R. Westwick), *Extremal Properties of Hermitian Matrices II*, Canad. J. Math., 1959.

ТЕОРИЯ УСТОЙЧИВОСТИ

§ 1. Введение. Чрезвычайно важной задачей является исследование поведения физической системы в окрестности положения равновесия. Если система, выведенная из положения равновесия малыми возмущениями, возвращается в это положение, то эта система называется *устойчивой*. В противном случае система называется *неустойчивой* *).

Хотя очень часто наличие этого свойства можно проверить экспериментально, такой эксперимент обычно слишком дорог и требует больших затрат времени. Следовательно, при построении системы всегда желательно иметь математический критерий устойчивости.

Как мы уже отмечали в гл. 11, линейные уравнения вида

$$\frac{dx}{dt} = Ax, \quad x(0) = c, \quad (1)$$

во многих случаях могут быть использованы для исследования поведения системы в окрестности положения равновесия, которое в данном случае соответствует точке $x = 0$.

Поэтому мы начнем с вывода необходимых и достаточных условий того, что решение системы (1) стремится к нулю при $t \rightarrow \infty$. Реальная экономическая, техническая или физическая задача на самом деле много сложнее, поскольку уравнения, описывающие процесс, являются нелинейными. В частности, эти уравнения могут иметь вид

$$\frac{dy}{dt} = Ay + g(y), \quad y(0) = c. \quad (2)$$

Возникает вопрос, может ли критерий, выведенный для линейных систем, служить для исследования устойчивости в общем нелинейном случае. Оказывается, что при довольно естественных предположениях на этот вопрос можно ответить положительно. Этот вывод является одним из результатов классических работ

*) См. примечание к § 1 гл. 10. (Прим. ред.)

Пуанкаре и Ляпунова. Однако здесь мы не будем заниматься нелинейным случаем и ограничимся рассмотрением линейных уравнений.

§ 2. Необходимые и достаточные условия устойчивости. Начнем с доказательства основного результата.

Теорема 1. Для того чтобы решение уравнения

$$\frac{dx}{dt} = Ax, \quad x(0) = c, \quad (1)$$

стремилось к нулю при $t \rightarrow \infty$ и любом c , необходимо и достаточно, чтобы собственные значения матрицы A имели отрицательные вещественные части.

Доказательство. Если среди собственных значений матрицы нет одинаковых, то представление

$$e^{At} = T \begin{pmatrix} e^{\lambda_1 t} & & & 0 \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_N t} \\ 0 & & & & 0 \end{pmatrix} T^{-1} \quad (2)$$

делает утверждение теоремы очевидным. К сожалению, мы не можем прибегнуть к соображениям непрерывности, чтобы доказать сформулированный результат для общего случая. Поэтому мы поступим следующим образом.

Вместо того чтобы приводить матрицу A к диагональному виду, преобразованием подобия приведем ее к треугольной форме, $T^{-1}AT = B$.

Система (1) в этом случае принимает вид

$$\frac{dz}{dt} = Bz, \quad z(0) = c', \quad (3)$$

где B — треугольная матрица, а $x = Tz$, или, в скалярной форме,

$$\begin{aligned} \frac{dz_1}{dt} &= b_{11}z_1 + b_{12}z_2 + \dots + b_{1N}z_N, & z_1(0) &= c'_1, \\ \frac{dz_2}{dt} &= & b_{22}z_2 + \dots + b_{2N}z_N, & z_2(0) &= c'_2, \\ & \dots & & & \\ \frac{dz_N}{dt} &= & & & b_{NN}z_N, & z_N(0) &= c'_N. \end{aligned} \quad (4)$$

Поскольку элементы b_{ii} являются собственными значениями матрицы A , то по предположению $\operatorname{Re}(b_{ii}) < 0$ для $i = 1, 2, \dots, N$.

Последнее из уравнений (4) дает

$$z_N = c'_N e^{b_N t} N N^t, \quad (5)$$

откуда следует, что $z_N \rightarrow 0$ при $t \rightarrow \infty$.

Для того чтобы показать, что все $z_i \rightarrow 0$ при $t \rightarrow \infty$, мы воспользуемся индукцией, основанной на следующем результате.

Если $v(t) \rightarrow 0$ при $t \rightarrow \infty$, то решение $u(t)$ системы

$$\frac{du}{dt} = b_1 u + v(t), \quad u(0) = a_1, \quad (6)$$

стремится к нулю при условии, что $\operatorname{Re}(b_1) < 0$. Это сразу же следует из того, что

$$u(t) = a_1 e^{b_1 t} + e^{b_1 t} \int_0^t e^{-b_1 s} v(s) ds. \quad (7)$$

Поскольку z_N стремится к нулю при $t \rightarrow \infty$, то в силу сказанного выше мы получаем, что все $z_i \rightarrow 0$ при $t \rightarrow \infty$.

Упражнения

1. Доказать теорему 1, используя каноническую форму Жордана.

§ 3. Устойчивые матрицы. Чтобы избежать ненужных повторений, введем новый термин.

Определение. Матрица A называется *устойчивой*, если ее собственные значения имеют отрицательные действительные части.

Упражнения

1. Выразить необходимые и достаточные условия устойчивости действительной матрицы в терминах матрицы $\|\operatorname{Sp}(A^{t+j})\|$. (Басс.)

2. Каковы соответствующие условия в случае, если матрица A комплексная?

§ 4. Метод Ляпунова. Посмотрим теперь, как квадратичные формы могут быть использованы при изучении асимптотического поведения решений линейных дифференциальных уравнений. Этот метод принадлежит Ляпунову и играет главную роль в современной теории устойчивости нелинейных функциональных уравнений всех видов.

Рассмотрим уравнение

$$\frac{dx}{dt} = Ax, \quad x(0) = c, \quad (1)$$

где c и A предполагаются действительными, а также квадратичную форму

$$u = (x, Yx), \quad (2)$$

где Y — симметрическая постоянная матрица, элементы которой пока не определены. Имеем

$$\begin{aligned} \frac{du}{dt} &= (\mathbf{x}', Y\mathbf{x}) + (\mathbf{x}, Y\mathbf{x}') = (A\mathbf{x}, Y\mathbf{x}) + (\mathbf{x}, YA\mathbf{x}) = \\ &= (\mathbf{x}, (A'Y + YA)\mathbf{x}). \end{aligned} \quad (3)$$

Предположим, что Y является решением уравнения

$$A'Y + YA = -I, \quad (4)$$

к тому же мы потребуем, чтобы матрица Y была положительно определенной. Тогда (3) приобретает вид

$$\frac{du}{dt} = -(\mathbf{x}, \mathbf{x}), \quad (5)$$

откуда с учетом (2)

$$\frac{du}{dt} \leq -\lambda_N u, \quad (6)$$

где λ_N — наибольшее собственное значение матрицы Y . Из (6) следует, что $u \leq u(0)e^{-\lambda_N t}$, т. е. $u \rightarrow 0$ при $t \rightarrow \infty$. Положительная определенность матрицы Y гарантирует стремление к нулю всех компонент вектора \mathbf{x} при $t \rightarrow \infty$.

Вместе с тем мы знаем (см. § 13 гл. 12), что если матрица A устойчива, то уравнение (4) имеет единственное симметрическое* решение Y . Поскольку при этом Y определяется формулой

$$Y = \int_0^{\infty} e^{A't_1} e^{At_1} dt_1, \quad (7)$$

то

$$(\mathbf{x}, Y\mathbf{x}) = \int_0^{\infty} (\mathbf{x}, e^{A't_1} e^{At_1} \mathbf{x}) dt_1 = \int_0^{\infty} (e^{At_1} \mathbf{x}, e^{At_1} \mathbf{x}) dt_1. \quad (8)$$

Таким образом, матрица Y является положительно определенной, так как матрица e^{At} всегда невырожденная.

Упражнения

1. Рассмотрим уравнение $\frac{dx}{dt} = A\mathbf{x} + g(\mathbf{x})$, $\mathbf{x}(0) = \mathbf{c}$, где

- (а) матрица A устойчивая;
- (б) $\|g(\mathbf{x})\|/\|\mathbf{x}\| \rightarrow 0$ при $\|\mathbf{x}\| \rightarrow 0$;
- (в) $\|\mathbf{c}\|$ достаточно мало.

Пусть матрица Y обладает свойствами, указанными в § 4. Доказать, что если \mathbf{x} является решением приведенного нелинейного уравнения и

* Симметричность Y следует из вида уравнения (4): достаточно в (4) перейти к транспонированным матрицам. См. также (7). (Прим. ред.)

выполнены указанные условия, то

$$\frac{d}{dt}(x, Yx) < -r_1(x, Yx),$$

где r_1 — положительная константа. Вывести отсюда, что $x \rightarrow 0$ при $t \rightarrow \infty$.

2. Обобщить предыдущие рассуждения на случай комплексной матрицы A .

§ 5. Среднеквадратичное отклонение. Предположим, что A — устойчивая матрица, и нам требуется вычислить интеграл

$$J = \int_0^{\infty} (x, Bx) dt, \quad (1)$$

где x является решением уравнения (4.1). Здесь интересно отметить, что интеграл J можно вычислить как рациональную функцию элементов матрицы A , не зная решения линейного дифференциального уравнения. В частности, нет необходимости вычислять собственные значения матрицы A .

Определим постоянную матрицу Y с помощью соотношения

$$(x, Bx) = \frac{d}{dt}(x, Yx). \quad (2)$$

Ясно, что

$$B = A'Y + YA. \quad (3)$$

Определив таким образом матрицу Y , мы получаем следующее выражение для интеграла J :

$$J = -(c, Yc). \quad (4)$$

Поскольку матрица A устойчива, уравнение (3) имеет единственное решение, которое может быть выражено через определители.

§ 6. Некоторые эффективные критерии устойчивости. Задача проверки матрицы на устойчивость является довольно трудной, и до сих пор не получено сколько-нибудь простого ее решения. Эту задачу осложняет то обстоятельство, что мы не столько заинтересованы в ее решении для данной конкретной матрицы A , сколько в получении условий, которые позволяли бы нам утверждать, что различные представители класса матриц $A(\mu)$ являются устойчивыми. Вопросы такого типа постоянно возникают при конструировании систем управления, в математической экономике, а также при изучении вычислительных алгоритмов.

Если характеристический многочлен матрицы A получен, то существует множество критериев, с помощью которых мы можем судить, имеют ли собственные значения матрицы A отрицательные вещественные части или нет. По-видимому, наиболее полезным из них является критерий Гурвица.

Рассмотрим уравнение

$$|\lambda I - A| = \lambda^N + a_1 \lambda^{N-1} + \dots + a_{N-1} \lambda + a_N = 0. \quad (1)$$

Для многочлена, стоящего в правой части (1), составим бесконечную таблицу

$$\begin{array}{cccccccc} a_1 & 1 & 0 & 0 & 0 & 0 & \dots & \\ a_2 & a_2 & a_1 & 1 & 0 & 0 & \dots & \\ a_3 & a_3 & a_2 & a_2 & a_1 & 1 & \dots & \\ a_4 & a_4 & a_3 & a_3 & a_2 & a_2 & \dots & \\ \dots & \end{array} \quad (2)$$

где коэффициенты k полагаются равными нулю, если $k > N$. Для того чтобы все корни уравнения (1) имели отрицательные вещественные части, необходимо и достаточно, чтобы все определители из последовательности

$$h_1 = |a_1|, \quad h_2 = \begin{vmatrix} a_1 & 1 \\ a_2 & a_2 \end{vmatrix}, \quad h_3 = \begin{vmatrix} a_1 & 1 & 0 \\ a_2 & a_2 & a_1 \\ a_3 & a_3 & a_2 \end{vmatrix}, \quad \dots \quad (3)$$

были положительны.

Не существует сколько-нибудь простого доказательства этого результата. В следующем параграфе мы наметим в общих чертах один из возможных путей, а в приложении Б коснемся некоторых идей Эрмита, которые послужили основой для доказательства Гурвица. Оба эти способа опираются на квадратичные формы.

Ссылки на работы, где приводятся другие доказательства, даны в конце главы.

Причина, почему этот результат не всегда является эффективным, состоит в необходимости вычисления определителя $|\lambda I - A|$, что является крайне нежелательным, если порядок матрицы A достаточно велик.

Упражнения

1. Будем называть многочлен *устойчивым*, если все его корни имеют отрицательные вещественные части. Используя приведенный критерий, показать, что необходимым и достаточным условием устойчивости квадратного трехчлена $\lambda^2 + a_1 \lambda + a_2$ является положительность коэффициентов: $a_1 > 0$, $a_2 > 0$.

2. Для многочлена $\lambda^3 + a_1 \lambda^2 + a_2 \lambda + a_3$ показать, что соответствующие условия устойчивости состоят в выполнении неравенств

$$a_1, a_2, a_3 > 0, \quad a_1 a_2 > a_3.$$

3. Показать, что для многочлена $\lambda^4 + a_1 \lambda^3 + a_2 \lambda^2 + a_3 \lambda + a_4$ эти условия сводятся к неравенствам $a_1, a_4 > 0$, $a_1 a_2 > a_3$, $a_2 (a_1 a_2 - a_3) > a_4^2$.

§ 7. Необходимое и достаточное условие устойчивости матриц. Покажем теперь, что из результата, установленного в § 4, следует Теорема 2. Пусть Y определяется соотношением

$$A'Y + YA = -I, \quad (1)$$

тогда необходимым и достаточным условием устойчивости матрицы A является положительная определенность матрицы Y .

Доказательство. Как и в § 5, нетрудно убедиться, что

$$\int_0^T (\mathbf{x}, \mathbf{x}) dt = (\mathbf{x}(0), Y\mathbf{x}(0)) - (\mathbf{x}(T), Y\mathbf{x}(T)) \quad (2)$$

или

$$(\mathbf{x}(T), Y\mathbf{x}(T)) + \int_0^T (\mathbf{x}, \mathbf{x}) dt = (\mathbf{x}(0), Y\mathbf{x}(0)). \quad (3)$$

Здесь \mathbf{x} является решением уравнения $\frac{d\mathbf{x}}{dt} = A\mathbf{x}$. Если матрица Y

положительно определена, то $\int_0^T (\mathbf{x}, \mathbf{x}) dt$ равномерно ограничен,

откуда следует, что $\mathbf{x}(t) \rightarrow 0$ при $t \rightarrow \infty^*$, а следовательно, матрица A является устойчивой. Тот факт, что матрица Y является положительно определенной, если матрица A устойчива, мы уже доказали ранее.

§ 8. Дифференциальные уравнения и собственные значения. Рассмотрим дифференциальное уравнение

$$A\mathbf{x}'' + 2B\mathbf{x}' + C\mathbf{x} = 0, \quad \mathbf{x}(0) = \mathbf{c}^1, \quad \mathbf{x}'(0) = \mathbf{c}^2, \quad (1)$$

где матрицы A , B и C являются неотрицательно определенными. Такие уравнения часто возникают при исследовании электрических цепей, состоящих из емкостей, индуктивностей и сопротивлений. Это обстоятельство наводит на мысль, что должен оказаться справедливым следующий результат:

Теорема 3. Если матрицы A , B и C являются неотрицательно определенными и, кроме того, либо матрица C , либо матрица A являются положительно определенными, то уравнение

$$|\lambda^2 A + 2\lambda B + C| = 0 \quad (2)$$

не имеет корней с положительной действительной частью.

Если же матрицы A и C неотрицательно определенные, а матрица B положительно определена, то единственным корнем с нулевой действительной частью является $\lambda = 0$.

Доказательство. Приведем доказательство, которое опирается на физические соображения.

*) Это связано с тем, что $(\mathbf{x}(t), \mathbf{x}(t))$ представляет собой комбинацию экспонент и полиномов параметра t . (Прим. ред.)

Имеем

$$(x', Ax') + 2(x', Bx') + (x', Cx) = 0. \quad (3)$$

Тогда для любого $s > 0$

$$\int_0^s [(x', Ax') + 2(x', Bx') + (x', Cx)] dt = 0 \quad (4)$$

или

$$(x', Ax') \Big|_0^s + 4 \int_0^s (x', Bx') dt + (x, Cx) \Big|_0^s = 0. \quad (5)$$

Это эквивалентно уравнению

$$(x'(s), Ax'(s)) + 4 \int_0^s (x'(t), Bx'(t)) dt + (x(s), Cx(s)) = c_3, \quad (6)$$

где $c_3 = (c^2, Ac^2) + (c^1, Cc^1)$.

Если λ — корень уравнения (2), то уравнение (1) имеет решение вида $e^{\lambda t}c$. Если λ действительно, то и c — действительный вектор. Если λ комплексно, $\lambda = r_1 + ir_2$, то действительная часть выражения $e^{\lambda t}c$, равная $e^{r_1 t}(a^1 \cos r_2 t + a^2 \sin r_2 t)$, также является решением. Подставляя в (6), мы получаем

$$e^{2r_1 s}(b^1, Ab^1) + 4 \int_0^s e^{2r_1 t}(b^2(t), Bb^2(t)) dt + e^{2r_1 s}(b^3, Cb^3) = c_3, \quad (7)$$

где b^1 и b^3 — постоянные векторы, а $b^2(t)$ — переменный вектор, равный $(a^1 r_1 + a^2 r_2) \cos r_2 t + (a^2 r_1 - a^1 r_2) \sin r_2 t$.

Если A или C положительно определены и при этом $B \geq 0$, то положительность r_1 приводит к противоречию при $s \rightarrow \infty$.

Если $A, C \geq 0$, то из положительной определенности матрицы B вытекает, что $r_1 \leq 0$. Поскольку, однако, функция $b^2(t)$

является периодической, интеграл $\int_0^s (b^2(t), Bb^2(t)) dt$ расходится при $s \rightarrow \infty$, если только r_2 не равно нулю.

Мы детально остановились на этом доказательстве, поскольку оно может быть обобщено на случай, когда матрицы A, B и C переменные.

Упражнения

1. Следуя Анке ¹⁾, воспользуемся описанным приемом для вычисления интеграла $\int_0^{\infty} u^2 dt$ при условии, что u является решением уравнения $u'' +$

¹⁾ Z. angew. Math. u. Phys. 6 (1955), 327—332.

$+a_2u'' + a_1u' + a_0u = 0$, все решения которого стремятся к нулю. Пусть $u(0) = c_0$, $u'(0) = c_1$, $u''(0) = c_2$. Показать, что при этих условиях имеют место следующие соотношения:

$$\int_0^{\infty} u''' u dt = u'u \Big|_0^{\infty} - \int_0^{\infty} u'' u' dt = -c_0 c_2 + c_1^2/2,$$

$$\int_0^{\infty} u'' u dt = u'u \Big|_0^{\infty} - \int_0^{\infty} u'^2 dt = -c_0 c_1 - \int_0^{\infty} u'^2 dt,$$

$$\int_0^{\infty} u' u dt = -c_0^2/2.$$

2. Из уравнений

$$\int_0^{\infty} u^{(i)} (u''' + a_2 u'' + a_1 u' + a_0 u) dt = 0, \quad i = 0, 1, 2,$$

и результата предыдущего упражнения получить линейные уравнения для величин $\int_0^{\infty} u^2 dt$, $\int_0^{\infty} (u')^2 dt$, $\int_0^{\infty} (u'')^2 dt$.

3. Используя полученные линейные уравнения, представить интеграл $\int_0^{\infty} u^2 dt$ как квадратичную форму от c_0 , c_1 , c_2 .

4. Используя указанную квадратичную форму, получить необходимые и достаточные условия устойчивости многочлена $r^3 + a_2 r^2 + a_1 r + a_0$.

5. Показать, что найденные условия эквивалентны условиям Гурвица из упражнения 2 к § 6.

§ 9. Эффективные условия устойчивости матриц. Как уже отмечалось в предыдущих параграфах, существуют достаточно эффективные критерии устойчивости данного многочлена. Но поскольку сама задача нахождения характеристического многочлена матрицы очень трудна, нельзя считать, что мы располагаем удовлетворительными критериями устойчивости матриц. В некоторых частных случаях известны весьма элегантные критерии. Так, например, имеет место

Теорема 4. Если матрица A имеет вид

$$A = \begin{vmatrix} a_1 + b_1 & a_2 & 0 & 0 & \dots \\ -1 & b_2 & a_2 & 0 & \dots \\ 0 & -1 & b_2 & a_4 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & -1 & b_{N-1} & a_N \\ \dots & \dots & 0 & -1 & b_N \end{vmatrix}, \quad (1)$$

где все элементы, не стоящие на главной диагонали и двух соседних с ней, равны нулю, причем элементы a_i действительные, а b_i либо равны нулю, либо являются чисто мнимыми, то число положительных членов в последовательности произведений $a_1, a_1 a_2, \dots, a_1 a_2 \dots a_{N-1} a_N$ равно числу собственных значений матрицы A , имеющих положительные действительные части.

Доказательство, которое проводится с помощью теории последовательностей Штурма, читатель найдет в работе Шварца, ссылка на которую имеется в конце этой главы. В этой работе показано также, что любая комплексная матрица A может быть приведена к виду (1) преобразованием подобия: $A = T^{-1} B T$.

Как мы увидим в гл. 16, теория положительных матриц дает другой подход к вопросам устойчивости в случае, когда все недиагональные элементы матрицы A неотрицательны.

Упражнения к гл. 13

1. Решить скалярное уравнение $u'' + u = f_1(t)$, для чего положить $u' = v$, $v' = -u + f_1(t)$ и вычислить элементы матрицы e^{At} , где

$$A = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}.$$

2. Пусть собственные значения матрицы A различны, а $B(t) \rightarrow 0$ при $t \rightarrow \infty$. Тогда собственные значения матрицы $A + B(t)$, которые мы обозначим $\lambda_i(t)$, различны при $t \geq t_0$.

3. Если, кроме того, $\int_0^{\infty} \|B'(t)\| dt < \infty$, то существует матрица $T(t)$, обладающая тем свойством, что замена переменных $y = Tz$ приводит уравнение $y' = (A + B(t))y$ к виду $z' = (L(t) + C(t))z$, где

$$L(t) = \begin{vmatrix} \lambda_1(t) & & & 0 \\ & \lambda_2(t) & & \\ & & \ddots & \\ 0 & & & \lambda_N(t) \end{vmatrix}.$$

а

$$\int_0^{\infty} \|C(t)\| dt < \infty.$$

4. Если $\sum_{i,j=1}^N \int_0^{\infty} |a_{ij}(t) + a_{ji}(t)| dt < \infty$, то все решения уравнения $y' = A(t)y$ ограничены при $t \rightarrow \infty$.

5. Существует ортогональная матрица $B(t)$ такая, что замена $y = B(t)z$ преобразует уравнение $y' = A(t)y$ к виду $z' = C(t)z$, где матрица $C(t)$ треугольная. (Дилиберто *.)

*) Ср. упражнение 20 к гл. 11. (Прим. ред.)

6 (продолже ие). Существует ограниченная невырожденная матрица $B(t)$ такая, что матрица $C(t)$ является диагональной.

7. Обычное правило дифференцирования произведения двух функций u и v дает $d(uv)/dt = u dv/dt + v du/dt$. Рассмотрим линейную матричную функцию $d(X)$, определяемую соотношением

$$d(X) = AX - XA,$$

где A — фиксированная матрица. Показать, что

$$d(XY) = Xd(Y) + d(X)Y.$$

8. Получить представление для $d(X_1, X_2, \dots, X_N)$ и тем самым для $d(X^N)$.

9. Пусть $d_A(X) = AX - XA$, $d_B(X) = BX - XB$. Показать, что

$$d_A(d_B(X)) = d_B(d_A(X)).$$

10. Когда уравнение $dX = \lambda X$, где λ — скаляр, имеет решение и каков вид этого решения? Рассмотрение этих вопросов читатель найдет в книге И. А. Лаппо-Данилевского¹⁾.

11. Когда уравнения

$$d(X_1) = a_{11}X_1 + a_{12}X_2,$$

$$d(X_2) = a_{21}X_1 + a_{22}X_2$$

имеют решения и каковы эти решения?

12. Поскольку $d(X)$ является аналогом производной, то какой вид имеет аналог ряда Тейлора?

13. Пусть дана действительная матрица A . Всегда ли существует диагональная матрица B с элементами, равными ± 1 , такая, что решения уравнения $B dx/dt = Ax$ стремятся к нулю при $t \rightarrow \infty$? (Брок.)

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 1. Мы используем здесь термин «устойчивость», понимая под этим следующее: все решения уравнения $dx/dt = Ax$ стремятся к нулю при $t \rightarrow \infty$. Более широкое определение и дальнейшие результаты читатель найдет в книге

Беллман (R. Bellman), *Stability Theory of Differential Equations*, McGraw-Hill Book Company, Inc., New York, 1953. [Русский перевод: Теория устойчивости решений дифференциальных уравнений, ИЛ, 1954.]

Задача исследования асимптотического поведения решения дифференциального уравнения $dx/dt = A(t)x$ с помощью собственных чисел матрицы $A(t)$ как функций времени t значительно сложнее. См. цитированную выше книгу, а также

Коддингтон и Левинсон (E. A. Coddington and N. Levinson), *Theory of Ordinary Differential Equations*, McGraw-Hill Book Company, Inc., New York, 1955. [Русский перевод: Теория обыкновенных дифференциальных уравнений, ИЛ, 1958.]

¹⁾ И. А. Лаппо-Данилевский, *Применение функций от матриц к теории линейных систем обыкновенных дифференциальных уравнений*, Гостехиздат, 1957.

§ 2. Обобщением этого результата является теорема Пуанкаре—Ляпунова, которая утверждает, что все решения уравнения $dx/dt = Ax + g(x)$, для которых $\|x(0)\|$ достаточно мала, стремятся к нулю при $t \rightarrow \infty$ при условии, что A —устойчивая матрица, а нелинейная функция $g(x)$ такая, что $\|g(x)\|/\|x\| \rightarrow 0$ при $\|x\| \rightarrow 0$. См. по этому поводу цитированные выше книги. Одно из доказательств этого результата намечено в упражнении 1 к § 4.

§ 4. Этот метод играет фундаментальную роль в современной теории устойчивости функциональных уравнений. Впервые этот метод был опубликован в мемуаре Ляпунова в 1897 г. См. книгу

А. М. Ляпунов, Общая задача об устойчивости движения, Гостехиздат, 1950.

Имеется обзор недавних результатов:

Антосевич (H. Antosiewicz), A Survey of Lyapunov's Second Method, Applied Mathematics Division, National Bureau of Standards, 1956.

См. также

Антосевич (H. Antosiewicz), Some Applications of Lyapunov's Conditions for Stability, J. Rat. Mech. Analysis 3 (1954), 447—457.

§ 5. Здесь мы следуем работам:

Беллман (R. Bellman), Notes on Matrix Theory, X: A. Problem in Control, Quart. Appl. Math. 14 (1957), 417—419;

Анке (K. Anke), Eine neue Berechnungsmethode der quadratischen Regelfläche, Z. angew. Math. u. Phys. 6 (1955), 327—331;

Букнер (H. Buckner), A Formula for an Integral Occuring in the Theory of Linear Servomechanisms and Control Systems, Quart. Appl. Math. 10 (1952);

Хазенбрук и Ван дер Варден (P. Hazenbroeck and B. L. Van der Waerden), Theoretical Considerations in the Optimum Adjustment of Regulators, Trans. Amer. Soc. Mech. Engrs. 72 (1950).

§ 6. Гурвиц получил эти результаты, исследуя задачу устойчивости многочлена. См. работу

Гурвиц (A. Hurwitz), Über die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt, Math. Ann. 46 (1895) (Werke. vol. 2, 533—545).

С тех пор этот результат был доказан многими различными способами. Обзор критериев устойчивости, новых результатов, а также обширную библиографию можно найти в работе

Кремер и Эфферц (H. Cremer and F. H. Effertz), Über die algebraische Kriterien für die Stabilität von Regelungssystemen, Math. Ann. 137 (1959), 328—350.

§ 7. Обширные исследования этого вопроса провел Басс. Результаты его не опубликованы. Рассмотрение вопросов устойчи-

ности в том виде, в каком они возникают в математической экономике, имеются в работах

Энтховен и Эрроу (A. C. Enthoven and K. J. Arrow), A Theorem on Expectations and the Stability of Equilibrium, *Econometrica* 24 (1956), 288—293;

Эрроу и Нерлав (K. J. Arrow and M. Nerlove), A Note on Expectations and Stability, *Econometrica* 26 (1958), 297—305.
где также имеется ссылка на более раннюю работу Мецлера.

§ 9. См. статью

Шварц (H. R. Schwarz), Ein Verfahren zur Stabilitätsfrage bei Matrizen-Eigenwert-Problemen, *Z. angew. Math. u. Phys.* 7 (1953), 473—500.

Задача приведения данной матрицы к якобневой рассмотрена в цитированной выше работе Шварца и в работе

Гивенс (W. Givens), Numerical Computation of the Characteristic values of a Real Symmetric Matrix, Oak Ridge National Laboratory, Report ORNL-1574, 1954.

Ряд других тесно связанных с этим результатом читатель найдет в статьях

Франк (E. Frank), On the Zeros of Polynomials with Complex Coefficients, *Bull. Amer. Math. Soc.* 52 (1946), 144—157;

Франк (E. Frank), The Location of the Zeros of Polynomials with Complex Coefficients, *Bull. Amer. Math. Soc.* 52 (1946), 890—898.

В работах

Таусски (O. Taussky), Analytical Methods in Hypercomplex Systems, *Comp. Math.*, 1937;

Таусски и Тодд (O. Taussky and J. Todd), Infinite Powers of Matrix, *J. London, Math. Soc.*, 1943

рассматриваются матрицы, собственные значения которых лежат внутри единичного круга.

Мы совершенно не затронули здесь интересного и важного вопроса об асимптотическом поведении решения векторно-матричного уравнения $x(n+1) = (A + B(n))x(n)$ при $n \rightarrow \infty$, где элементы матрицы $B(n)$ стремятся к нулю при $n \rightarrow \infty$. Изучение этого вопроса было начато Пуанкаре, см.

Пуанкаре (H. Poincaré), Sur les équations lineaires aux différentielles ordinaires et aux différences finies, *Amer. J. Math.* 7 (1885), 203—258,
и продолжено Перроном:

Перрон (O. Perron), Über die Poincarésche lineare Differenzgleichung, *J. reine angew. Math.* 137 (1910), 6—64

и его учеником Та Ли:

Та Ли (Ta Li), Die Stabilitätsfrage bei Differenzgleichungen, *Acta Math.* 63 (1934), 99—141.

Последние результаты в этой области излагаются в работе

Г. А. Фрайман, О теоремах Пуанкаре и Перрона, *УМН* 12, № 3 (75) (1957), 241—246.

См. также

Беллман (R. Bellman), A Survey of the Theory of the Boundedness, Stability and Asymptotic Behavior of Solutions of Linear and Non-linear Differential and Difference Equations, Office of Naval Research, Department of the Navy, January 1949.

Эти вопросы, так же как и соответствующие вопросы для дифференциальных уравнений, больше относятся к теории линейных дифференциальных и разностных уравнений, чем к теории матриц как таковой.

Рассмотрение вопросов, связанных с устойчивостью линейных дифференциальных уравнений со случайными коэффициентами, имеется в работах

Сефл (O. Seffl), On Stability of a Randomized Linear Systems, Sci. Sinica 7 (1958), 1027—1034;

Самуэлс (J. C. Samuels) On the Mean Square Stability of Random Linear Systems, IRE Trans. on Circuit Theory, May 1959, Special Supplement (Transactions of the 1959 International Symposium on Circuit and Information Theory), 248—259.

где имеются ссылки на более раннюю статью Розенблюма и другие работы.

МАРКОВСКИЕ МАТРИЦЫ
И ТЕОРИЯ ВЕРОЯТНОСТЕЙ

§ 1. Введение. В этой и следующих главах мы будем изучать класс матриц, возникающих с связи с некоторыми фундаментальными вопросами теории вероятностей и математической экономики. Методы, которые мы будем здесь применять, совершенно отличны от используемых при изучении квадратичных форм и дифференциальных уравнений.

Основным понятием теперь будет неотрицательность — неотрицательные матрицы и неотрицательные векторы. Матрицу $M = \|m_{ij}\|$ называют *неотрицательной*, если $m_{ij} \geq 0$ для всех i и j . Аналогично говорят, что вектор x *неотрицателен*, если все его компоненты неотрицательны, $x_i \geq 0$.

Неотрицательные матрицы, для которых выполняется более сильное условие $m_{ij} > 0$, называются *положительными*. Эти матрицы обладают особенно интересными и изящными свойствами.

Термины «положительный» и «неотрицательный» часто используются для обозначения того, что мы ранее называли «положительно определенный» и «неотрицательно определенный». Поскольку в последующем эти два типа матриц вместе появляться не будут, путаницы возникнуть не должно. Прилагательное «положительный» представляет собой настолько распространенный и важный термин, что нет нужды пояснять, почему оно столь часто употребляется в разных аспектах.

В этой книге мы ограничимся рассмотрением лишь элементарных свойств неотрицательных матриц. Детальное изучение вопросов, связанных с теорией вероятностей и математической экономикой, можно найти в литературе, приведенной в конце главы.

Эта и следующая за ней главы посвящены стохастическим процессам; в последующей главе рассматриваются различные аспекты теории матриц и математической экономики.

§ 2. Простой стохастический процесс. Для того чтобы подготовить почву для рассмотрения марковских матриц, остановимся на некоторых основных моментах, связанных с понятием детерминированного процесса. Рассмотрим систему S , состояние которой изменяется таким образом, что в любой момент t оно

может быть описано конечномерным вектором x . Предположим далее, что состояние системы в любой момент $s+t$, $s > 0$, может быть представлено в виде некоторой функции следующей пары аргументов: состояния в момент t и текущего времени s , именно

$$x(s+t) = g(x(t), s). \quad (1)$$

При естественных допущениях относительно гладкости функции $g(t)$ мы можем получить, разлагая обе части (1) по степеням s , дифференциальное уравнение для $x(t)$

$$\frac{dx}{dt} = h(x(t)). \quad (2)$$

Конечномерные детерминированные системы рассматриваемого типа эквивалентны, таким образом, системам, описываемым обыкновенными дифференциальными уравнениями. Если мы введем функции более сложного типа, зависящие от предыстории процесса, то придем к более сложным функциональным уравнениям.

Предположение о том, что настоящее состояние системы полностью определяет ее будущие состояния, является чрезвычайно ограничительным. Ясно, что в любом реальном процессе оно в той или иной степени не выполняется. Из-за простоты лежащих в ее основе понятий мы все же принимаем детерминированную модель в тех случаях, когда она позволяет получить результаты, согласующиеся с экспериментом. В настоящее время многие наблюдаемые явления уже не могут быть объяснены с позиции теории детерминированных процессов; это обстоятельство вынуждает нас строить другие типы математических моделей.

Рассмотрим теперь следующий стохастический процесс.

Для упрощения формулировки и во избежание ряда затруднений мы предположим, что рассматриваемая физическая система S может находиться в одном из состояний, число которых конечно, и переходить из одного состояния в другое только в дискретные моменты времени.

Обозначим состояния целыми числами $1, 2, \dots, N$, а моменты времени — через $t=0, 1, 2, \dots$. Элемент случайности мы введем, предположив, что существует фиксированная вероятность того, что система, находящаяся в состоянии j в момент t , перейдет в состояние i в момент $t+1$. Это, конечно, также является очень сильным требованием регулярности. На математическом языке случайный процесс обозначает отнюдь не то, что мы обычно подразумеваем под случайным процессом в обыденной речи.

В соответствии со сказанным введем матрицу переходов $M = \|m_{ij}\|$, где m_{ij} — вероятность того, что система, находящаяся в момент t в состоянии j , перейдет в момент $t+1$ в состояние i .

Отметим, что мы рассматриваем матрицу переходов M , не зависящую от времени. Это наиболее важный и интересный случай.

Введя, таким образом, матрицу M , мы, очевидно, должны наложить следующие условия:

$$m_{ij} \geq 0, \quad (3a)$$

$$\sum_{i=1}^N m_{ij} = 1, \quad j = 1, 2, \dots, N. \quad (3b)$$

Неотрицательность элементов m_{ij} необходима для того, чтобы их можно было рассматривать как вероятности. Второе условие отражает тот факт, что частица, находящаяся в момент t в состоянии j , должна перейти в одно из допустимых состояний в момент $t + 1$.

§ 3. Марковские матрицы и вероятностные векторы. Введем теперь некоторые обозначения. Матрицу M , элементы которой удовлетворяют условиям (2.3a) и (2.3b), будем называть *марковской матрицей*.

Вектор x с компонентами x_i , удовлетворяющими условиям

$$x_i \geq 0, \quad (1a)$$

$$\sum_{i=1}^N x_i = 1, \quad (1b)$$

будем называть *вероятностным* вектором. Вообще, вектор, все компоненты которого положительны, мы будем называть *положительным* вектором.

Упражнения

1. Доказать, что при $0 < \lambda < 1$ матрица $\lambda P + (1 - \lambda) Q$ является марковской всякий раз, когда матрицы P и Q марковские.

2. Доказать, что при тех же условиях матрица PQ марковская.

3. Доказать, что если x — вероятностный вектор, а M — марковская матрица, то вектор Mx вероятностный.

4. Доказать, что марковскую матрицу можно определить следующим образом: матрица M является марковской тогда и только тогда, когда при любом вероятностном векторе x вектор Mx также является вероятностным.

5. Доказать, что если матрица M' является транспонированной к марковской матрице M и $M' = \|a_{ij}\|$, то при $i = 1, 2, \dots, N$ имеем $\sum_{j=1}^N a_{ij} = 1$.

6. Доказать, что положительная марковская матрица преобразует неотрицательный вектор с неотрицательными компонентами в вектор с положительными компонентами.

§ 4. Аналитическое описание дискретных марковских процессов. Стохастический процесс, описанный в § 2, обычно называют *дискретным марковским* процессом. Посмотрим теперь, как этот процесс можно описать аналитически.

Поскольку состояние системы S в любой момент времени t представляет собой случайную величину, принимающую любое из значений $i = 1, 2, \dots, N$, мы введем N функций времени t следующим образом:

$x_i(t)$ — вероятность того, что система находится в состоянии i в момент t . (1)

При $t = 0$ имеем

$$x_i(0) = \delta_{ik}, \quad (2)$$

где k — начальное состояние системы S .

Тогда должны выполняться следующие соотношения:

$$x_i(t+1) = \sum_{j=1}^N m_{ij} x_j(t), \quad i = 1, 2, \dots, N. \quad (3)$$

Один из аспектов задачи предсказания поведения системы S состоит в изучении поведения решений системы уравнений (3), которую мы можем записать более компактно в виде

$$\mathbf{x}(t+1) = M\mathbf{x}(t), \quad i = 1, 2, \dots, N. \quad (4)$$

В целом эта задача чрезвычайно сложна и интересна и сопровождается многочисленными аналитическими, алгебраическими, топологическими и физическими «обертонами». На эту тему был написан ряд монографий, и все же здесь, по-видимому, еще непочатый край работы.

Особенно простые и элегантные результаты могут быть получены в случае, когда все элементы матрицы M положительны. Поэтому мы в основном и ограничимся рассмотрением этого случая, воспользовавшись несколькими методами, и лишь слегка коснемся общего случая (неотрицательная матрица).

Упражнения

1. Доказать, что $\mathbf{x}(t) = M^t \mathbf{x}(0)$; $\mathbf{x}(t+s) = M^t \mathbf{x}(s)$.

§ 5. Асимптотическое поведение. Следующие несколько параграфов мы посвятим одному замечательному результату, приведя два его доказательства.

Теорема 1. Если M — положительная марковская матрица и $\mathbf{x}(t)$ удовлетворяет условию (4.4), то

$$\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{y}, \text{ где } \mathbf{y} \text{ — вероятностный вектор;} \quad (1a)$$

$$\mathbf{y} \text{ не зависит от } \mathbf{x}(0); \quad (1b)$$

\mathbf{y} является собственным вектором матрицы M , принадлежащим характеристическому числу 1. (1в)

Тот факт, что $x(t)$ сходится к фиксированному вероятностному вектору y , очень интересен и, возможно, не является неожиданным, если учесть «перемешивающее» свойство системы, которое вытекает из условия $m_{ij} > 0$. Независимость же этого предела от начального состояния поистине удивительна.

§ 6. Первое доказательство. В этой главе мы дадим два доказательства этой теоремы. Третье доказательство можно получить как следствие из результатов, касающихся общих положительных матриц, которые будут выведены в гл. 16. Первое доказательство (мы приводим его первым из-за его простоты) иллюстрирует то обстоятельство, которое мы подчеркивали ранее, а именно полезность рассмотрения линейных преобразований совместно с их сопряженными.

Поскольку t принимает только дискретные значения $0, 1, 2, \dots$, заменим его на n и будем рассматривать $x(n)$ вместо $x(t)$. Чтобы установить наличие предела при $n \rightarrow \infty$ у $x(n)$, рассмотрим скалярное произведение $x(n)$ на произвольный вектор b . Так как в силу (4.4) $x(n) = M^n x(0)$, то, полагая $x(0) = c$, имеем

$$(x(n), b) = (M^n c, b) = (c, (M^n)' b) = (c, (M')^n b), \quad (1)$$

где матрица M' является транспонированной к M .

Если мы покажем, что $(M')^n b$ при $n \rightarrow \infty$ сходится при любом векторе b , то тем самым мы покажем, что при $n \rightarrow \infty$ сходится $x(n)$, так как можно сначала взять вектор b со всеми нулевыми компонентами, кроме первой, затем кроме второй и т. д.

Введем вектор

$$z(n) = (M')^n b, \quad (2)$$

удовлетворяющий разностному уравнению

$$z(n+1) = M' z(n), \quad z(0) = b. \quad (3)$$

Пусть $u(n)$ есть наибольшая компонента вектора $z(n)$ при каждом n , а $v(n)$ — наименьшая. Покажем, что из свойств матрицы M' вытекает, что $u(n) - v(n) \rightarrow 0$ при $n \rightarrow \infty$. Так как

$$z_i(n+1) = \sum_{j=1}^N m_{ji} z_j(n) \quad (4)$$

и

$$\sum_{j=1}^N m_{ij} = 1, \quad m_{ij} \geq 0, \quad (4')$$

то

$$\begin{aligned} u(n+1) &\leq u(n), \\ v(n+1) &\geq v(n). \end{aligned} \quad (5)$$

Поскольку $\{u(n)\}$ — монотонно убывающая последовательность, ограниченная снизу нулем, а $\{v(n)\}$ — монотонно возрастающая последовательность, ограниченная сверху единицей, эти последовательности сходятся. Пусть

$$u(n) \rightarrow u, \quad v(n) \rightarrow v. \quad (6)$$

Чтобы показать, что $u = v$, поступим следующим образом.

Используя формулы (4) и (4'), получаем неравенства

$$\begin{aligned} u(n+1) &\leq (1-d)u(n) + dv(n), \\ v(n+1) &\geq (1-d)v(n) + du(n), \end{aligned} \quad (7)$$

где d — положительная нижняя граница элементов m_{ij} , $i, j = 1, 2, \dots, N^*$). Из (7) следует оценка

$$u(n+1) - v(n+1) \leq [(1-d)u(n) + dv(n) - (1-d)v(n) - du(n)] \leq (1-2d)(u(n) - v(n)), \quad (8)$$

откуда имеем

$$u(n) - v(n) < (1-2d)^n (u(0) - v(0)). \quad (9)$$

Следовательно, так как $d \leq 1/2$, если $N \geq 2$, то $u(n) - v(n) \rightarrow 0$ при $n \rightarrow \infty$.

Таким образом, $z(n)$ не только сходится при $n \rightarrow \infty$, но и все координаты предельного вектора равны между собой. Из сходимости $z(n)$ немедленно вытекает сходимость $x(n)$. Из равенства компонент $\lim z(n)$ вытекает независимость предельного вектора от начального состояния.

Пусть $\lim_{n \rightarrow \infty} z(n) = z$, а $\lim_{n \rightarrow \infty} x(n) = y$. Как мы только что убедились, все компоненты вектора z равны между собой. Пусть, например, они равны a_1 . Тогда

$$(y, b) = \lim_{n \rightarrow \infty} (x(n), b) = (c, z) = a_1 [c_1 + c_2 + \dots + c_n] = a_1, \quad (10)$$

где величина константы a_1 зависит только от b . Следовательно, y не зависит от c .

Легко показать, что y является собственным вектором матрицы M , соответствующим характеристическому числу 1. В самом деле (**),

$$y = \lim_{n \rightarrow \infty} M^{n+1} c = M \lim_{n \rightarrow \infty} M^n c = My. \quad (11)$$

*) Следует взять то из равенств (4), в левой части которого стоит $u(n+1)$. В правой части его все компоненты вектора $z(n)$, кроме $v(n)$, заменить на $u(n)$, после чего с учетом (4') первое неравенство (7) следует немедленно. Аналогично доказывается второе неравенство. (Прим. ред.)

***) Поскольку y — вероятностный вектор $\left(\sum_{i=1}^n y_i = 1, y_i \geq 0 \right)$, то он отличен от нуля. (Прим. ред.)

Другими словами, y есть «неподвижная точка» преобразования, осуществляемого матрицей M .

Упражнение

1. Используя тот факт, что I — неотрицательная марковская матрица, показать, что условия положительности не могут быть полностью устранены.

§ 7. Второе доказательство независимости от начального состояния. Другой метод доказательства независимости от начального состояния состоит в следующем. Как мы уже знаем, $\lim_{n \rightarrow \infty} M^n c$ существует при любом начальном вероятностном век-

торе c . Пусть y и z — два предельных вектора, соответствующих различным начальным вероятностным векторам c и d . Выберем скаляр t_1 так, чтобы вектор $y - t_1 z$ имел хотя бы одну нулевую компоненту, а все остальные компоненты положительные^{*)}. В этом случае вектор $M(y - t_1 z)$ является положительным, если вектор $y - t_1 z$ нетривиальный. Вместе с тем

$$M(y - t_1 z) = My - t_1 Mz = y - t_1 z. \quad (1)$$

Это равенство содержит противоречие, если только не выполняется условие $y - t_1 z = 0$. Последнее же означает, что $t_1 = 1$, так как y и z — вероятностные векторы.

§ 8. Некоторые свойства положительных марковских матриц. Опираясь на результаты и используя методы предыдущих пунктов, мы установим некоторые интересные свойства положительных марковских матриц.

Отметим прежде всего, что в предыдущем параграфе мы показали, что у положительной марковской матрицы не может быть двух линейно независимых положительных собственных векторов, принадлежащих характеристическому числу 1.

Более того, с помощью тех же самых рассуждений легко показать, что положительная марковская матрица не может иметь собственного вектора^{**)}, принадлежащего характеристическому числу 1 и не пропорционального вероятностному вектору y , полученному выше. В самом деле, пусть z будет таким вектором¹⁾. Тогда при достаточно большом t_1 вектор $z + t_1 y$ станет положительным. Разделив каждую компоненту на скаляр $t_2 = (z + t_1 y, c)$, где c — вектор, все компоненты которого равны единице, мы получим вероятностный вектор. Следовательно, $(z + t_1 y)/t_2 = y$, а это означает линейную зависимость векторов z и y .

^{*)} Существование такого значения t_1 легко показать, введя вектор $y - t_2 z$ и увеличивая параметр t от нуля. (Прим. ред.)

^{**)} Не обязательно положительного. (Прим. ред.)

¹⁾ Легко видеть, что для доказательства достаточно взять вектор z действительным.

Нетрудно показать, что никакое характеристическое число матрицы M по абсолютной величине не может превосходить единицы. Действительно, пусть x — собственный вектор матрицы M' , λ — характеристическое число, а m — модуль наибольшей по величине компоненты вектора x . Тогда из соотношения $\lambda x = M'x$ следует неравенство

$$|\lambda| m \leq m \sum_{j=1}^N m_{ji} = m, \quad (1)$$

откуда $|\lambda| \leq 1$.

Можно легко показать, что если M — положительная марковская матрица, то $\lambda = 1$ является единственным характеристическим числом, модуль которого равен единице. В самом деле, пусть μ — другое такое характеристическое число, а $w + iz$ — собственный вектор, принадлежащий этому характеристическому числу, где w и z — действительные векторы. Выберем t_1 настолько большим, чтобы векторы $w + t_1 y$ и $z + t_1 y$ стали заведомо положительными. Тогда

$$M(w + iz + t_1(1+i)y) = \mu(w + iz) + t_1(1+i)y. \quad (2)$$

Таким образом, с одной стороны, мы имеем

$$M^n(w + iz + t_1(1+i)y) = \mu^n(w + iz) + t_1(1+i)y. \quad (3)$$

С другой стороны, при $n \rightarrow \infty$ последовательность

$$M^n(w + iz + t_1(1+i)y) = M^n(w + t_1 y + i(z + t_1 y)) \quad (4)$$

в силу положительности векторов $w + t_1 y$ и $z + t_1 y$ сходится к вектору y , умноженному на некоторый скалярный множитель.

Если же по предположению абсолютная величина μ равна единице, то вектор $\mu^n(w + iz)$ сходится при $n \rightarrow \infty$ только тогда, когда $\mu = 1$. Этим завершается доказательство.

Таким образом, нами доказана

Теорема 2. Если M — положительная марковская матрица, то характеристическое число с наибольшей абсолютной величиной есть 1. Любой собственный вектор, принадлежащий этому характеристическому числу, с точностью до скалярного множителя совпадает с вероятностным вектором.

Упражнение

1. Если λ — характеристическое число, а принадлежащий ему собственный вектор положителен, то $\lambda = 1$.

§ 9. Второе доказательство сходимости. Сейчас мы дадим другое прямое доказательство того факта, что $x(n)$ имеет предел при $n \rightarrow \infty$. Выпишем уравнения, связывающие компоненты

векторов $\mathbf{x}(n)$ и $\mathbf{x}(n+1)$:

$$x_i(n+1) = \sum_{j=1}^N m_{ij} x_j(n). \quad (1)$$

Отсюда имеем

$$x_i(n+1) - x_i(n) = \sum_{j=1}^N m_{ij} [x_j(n) - x_j(n-1)]. \quad (2)$$

Поскольку $\sum_{j=1}^N x_j(n) = 1$ при всех n , то неравенства $x_j(n) \geq x_j(n-1)$ не могут выполняться для всех j , если только они не сводятся к равенствам. В этом последнем случае $\mathbf{x}(n) = \mathbf{x}(n-1)$, откуда $\mathbf{x}(m) = \mathbf{x}(n-1)$ при $m \geq n$, и, следовательно, доказываемая сходимость установлена.

В общем случае обозначим через $S(n)$ множество тех j , для которых $x_j(n) \geq x_j(n-1)$, а через $T(n)$ — множество таких j , для которых $x_j(n) < x_j(n-1)$. Из предыдущего ясно, что в $S(n)$ и $T(n)$ при каждом n входит по крайней мере один элемент.

На основании (2) легко заключить, что

$$\begin{aligned} \sum_{i \in T(n)} m_{ij} [x_j(n) - x_j(n-1)] &\leq x_i(n+1) - x_i(n) \leq \\ &\leq \sum_{i \in S(n)} m_{ij} [x_j(n) - x_j(n-1)]. \end{aligned} \quad (3)$$

Следовательно,

$$\sum_{i \in S(n+1)} [x_i(n+1) - x_i(n)] \leq \sum_{i \in S(n)} \left\{ \sum_{i \in S(n+1)} m_{ij} \right\} [x_j(n) - x_j(n-1)], \quad (4)$$

откуда в силу того, что $m_{ij} \geq d > 0$ при всех i и j , следует неравенство

$$\sum_{i \in S(n+1)} [x_i(n+1) - x_i(n)] \leq (1-d) \sum_{i \in S(n)} [x_j(n) - x_j(n-1)]. \quad (5)$$

Точно так же, суммируя по $i \in T(n+1)$, имеем

$$(1-d) \sum_{i \in T(n)} [x_j(n) - x_j(n-1)] \leq \sum_{i \in T(n+1)} [x_i(n+1) - x_i(n)]. \quad (6)$$

Сравнивая (5) и (6), получаем

$$\sum_{i=1}^N |x_j(n+1) - x_j(n)| \leq (1-d) \sum_{i=1}^N |x_j(n) - x_j(n-1)|, \quad (7)$$

откуда следует, что ряд $\sum_{n=1}^{\infty} \|\mathbf{x}(n) - \mathbf{x}(n-1)\|$ сходится.

Этим завершается доказательство сходимости $\mathbf{x}(n)$.

Упражнения

1. Показать, что если M — марковская матрица, а M^2 — положительная марковская матрица, то последовательности $\{M^{2n}c\}$ и $\{M^{2n+1}c\}$ сходятся. Сходится ли последовательность $\{M^n c\}$?

2. Чему равен предел M^n , если M — положительная марковская матрица?

§ 10. Марковские матрицы общего вида. Как уже отмечалось выше, общая теория марковских матриц очень сложна; ее изучение лучше всего проводить в рамках теории вероятностей.

Приведем простой пример, который покажет нам, как могут возникать характеристические числа с абсолютной величиной, равной единице. Рассмотрим систему, имеющую только два состояния, 1 и 2, причем состояние 1 всегда переходит в состояние 2, а состояние 2 — в состояние 1. В этом случае матрица переходов имеет вид

$$M = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}. \quad (1)$$

Матрица M имеет своими характеристическими числами $\lambda_1 = 1$ и $\lambda_2 = -1$.

Рассматривая циклические случаи более высокого порядка, мы можем получить характеристические числа, представляющие собой корни из единицы произвольной степени.

Остановимся все же, хотя и вкратце, на одном важном результате, который утверждает, что некоторый усредненный предел существует и в том случае, когда отсутствует предел в доказанном выше смысле.

Предположим, что M имеет только простые характеристические числа; тогда, записав M в виде

$$M = T \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \\ 0 & & & & \end{vmatrix} T^{-1}, \quad (2)$$

получим

$$\left(\sum_{k=1}^n M^k \right) / n = T \begin{vmatrix} \sum_{k=1}^n \lambda_1^k / n & & & 0 \\ & \sum_{k=1}^n \lambda_2^k / n & & \\ & & \ddots & \\ & & & \sum_{k=1}^n \lambda_N^k / n \\ 0 & & & & \end{vmatrix} T^{-1}. \quad (3)$$

Если $|\lambda_j| < 1$, то $\sum_{k=1}^n \lambda_j^k / n \rightarrow 0$ при $n \rightarrow \infty$. Если же $|\lambda_j| = 1$, то предел

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n \lambda_j^k / n = a_j; \quad (4)$$

при этом $a_j = 0$ при $\lambda_j \neq 1$ и $a_j = 1$ при $\lambda_j = 1$.

При всех случаях, если предположить, что M имеет только простые характеристические числа, можно утверждать, что существует квазипредельное распределение

$$y = \lim_{n \rightarrow \infty} \left(\frac{x + Mx + \dots + M^n x}{n+1} \right), \quad (5)$$

являющееся собственным вектором матрицы M ; точнее, $My = y$. Предельный вектор y является вероятностным вектором, если x — вероятностный вектор.

Рассмотрение общего случая более сложно, но проводится точно таким же образом (см. упражнения в конце этой главы).

Упражнение

1. Всякое ли характеристическое число марковской матрицы M такое, что $|\lambda| = 1$, представимо в виде $\lambda = \sqrt[n]{\Gamma}$?

§ 11. Непрерывный стохастический процесс. Обратимся теперь к стохастическому процессу, описанному в § 2, и рассмотрим его в предположении, что система наблюдается непрерывно. Начнем с дискретного процесса, в котором наблюдения производятся в моменты времени $t = 0, \Delta, 2\Delta, \dots$. Чтобы придать смысл непрерывному процессу, мы примем гипотезу непрерывности, состоящую в следующем: вероятность того, что система не изменит своего состояния в течение интервала времени Δ , равна $1 - o(\Delta)$. Чтобы сделать это утверждение строгим, введем следующие величины:

$a_{ij}\Delta$ — вероятность того, что система S находится в состоянии i в момент $t + \Delta$, при условии что в момент t она находится в состоянии j , $i \neq j$. (1)

$1 - a_{ii}\Delta$ — вероятность того, что система S находится в состоянии i в момент $t + \Delta$, при условии что в момент t она находится в состоянии i . (2)

Предполагается, что величины a_{ij} удовлетворяют условиям

$$a_{ij} \geq 0. \quad (2a)$$

$$a_{ii} = \sum_{j \neq i} a_{ji}. \quad (26)$$

Тогда уравнения, описывающие поведение системы S , таковы:

$$x_i(t + \Delta) = (1 - a_{ii}\Delta) x_i(t) + \Delta \sum_{j \neq i} a_{ij} x_j(t), \quad i = 1, 2, \dots, N. \quad (3)$$

Чисто формально используя соотношение

$$x_i(t + \Delta) = x_i(t) + \Delta x'_i(t) + o(\Delta^2), \quad (4)$$

устремим в (3) Δ к нулю. В результате мы придем к следующей системе дифференциальных уравнений:

$$\frac{dx_i}{dt} = -a_{ii}x_i + \sum_{j \neq i} a_{ij}x_j, \quad x_i(0) = c_i, \quad (5)$$

$i = 1, 2, \dots, N$, где c_i , $i = 1, 2, \dots, N$, — начальные вероятности.

Все сложные проблемы, связанные с определением непрерывного стохастического процесса, мы оставим в стороне и будем рассматривать процесс в терминах полученной системы дифференциальных уравнений. Для того чтобы принятое определение было корректным, нужно показать, что функции, порождаемые системой (5), ведут себя как вероятности. Этим мы займемся ниже.

Наше рассмотрение связано с рядом интересных вопросов, которые здесь изучаться не будут. Вот некоторые из них.

1. Как непосредственно определить непрерывный стохастический процесс и вывести дифференциальные уравнения (5)?

2. В каком смысле непрерывный стохастический процесс, определенный системой (5), можно рассматривать как предел дискретного стохастического процесса, определенного системой (3)?

Поскольку эти вопросы скорее относятся к теории вероятностей, мы лишь ограничимся их упоминанием и сосредоточим свое внимание на матричных аспектах системы (5). Тем не менее мы, как обычно, не будем стесняться использовать вероятностные соображения.

§ 12. Доказательство вероятностных свойств. Покажем теперь, что система дифференциальных уравнений

$$\frac{dx_i}{dt} = -a_{ii}x_i + \sum_{j \neq i} a_{ij}x_j, \quad x_i(0) = c_i, \quad i = 1, 2, \dots, N, \quad (1)$$

где

$$c_i \geq 0, \quad \sum_i c_i = 1, \quad (2a)$$

$$a_{ij} \geq 0, \quad (2b)$$

$$a_{ii} = -\sum_{j \neq i} a_{ji}, \quad (2b)$$

порождает набор функций, удовлетворяющих условиям

$$x_i(t) \geq 0, \quad t \geq 0, \quad (3a)$$

$$\sum_i x_i(t) = 1. \quad (36)$$

Докажем сначала свойство (36). Уравнения (1) с учетом свойства (2в) дают

$$\frac{d}{dt} \left(\sum_i x_i \right) = \sum_i \left(-a_{ii}x_i + \sum_{j \neq i} a_{ij}x_j \right) = 0. \quad (4)$$

Следовательно, при всех t

$$\sum_i x_i(t) = \sum_i x_i(0) = 1. \quad (5)$$

Для того чтобы показать, что $x_i \geq 0$, запишем (1) в виде

$$\frac{d}{dt} (x_i e^{a_{ii}t}) = e^{a_{ii}t} \sum_{j \neq i} a_{ij} x_j.$$

Поскольку $x_j \geq 0$ в момент $t=0$, из последнего соотношения следует, что величины $x_i e^{a_{ii}t}$ монотонно возрастают.

Мы могли бы дать другое доказательство, основываясь на том легко устанавливаемом факте, что решения системы (1) являются пределами при Δ , стремящемся к нулю, решений системы (11.3). Следовательно, свойства решений системы (11.3), имеющие место при всех Δ , должны сохраняться и для решений системы (1). Рекомендуем читателю провести строгое доказательство этим путем.

§ 13. Обобщенные вероятности: унитарные преобразования. Любой набор неотрицательных величин $\{x_i\}$, удовлетворяющих условиям

$$x_i \geq 0, \quad (1a)$$

$$\sum_{i=1}^N x_i = 1, \quad (16)$$

можно рассматривать как набор вероятностей, где x_i есть вероятность того, что система S находится в состоянии i .

Преобразованием

$$x'_i = g_i(x_1, x_2, \dots, x_N), \quad i = 1, 2, \dots, N, \quad (2)$$

сохраняющие соотношения (1), можно рассматривать как аналитическое описание физических изменений, происходящих в системе S .

Линейные однородные преобразования

$$x'_i = \sum_{j=1}^N a_{ij} x_j \quad (3)$$

являются простейшими преобразованиями. Как мы знаем, необходимым и достаточным условием того, чтобы свойства (1) сохранялись, является марковость матрицы $A = \|a_{ij}\|$.

Рассмотрим еще квадратичные преобразования. Пусть x — комплексный вектор с компонентами x_i , и пусть

$$x' = Tx, \quad (4)$$

где T — унитарное преобразование. Тогда $(x', \bar{x}') = (x, \bar{x})$.

Следовательно, если мы рассматриваем $|x_1|^2, |x_2|^2, \dots, |x_N|^2$ как набор вероятностей, то и $|x'_1|^2, |x'_2|^2, \dots, |x'_N|^2$ можно считать таковым.

Мы оставляем читателю доказательство соответствующих предельных теорем для последовательности $\{x(n)\}$, рекуррентно определенной соотношением

$$x(n+1) = Tx(n), \quad (5)$$

где матрица T унитарная, а $(x(0), \bar{x}(0)) = 1$.

§ 14. Обобщенные вероятности: матричные преобразования. Обобщим теперь понятие вероятности следующим образом. Если $\{X_i\}$ представляют собой конечную совокупность симметрических матриц, удовлетворяющих условиям

$$X_i \geq 0, \quad i = 1, 2, \dots, N, \quad (1a)$$

$$\sum_{i=1}^N \text{Sp}(X_i) = 1, \quad (16)$$

то мы назовем их набором вероятностей. Условие $X_i \geq 0$ означает здесь неотрицательную определенность матрицы X_i .

Рассмотрим преобразование

$$Y_i = \sum_{j=1}^N A_{ij} X_j A_{ij}, \quad i = 1, 2, \dots, N. \quad (2)$$

Легко установить следующий аналог полученного для марковских матриц результата:

Теорема 3. *Необходимые и достаточные условия того, чтобы преобразование (2) сохраняло свойства (1), даются следующими соотношениями:*

$$\sum_{i=1}^N A_{ij} A_{ij} = I, \quad j = 1, 2, \dots, N. \quad (3)$$

приложений этого результата можно найти в работе Купманса и Бекмана¹⁾.

Некоторые дальнейшие результаты, касающиеся двояко стохастических матриц, имеются в работах Шрайбера²⁾ и Мирского³⁾. Мирский дал простое доказательство сформулированного выше результата Биркгофа, а также указал на интересную связь между двояко стохастическими матрицами и теорией матричных неравенств.

8. Пусть A — марковская матрица второго порядка. Показать, что A является квадратом марковской матрицы тогда и только тогда, когда $a_{11} \geq a_{12}$.

9. Показать, далее, что, если A является квадратом некоторой марковской матрицы, то ее можно представить как 2^n -ю степень марковской матрицы, где $n = 1, 2, \dots$.

10. Используя предыдущие результаты, показать, что если матрица A имеет квадратный корень, являющийся марковской матрицей, то и среди корней любой фиксированной степени из этой матрицы есть корень, также являющийся марковской матрицей⁴⁾.

11. Вывести отсюда, что можно построить семейство марковских матриц $A(t)$, удовлетворяющее условиям $A(s+t) = A(s)A(t)$ при $s, t \geq 0$, $A(0) = I$ и $A(1) = A^0$.

12. Используя этот результат или каким-либо другим способом показать, что $A = e^B$, где B имеет вид

$$B = \begin{vmatrix} -a & b \\ a & -b \end{vmatrix}, \quad a, b \geq 0.$$

13. Если B — матрица типа, описанного в упражнении 12, то мы знаем, что матрица e^B является марковской. Пусть B_1 и B_2 — матрицы такого типа, и пусть $A_1 = e^{B_1}$, $A_2 = e^{B_2}$. Является ли произведение $A_1 A_2$ марковской матрицей этого типа?

14. Пусть A — марковская матрица порядка N , обладающая тем свойством, что среди корней любой фиксированной степени имеется корень, являющийся марковской матрицей. Пусть A_n — некоторый марковский корень степени 2^n из матрицы A , пусть $A(t) = A_{n_1} A_{n_2} \dots$, где $t = 2^{-n_1} + 2^{-n_2} + \dots$, причем сумма конечна. Можно ли определить $A(t)$ для $0 < t < 1$ как предел значений $A(t)$, где t есть сумма указанного выше вида? Если это возможно, то показать далее, что $A(s+t) = A(s)A(t)$ при $0 < s, t < 1$. Положим еще $A(0) = I$. Пусть все перечисленные соотношения выполняются; является ли в таком случае $A(t)$ решением дифференциального уравнения $dX/dt = BX$, $X(0) = I$, где B — матрица типа, описанного в § 12 (см. (26) и (2в))?

15. Пусть $p_i(n)$ — вероятность перехода из состояния a в состояние i за n шагов, а $q_i(n)$ — вероятность перехода из состояния b в состояние i за n шагов. Рассмотрим определитель

$$d_{ij}(n) = \begin{vmatrix} p_i(n) & p_j(n) \\ q_i(n) & q_j(n) \end{vmatrix}.$$

¹⁾ T. C. Koopmans and M. J. Beckman, Assignment Problems and the Location of Economic Activities, *Econometrica* 25 (1957), 53—76.

²⁾ S. Schreiber, On a Result of S. Sherman Concerning Doubly Stochastic Matrices, *Proc. Amer. Math. Soc.* 9 (1953), 350—353.

³⁾ L. Mirsky, Proofs of Two Theorems on Doubly Stochastic Matrices, *Proc. Amer. Math. Soc.* 9 (1958), 371—374.

⁴⁾ В упражнениях 9—13 речь идет о матрицах второго порядка.

(Прим. ред.)

Доказать, что

$$d_{ij}(n) = \sum_{r,s} a_{ir} a_{js} d_{rs}(n-1)$$

(см. § 14 гл. 12) *).

16. Пусть $\varphi_{ij}(n)$ — вероятность перехода двух частиц из различных состояний a и b в состояния i и j за время n без промежуточного попадания обеих частиц в одно и то же состояние. Показать, что

$$\varphi_{ij}(n) = \sum_{r,s} a_{ir} a_{js} \varphi_{rs}(n-1) *).$$

Какова связь между $d_{ij}(n)$ и $\varphi_{ij}(n)$? См. Карлин и Мак-Грегор (S. Karlin and J. MacGregor), Coincidence Probabilities, Stanford University Department of Statistics, Tech. Rept. 8.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§§ 1—12. Результаты, приведенные в этой главе, являются классическими для теории конечных марковских цепей. То же можно сказать и о данных нами доказательствах. Детальное рассмотрение и многочисленные дополнительные библиографические ссылки можно найти в следующих монографиях и работах **):

Феллер (W. Feller), An Introduction to Probability Theory and Applications, John Wiley & Sons, Inc., New York, 1950 [русский перевод: Введение в теорию вероятностей и ее применения, ИЛ, 1951];

Бартлетт (M. S. Bartlett), An Introduction to Stochastic Processes, Cambridge University Press, New York, 1955;

Арли (N. Arley), On the Theory of Stochastic Processes and Their Applications to the Theory of Cosmic Radiation, Copenhagen, 1943;

Ледерман (W. Ledermann), On the Asymptotic Probability Distribution for Certain Markoff Processes, Proc. Cambridge Phil. Soc. 46 (1950), 581—594;

Ледерман и Ройтер (W. Ledermann and G. E. Reuter), Spectral Theory for the Differential Equations of Simple Birth and Death Processes, Phil. Trans. Royal. Soc. London, ser. A 246 (1953—1954), 321—369.

Интересное обсуждение теории матриц и задач случайного блуждания имеется в работе

Кац (M. Kac), Random Walk and the Theory of Brownian Motion, Amer. Math Monthly 54 (1947), 369—391.

§ 7. Перемешивание карт, по замыслу, должно быть марковским процессом, делающим каждое новое распределение карт по игрокам независимым от предыдущего. На самом деле большинство игроков не очень тщательно перемешивают карты, и поэтому, запомнив некоторые моменты, касающиеся предшествующей сдачи, можно получить значительную информацию о настоящей партии.

*) В упражнениях 15 и 16 через a_{ij} обозначена вероятность перехода из состояния j в состояние i . (Прим. ред.)

**) См. также Е. Б. Дынкин, Марковские процессы, Физматгиз, 1963. (Прим. ред.)

Понятие марковских цепей было введено Пуанкаре. См.

Фреше (M. Fréchet), *Traité du calcul des probabilités*, tome 1, fasc. III, 2-e livre, Gauthier-Villrs & Cie, Paris, 1936,

где подробно рассматривается приложение теории матриц к марковским цепям.

§ 13. Обсуждение этих обобщенных вероятностей и их связи с Фейнмановской формулировкой квантовой механики читатель найдет в работе

Монтролл (E. Montroll), *Markoff Chains, Wiener Integrals, and Quantum Theory*, *Comm. Pure Appl. Math.* 5 (1952), 415—453.

§ 14. Это матричное обобщение вероятностей было введено в работе

Беллман (R. Bellman), *On a Generalization of Classical Probability Theory, I: Markoff Chains*, *Proc. Nat. Acad. Sci. U. S. A.* 39 (1953), 1075—1077.

См. также

Беллман (R. Bellman), *On a Generalization of the Stieltjes Integral to Matrix Functions*, *Rend. circ. mat. Palermo*, ser. II 5 (1956), 1—6;

Беллман (R. Bellman), *On Positive Definite Matrices and Stieltjes Integral*, *Rend. circ. mat. Palermo*, ser. II 6 (1957), 254—258.

Другое обобщение марковских матриц имеется в статье

Хейнсворт (E. V. Hainsworth), *Quasi-stochastic matrices*, *Duke Math. J.* 22 (1955), 15—24.

Некоторые недавние результаты, касающиеся марковских цепей и вполне положительных матриц в смысле Ф. Р. Гантамахера и М. Г. Крейна, имеются в работе

Карлин и Мак-Грегор (S. Karlin and J. MacGregor), *Coincidence Probabilities*, *Stanford University Tech. Rept.* 8, 1958.

СЛУЧАЙНЫЕ МАТРИЦЫ

§ 1. Введение. В этой главе мы очень бегло обсудим некоторые вопросы, связанные со случайными матрицами, и то, каким образом матрицы этого типа возникают при изучении дифференциальных и разностных уравнений. Мы увидим также, что при исследовании моментов решений линейных функциональных уравнений указанного типа естественно ввести кронекеровское произведение матриц.

§ 2. Предельное поведение физических систем. Рассмотрим, как обычно, физическую систему S , состояние которой в любой момент времени t описывается вектором $\mathbf{x}(t)$. В предыдущих главах мы рассматривали случай, когда состояние $\mathbf{x}(t + \Delta)$ определяется состоянием $\mathbf{x}(t)$ посредством линейного преобразования

$$\mathbf{x}(t + \Delta) = (I + Z\Delta) \mathbf{x}(t). \quad (1)$$

Предельная форма этого преобразования при бесконечно малом Δ

$$\frac{d\mathbf{x}}{dt} = Z\mathbf{x}. \quad (2)$$

Предположим теперь, что Z является случайной матрицей*). Под этим мы понимаем, что элементами матрицы Z служат случайные величины, распределения которых зависят от t . Поскольку понятия непрерывного стохастического процесса, и в особенности решения стохастического дифференциального уравнения, чрезвычайно сложны, мы ограничимся рассмотрением дискретного процесса.

Возвращаясь к уравнению (1), введем обозначение $\mathbf{x}(k\Delta) = \mathbf{x}_k$. Тогда (1) можно записать в виде

$$\mathbf{x}_{k+1} = (I + Z_k\Delta) \mathbf{x}_k. \quad (3)$$

Поскольку Z —случайная матрица, ее значение в каждый момент времени является функцией времени; этот факт мы отразили, снабдив в уравнении (3) Z индексом k .

*) В подлиннике «stochastic matrix». (Прим. ред.)

Очевидно, что

$$\mathbf{x}_n = \left[\prod_{k=0}^{n-1} (I + Z_k \Delta) \right] \mathbf{x}_0. \quad (4)$$

Если Δ мало, то можно написать

$$\mathbf{x}_n = \left[I + \Delta \sum_{k=0}^{n-1} Z_k + O(\Delta^2) \right] \mathbf{x}_0. \quad (5)$$

Таким образом, с точностью до членов порядка Δ^2 эффект повторных преобразований можно рассматривать как аддитивный. Аддитивные стохастические процессы подробно рассматриваются в теории вероятностей, и мы не будем заниматься ими. Остановимся лишь на предельном поведении случайной последовательности \mathbf{x}_k в случае, когда $Z_k \Delta$ нельзя считать малыми величинами.

Поскольку эта задача чрезвычайно сложна, мы ограничимся рассмотрением вопроса об асимптотическом поведении средних значений k -х степеней компонент вектора \mathbf{x}_n при $n \rightarrow \infty$ для $k = 1, 2, \dots$

§ 3. Средние значения. Для того чтобы пояснить идею и метод, достаточно рассмотреть двумерный случай, для чего запишем уравнение (3) в координатной форме:

$$\begin{aligned} u_{k+1} &= z_{11}u_k + z_{12}v_k, \\ v_{k+1} &= z_{21}u_k + z_{22}v_k; \end{aligned} \quad (1)$$

здесь u_k и v_k — координаты двумерного вектора \mathbf{x}_k . Предположим теперь, что Z_i являются независимыми случайными матрицами, обладающими одинаковым распределением. Значительно труднее случай, когда распределение матрицы Z_k зависит от значений, принимаемых матрицей Z_{k-1} . Этот вопрос будет рассмотрен в другой книге.

В силу независимости матриц Z_k из системы (1) следуют соотношения

$$\begin{aligned} E(u_{k+1}) &= e_{11}E(u_k) + e_{12}E(v_k), \\ E(v_{k+1}) &= e_{21}E(u_k) + e_{22}E(v_k), \end{aligned} \quad (2)$$

где $E(u_k)$ и $E(v_k)$ — средние значения компонент u_k и v_k , а e_{ij} — среднее значение Z_{ij} .

Из системы (2) следует, что

$$E(\mathbf{x}_{k+1}) = E(Z)E(\mathbf{x}_k) = E(Z)^{k+1}\mathbf{x}_0. \quad (3)$$

Это означает, что асимптотическое поведение вектора $E(\mathbf{x}_n)$ определяется характеристическими числами матрицы $E(Z)$.

Упражнение

1. Пусть матрица Z с одинаковой вероятностью принимает одно из двух возможных значений A или B , где A и B — положительные марковские матрицы. Доказать, что $E(x_n) \rightarrow y$, где y — вероятностный вектор, не зависящий от x_0 (x_0 по предположению — вероятностный вектор).

§ 4. Средние значения квадратов. Для определения средних значений квадратов u_n^2 и v_n^2 из (3.1) имеем

$$\begin{aligned} u_{k+1}^2 &= z_{11}^2 u_k^2 + 2z_{11}z_{12} u_k v_k + z_{12}^2 v_k^2, \\ v_{k+1}^2 &= z_{21}^2 u_k^2 + 2z_{21}z_{22} u_k v_k + z_{22}^2 v_k^2, \end{aligned} \quad (1)$$

откуда видно, что для определения $E(u_{k+1}^2)$ и $E(v_{k+1}^2)$ нужно знать величину $E(u_k v_k)$. Из (3.1) следует, что

$$u_{k+1} v_{k+1} = z_{11}z_{21} u_k^2 + (z_{11}z_{22} + z_{12}z_{21}) u_k v_k + z_{12}z_{22} v_k^2. \quad (2)$$

Следовательно,

$$\begin{vmatrix} E(u_{k+1}^2) \\ E(u_{k+1}v_{k+1}) \\ E(v_{k+1}^2) \end{vmatrix} = E(Z^{21}) \begin{vmatrix} E(u_k^2) \\ E(u_k v_k) \\ E(v_k^2) \end{vmatrix},$$

где $Z^{(2)}$ — кронекеровский квадрат матрицы $Z = |z_{ij}|$ (см. определение в § 8 гл. 12).

Этот пример показывает, что кронекеровский квадрат естественным образом возникает при изучении случайных матриц.

Упражнения к гл. 15

1. Показать, что при определении $E(u_k^2)$ возникает матрица $E(Z^{(2)})$.
2. Показать, что кронекеровский квадрат фигурирует в следующей задаче. Пусть требуется найти матрицу Y , обладающую тем свойством, что

$$E(Z'YZ) = \lambda Y.$$

Какие значения принимает λ ?

3. Пусть матрица Z имеет то же распределение, что и в упражнении 1 § 3. Сходятся ли последовательности $\{E(u_n^2)\}$ и $\{E(v_n^2)\}$, и если сходятся, то к каким пределам?

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

Задача исследования различных функций случайных матриц таких, как, например, определители или характеристические числа, чрезвычайно важна для различных разделов математической физики и статистики. В этой главе мы хотели лишь указать на один из аспектов этой задачи, а также на естественную роль кронекеровских произведений.

Обсуждение многих вопросов, связанных со случайными матрицами, имеется в работах

Уилкс (S. S. Wilks), *Mathematical Statistics*, Princeton University Press, Princeton, N. J., 1943;

Андерсон (T. W. Anderson), *The Asymptotic Distribution of Certain Characteristic Roots and Vectors*, Proc. Second Berkeley Symposium on Mathematical Statistics and Probability, 1950;

Андерсон (T. W. Anderson), *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, Inc., New York, 1958 [русский перевод: Введение в многомерный статистический анализ, Физматгиз, 1963],

а ссылки — в статье

Ингам (A. E. Ingham), *An Integral Which Occurs in Statistics*, Proc. Cambridge Phil. Soc. 29 (1933), 271—276.

См. также

Найквист и др. (H. Nyquist et al.), *The Distribution of Random Determinants* Quart. Appl. Math. 12 (1954), 97—104;

Беллман (R. Bellman), *A Note on Mean Value of Random Determinants*, Quart. Appl. Math. 13 (1955), 322—324.

Некоторые задачи математической физики, в которых фигурируют случайные матрицы, читатель найдет в статьях

Вигнер (E. P. Wigner), *Characteristic Vectors of Bordered Matrices with Infinite Dimensions*, Ann. Math. 62 (1955), 548—564;

Дайсон (F. J. Dyson), *Phys. Rev.* (2) 92 (1953), 1331—1338;

Беллман (R. Bellman), *Dynamics of Disordered Linear Chain*, Phys. Rev. (2) 101 (1958), 19.

На роль кронекеровских произведений при исследовании моментов указывалось также в работах

Кернер (E. H. Kerner), *The Band Structure of Mixed Linear Lattices*, Proc. Phys. Soc. 69 (1956), 234—244;

Джун-ичи Хори (Jun-ichi Hori), *On the Vibration of Disordered Linear Lattice, II*, Progr. Theoret. Phys. 18 (1957), 367—374.

В статье

Марадудин и Вайс (A. Maradudin and G. H. Weiss), *The Disordered Lattice Problem, A Review*, J. Soc. Ind. Appl. Math., 6 (1958), 302—319 [русский перевод: Динамическая теория кристаллической решетки в гармоническом приближении, «Мир», 1965]

показано, как применяются матрицы в задаче «о ближайшем соседе». См. также

Нанда (D. N. Nanda), *Distribution of a Root of a Determinantal Equation*, Ann. Math. Stat. 19 (1948), 47—57, 340—350;

Сю (P. L. Hsu), *On the Distribution of Roots of Certain Determinantal Equations*, Ann. Eugenics 3 (1939), 250—258;

Рой (S. N. Roy), *The Individual Sampling...*, Sankhya, 1943;

Муд (A. M. Mood), *On the Distribution of the Characteristic Roots of Normal Second-moment Matrices*, Ann. Math. Stat. 22 (1951), 26—273;

Кендалл (M. G. Kendall), *The Advanced Theory of Statistics*, vol. II, London, 1946;
 Энглман (R. Engelman), *The Eigenvalues of a Randomly Distributed Matrix*, *Nuovo cimento* 10 (1958), 615—621.

Отметим еще, что итерации случайных линейных преобразований играют важную роль в исследовании прохождения волн через случайную среду; по этому поводу см.

Беллман и Калаба (R. Bellman and R. Kalaba), *Invariant Imbedding, Wave Propagation, and WKB Approximation*, *Proc. Nat. Acad. Sci. U.S.A.* 44 (1958), 317—319;

Беллман и Калаба (R. Bellman and R. Kalaba), *Invariant Imbedding and Wave Propagation in Stochastic Media*, *J. Math. and Mech.*, 1959.

ПОЛОЖИТЕЛЬНЫЕ МАТРИЦЫ, ТЕОРЕМА ПЕРРОНА И МАТЕМАТИЧЕСКАЯ ЭКОНОМИКА

§ 1. Введение. В этой заключительной главе мы рассмотрим несколько задач, возникающих в математической экономике. Общими чертами этих задач являются линейность и положительность.

Однако начнем мы с описания некоторых простых ветвящихся процессов, связанных с ростом биологических популяций и такими физическими процессами, как каскады космических лучей и деление ядер.

Основной результат теории положительных матриц, как это ни кажется странным, был получен Перроном в связи с его исследованиями многомерных цепных дробей, введенных Якоби. Фробениус в ряде работ существенно обобщил результат Перрона.

Весьма примечателен тот факт, что результат, полученный в теоретико-числовых исследованиях, занимает теперь центральное место в математической экономике, в частности, на него опирается «вход—выход»-анализ Леонтьева. Этот результат играет также важную роль в теории процессов.

Вопросы, излагаемые в этой главе, являются малой и очень специальной частью современной теории положительных операторов, точно так же как результаты, относящиеся к симметрическим матрицам, являются частным случаем общих теорем теории эрмитовых операторов.

В конце главы мы коснемся основной задачи линейного программирования и укажем на ее связь с теорией игр Бореля и фон Неймана. В заключение мы упомянем некоторые марковские процессы принятия решений, возникающие в теории динамического программирования.

§ 2. Некоторые процессы простого роста. Рассмотрим следующую простую модель роста группы биологических объектов. Пусть имеется N различных типов этих объектов, которым мы присвоим номера $1, 2, \dots, N$, и пусть в моменты времени $t=0, 1, 2, \dots$ каждый из этих типов производит некоторое число потомков каждого из указанных типов.

В частности, можно рассматривать случай только двух типов — нормальный вид и мутантный вид. Другой интересный случай

представляет собой задача определения числа женских особей в различных возрастных группах. По прошествии каждого года женские особи возраста i переходят в группу женских особей возраста $i+1$, производя при этом время от времени женскую особь возраста нуль. Важно попытаться предсказать возрастное распределение на основе математической модели, так как действительные данные обычно очень трудно получить.

Введем величины

a_{ij} — число индивидуумов типа i , порожденных одним индивидуумом типа j ; $i, j = 1, 2, \dots, N$. (1)

Как обычно, мы будем интересоваться в первую очередь процессами роста, механизм которых не изменяется во времени.

Состояние системы в момент времени n определяется N величинами

$x_i(n)$ — число индивидуумов типа i в момент n , (2)

$i = 1, 2, \dots, N$. Два последовательных состояния системы связаны соотношениями

$$x_i(n+1) = \sum_{j=1}^N a_{ij} x_j(n), \quad i = 1, 2, \dots, N. \quad (3)$$

Начальное состояние системы считаем заданным: $x_i(0) = c_i$, $i = 1, 2, \dots, N$.

Задача, которую мы ставим перед собой, состоит в исследовании поведения компонент $x_i(n)$ при $n \rightarrow \infty$. Это поведение определяется характеристическими числами матрицы $A = \|a_{ij}\|$ и, в частности, наибольшим по модулю характеристическим числом.

Как мы уже знаем из рассмотрения марковских матриц, эта задача очень сложна. Можно надеяться, однако, что в предположении положительности элементов a_{ij} все существенно упрощается.

§ 3. Обозначения и определения. Как и в гл. 14, мы будем называть матрицу *положительной*, если положительны ее элементы. Если A положительна, то мы будем писать $A > 0$. Запись $A > B$ означает, что $A - B > 0$. Подобно этому вводятся неотрицательные матрицы, обозначаемые $A \geq 0$.

Эти обозначения, разумеется, вступают в противоречие с обозначениями, ранее принятыми при рассмотрении положительно определенных матриц. Мы оказываемся перед выбором: усилить ли бдительность и следить за тем, чтобы оба типа матриц не появлялись одновременно, или ввести новые обозначения типа $A \gg B$. Мы все же прибегнем к обозначениям более простым, но требующим от читателя внимания. В конце концов такая ситуация более предпочтительна.

Вектор x будем называть *положительным*, если все его компоненты положительны, и *неотрицательным*, если все его компоненты неотрицательны. Соответственно будем писать $x > 0$ и $x \geq 0$. Соотношение $x \geq y$ означает, что $x - y \geq 0$.

Упражнения

1. Показать, что из $x \geq y$, $A \geq 0$ следует, что $Ax \geq Ay$.
2. Доказать, что из условия $Ax \geq 0$ при всех $x \geq 0$ следует, что $A \geq 0$.

§ 4. Теорема Перрона. Сформулируем теперь фундаментальный результат Перрона.

Теорема 1. Если A — положительная матрица, то существует единственное характеристическое число матрицы A , $\lambda(A)$ с наибольшей абсолютной величиной. Это характеристическое число положительное и простое, а соответствующий ему собственный вектор может быть выбран положительным.

Имеется много различных доказательств этой теоремы. Мы приведем доказательство, являющееся во многих отношениях наиболее важным, поскольку попутно будет получено вариационное описание для $\lambda(A)$. Из этого описания непосредственно следуют многие свойства $\lambda(A)$, к тому же его можно обобщить на случай более общих операторов.

§ 5. Доказательство теоремы 1. Доказательство теоремы 1 будет получено в ходе доказательства теоремы 2.

Теорема 2. Пусть матрица A положительная, и пусть $\lambda(A)$ — наибольшее по модулю характеристическое число матрицы A . Пусть $S(\lambda)$ — множество всех неотрицательных λ , для каждого из которых найдется неотрицательный вектор x такой, что $Ax \geq \lambda x$ (*), $T(\lambda)$ — множество всех положительных λ таких, что $Ax \leq \lambda x$ при некотором положительном векторе x . Тогда

$$\begin{aligned} \lambda(A) &= \max \lambda, \quad \lambda \in S(\lambda); \\ \lambda(A) &= \min \lambda, \quad \lambda \in T(\lambda). \end{aligned} \quad (1)$$

Доказательство теоремы 2. Рассмотрим множество векторов, определенное соотношением

$$|x| = \sum_{i=1}^N x_i = 1. \quad (2)$$

Нулевой вектор, очевидно, не входит в это множество. Как и

ранее, положим $|A| = \sum_{i,j=1}^N a_{ij}$. Если $\lambda x \leq Ax$, то

$$\lambda |x| \leq |A| \cdot |x|$$

*) $x \neq 0$. (Прим. ред.)

Пусть $\lambda_N \leq \lambda_{N-1}$, и пусть $A_{N-1} = \|a_{ij}\|$, $i, j = 1, 2, \dots, N-1$, что не умаляет общности. Имеют место следующие соотношения:

$$\sum_{j=1}^{N-1} a_{ij} y_j = \lambda_{N-1} y_i, \quad i = 1, 2, \dots, N-1, \quad y_i > 0, \quad (2)$$

$$\sum_{j=1}^N a_{ij} x_j = \lambda_N x_i, \quad i = 1, 2, \dots, N, \quad x_i > 0. \quad (3)$$

Используя первые $N-1$ уравнений из системы (3), получаем

$$\sum_{j=1}^{N-1} a_{ij} x_j = \lambda_N x_i - a_{iN} x_N = \lambda_N (x_i - a_{iN} x_N / \lambda_N) < \lambda_N x_i, \quad (4)$$

что противоречит свойству минимальности λ_{N-1} *).

Воспользуемся теперь этим результатом для того, чтобы показать, что $\lambda(A)$ является простым корнем уравнения $f(\lambda) = |A - \lambda I| = 0$. Воспользовавшись правилом дифференцирования определителя, получаем формулу

$$f'(\lambda) = -|A_1 - \lambda I| - |A_2 - \lambda I| - \dots - |A_N - \lambda I|, \quad (5)$$

где через A_k обозначена матрица, полученная из матрицы A вычеркиванием k -го столбца и k -й строки. Поскольку $\lambda(A) > \lambda(A_k)$ для каждого k , а каждый определитель $|A_k - \lambda I|$ является многочленом от λ с одним и тем же коэффициентом при старшей степени, то каждый многочлен $|A_k - \lambda I|$ в точке $\lambda = \lambda(A)$ имеет один и тот же знак. Следовательно, $f'(\lambda(A)) \neq 0$.

§ 7. Доказательство свойства минимальности $\lambda(A)$. Выше мы дали доказательство свойства максимальности $\lambda(A)$, в предыдущем же пункте мы использовали свойство минимальности. Сейчас мы докажем свойство минимальности, используя свойство максимальности, хотя мы могли бы и здесь применить ту же схему, что и при доказательстве последнего. Метод, которым мы воспользуемся, — метод сопряженного оператора — является одним из наиболее важных и полезных в анализе. С этим методом мы уже встречались при рассмотрении марковских матриц.

Пусть A' , как обычно, обозначает матрицу, транспонированную к матрице A . Так как характеристические числа матриц A и A' совпадают, то $\lambda(A) = \lambda(A')$. Как мы знаем, $(Ax, y) = (x, A'y)$. Если $Ay \leq \lambda y$ при некотором $y > 0$, то для любого вектора $z \geq 0$

$$\lambda(z, y) \geq (z, Ay) = (A'z, y). \quad (1)$$

* См. ниже в § 7 завершение доказательства теоремы 2. (Прим. ред.)

Пусть z является собственным вектором матрицы A' , соответствующим характеристическому числу $\lambda(A)$. Тогда

$$\lambda(z, y) \geq (A'z, y) = \lambda(A)(z, y). \quad (2)$$

Так как $(z, y) > 0$, то $\lambda \geq \lambda(A)$. Это завершает доказательство свойства минимальности.

§ 8. Эквивалентное определение $\lambda(A)$. Вместо определения $\lambda(A)$, данного выше в теореме 1, мы можем дать следующее определение:

Теорема 3.

$$\lambda(A) = \max_x \min_i \left\{ \sum_{j=1}^N a_{ij} x_j / x_i \right\} = \min_x \max_i \left\{ \sum_{j=1}^N a_{ij} x_j / x_i \right\}. \quad (1)$$

Здесь x принимает значения на множестве всех неотрицательных векторов, отличных от нуля.

Доказательство мы оставляем читателю в качестве упражнения.

§ 9. Предельная теорема. Из теоремы 1 вытекает следующий важный результат:

Теорема 4. Пусть c — произвольный неотрицательный вектор $*$). Тогда предел

$$v = \lim_{n \rightarrow \infty} A^n c / \lambda(A)^n$$

существует и является собственным вектором матрицы A , соответствующим характеристическому числу $\lambda(A)$. Этот предел единствен с точностью до скалярного множителя, зависящего от выбора начального состояния c .

Эта теорема представляет собой обобщение соответствующего результата, доказанного в гл. 14 для специального случая марковских матриц.

Упражнение

1. Получить предыдущий результат, основываясь на его частном случае, установленном для марковских матриц.

§ 10. Стационарный рост. Посмотрим теперь, какое асимптотическое поведение вектора $x(n)$, рассмотренного в § 2, можно усмотреть на основании теоремы Перрона. Мы предположим, что матрица $A = \|a_{ij}\|$ положительная.

*) $c \neq 0$.

В этом случае асимптотическое поведение вектора $x(n)$ при $n \rightarrow \infty$ характеризуется соотношением

$$x(n) \sim \lambda \gamma, \quad (1)$$

где λ — перроново характеристическое число, а γ — собственный вектор, соответствующий λ . Мы знаем, что вектор γ положителен и отличается положительным скалярным множителем от нормированного собственного вектора δ . (Нормированным мы называем здесь вектор, сумма компонент которого равна единице.) Константа пропорциональности, связывающая векторы γ и δ , определяется значениями компонент c_i вектора c — начального распределения по видам.

Замечателен тот факт, что независимо от начального распределения по видам рост популяции асимптотически приобретает стационарность: размер популяции экспоненциально возрастает, но пропорции, в которых в нее входят различные виды, остаются постоянными.

Как мы увидим далее, это же явление имеет место в линейных моделях экономических систем.

§ 11. Непрерывные процессы роста. Возвращаясь к математической модели, описанной в § 2, предположим, что моменты времени, в которые производятся наблюдения над системой, приближаются друг к другу. В пределе мы получим непрерывный процесс роста.

Пусть

$a_{ij}\Delta$ — число индивидуумов типа i , произведенных индивидуумом типа j за время Δ , $j \neq i$;

$1 + a_{ii}\Delta$ — число индивидуумов типа i , произведенных индивидуумом типа i за время Δ ,
 $i = 1, 2, \dots, N.$ (1)

Тогда имеют место уравнения

$$x_i(t + \Delta) = (1 + a_{ii}\Delta)x_i(t) + \Delta \sum_{j \neq i} a_{ij}x_j(t), \quad i = 1, 2, \dots, N. \quad (2)$$

Заметим, что теперь a_{ij} представляют собой интенсивности рождения. Формально переходя к пределу при $\Delta \rightarrow 0$, получаем следующие дифференциальные уравнения:

$$\frac{dx_i}{dt} = a_{ii}x_i + \sum_{j \neq i} a_{ij}x_j, \quad i = 1, 2, \dots, N. \quad (3)$$

Мы можем теперь определить непрерывный процесс роста посредством этих уравнений, оговаривая, что процесс является пределом (в некотором смысле) дискретного процесса.

Асимптотическое поведение $x_i(t)$ при $t \rightarrow \infty$ определяется характеристическими числами матрицы A , имеющими наибольшую действительную часть.

§ 12. Аналог теоремы Перрона. Докажем теперь следующую теорему:

Теорема 5. Если

$$a_{ij} > 0, \quad i \neq j, \quad (1)$$

то характеристическое число матрицы A , имеющее наибольшую действительную часть, является простым и действительным. Этому характеристическому числу соответствует положительный собственный вектор, единственный с точностью до постоянного множителя. Кроме того,

$$\rho(A) = \max_x \min_i \left\{ \sum_{j=1}^N a_{ij} x_j / x_i \right\} = \min_x \max_i \left\{ \sum_{j=1}^N a_{ij} x_j / x_i \right\}. \quad (2)$$

Доказательство. Этот результат легче всего получить как предельный случай теоремы Перрона. Пусть $\rho(A)$ обозначает характеристическое число с наибольшей действительной частью. Очевидно, что характеристическое число матрицы $e^{\delta A}$ имеет вид

$$\lambda(e^{\delta A}) = e^{\delta \rho(A)} \quad (3)$$

при $\delta > 0$. Как это следует из § 15 гл. 10, матрица $e^{\delta A}$ положительна при наших предположениях, и, следовательно, характеристическое число $\rho(e^{\delta A})$ положительное и простое.

Используя вариационное представление для $\lambda(e^{\delta A})$, имеем

$$e^{\delta \rho(A)} = \max_x \min_i \left(1 + \delta \sum_{j=1}^N a_{ij} x_j / x_i \right) + O(\delta^2). \quad (4)$$

Поскольку δ в (4) сколь угодно мало, то мы получаем искомое представление (2).

Упражнения

1. Если $B \geq 0$, а A такая, что $a_{ij} > 0$ при $i \neq j$, то $\rho(A+B) \geq \rho(A)$.
2. Вывести теорему 5 непосредственно из теоремы Перрона 1, рассмотрев матрицу $sI + A$, где s выбрано так, что матрица $sI + A$ является положительной.

§ 13. Ядерный распад. Этот же тип матриц возникает в связи с некоторыми простыми моделями ядерного распада. В качестве объектов теперь могут выступать N различных видов элементарных частиц или одна частица, скажем, нейтрон, находящаяся в N различных энергетических состояниях.

Ссылки на работы в этой области имеются в конце этой главы.

§ 14. Математическая экономика. Рассмотрим простую модель экономической системы. Имеется N различных отраслей промышленности, функционирование каждой из которых зависит от состояния остальных. Мы предположим, что состояние системы (совокупности) отраслей может быть описано вектором $x(n)$, i -я компонента которого так или иначе описывает состояние i -й отрасли. Взаимодействие отраслей моделируется системой рекуррентных соотношений, или разностных уравнений вида

$$\begin{aligned} x_i(n+1) &= g_i(x_1(n), x_2(n), \dots, x_N(n)), \\ x_i(0) &= c_i, \quad i = 1, 2, \dots, N. \end{aligned} \quad (1)$$

Такая формулировка, однако, слишком туманна. Для того чтобы увидеть, как соотношения этого типа возникают в действительности, рассмотрим простую модель трех взаимодействующих отраслей, которые мы будем условно называть «автомобилестроительной», «сталелитейной» и «станкостроительной».

Предположим, что каждая из этих отраслей в любой момент времени может быть описана ее сырьевым ресурсом и ее производительной мощностью.

Введем следующие переменные состояния:

$x_1(n)$ — число автомобилей, произведенных к моменту времени n ;

$x_2(n)$ — производственная мощность автомобилестроительных заводов в момент времени n ;

$x_3(n)$ — запас стали в момент времени n ; (2)

$x_4(n)$ — производственная мощность сталелитейных заводов в момент времени n ;

$x_5(n)$ — количество (ресурс) станков в момент времени n ;

$x_6(n)$ — производственная мощность станкостроительных заводов в момент времени n .

Для того чтобы получить соотношения, связывающие $x_i(n+1)$ и $x_i(n)$, мы должны сделать некоторые допущения относительно экономической взаимозависимости рассматриваемых трех отраслей.

Эти допущения состоят в следующем:

Для того чтобы увеличить производственную мощность каждой из трех отраслей, требуются только станки и сталь. (3а)

Для производства автомобилей требуются сталь и производственная мощность автомобилестроительных заводов. (3б)

Для производства стали требуется только производственная мощность сталелитейных заводов. (3в)

Для производства станков требуются сталь и производственная мощность станкостроительных заводов. (3г)

Динамика процесса производства такова: в начале временного периода, протекающего от момента n до момента $n+1$, некоторые количества стали и станков размещаются на заводах трех отраслей для дальнейшего производства стали, станков и автомобилей, а также для увеличения существующих производственных мощностей, причем каждая отрасль использует для этого свои запасы.

Пусть

$z_i(n)$ — количество стали, предназначенное для увеличения $x_i(n)$ в момент времени n ,
 $i = 1, 2, \dots, 6$; (4)

$w_i(n)$ — количество станков, предназначенное для увеличения $x_i(n)$ в момент времени n , $i = 1, 2, \dots, 6$.

Обращаясь к ранее сделанным предположениям (3), мы видим, что

$$z_s(n) = 0, \quad (5a)$$

$$w_1(n) = w_2(n) = w_6(n) = 0. \quad (5b)$$

Для того чтобы получить соотношения, связывающие $x_i(n+1)$ с $x_j(n)$, $z_j(n)$ и $w_j(n)$, мы должны сделать некоторые предположения, касающиеся зависимости входа и выхода. Простейшим из такого рода предположений является предположение о линейности производства, когда выход продукции прямо пропорционален входу по наименее обеспеченному параметру. Конкретно это значит, что объем выпускаемой продукции пропорционален производственной мощности, если отсутствуют какие-либо ограничения на сырьевые материалы, и пропорционален количеству наименее обеспеченного сырьевого материала, если отсутствуют ограничения на производственную мощность.

Процессы этого типа называются «процессами с узкими местами».

Для получения уравнений, описывающих рассматриваемый процесс, воспользуемся принципом сохранения. Количество материала в момент $n+1$ равно количеству материала в момент n минус использованное за время $(n, n+1)$ плюс произведенное за время $(n, n+1)$.

Ограничения на выбор z_i и w_i состоят в том, что мы не можем использовать на каждой стадии больше материала, чем запасено; кроме того, нет смысла выделять больше сырья, чем можно освоить при данной производственной мощности.

С учетом эффекта узких мест уравнения сохранения имеют вид

$$\begin{aligned}x_1(n+1) &= x_1(n) + \min(\gamma_1 x_2(n), \alpha_1 z_1(n)), \\x_2(n+1) &= x_2(n) + \min(\alpha_2 z_2(n), \beta_2 w_2(n)), \\x_3(n+1) &= x_3(n) - z_1(n) - z_2(n) - z_3(n) - z_4(n) - z_5(n) + \gamma_3 x_4(n), \\x_4(n+1) &= x_4(n) + \min(\alpha_4 z_4(n), \beta_4 w_4(n)), \\x_5(n+1) &= x_5(n) - w_2(n) - w_4(n) - w_5(n) + \min(\gamma_5 x_6(n), \alpha_5 z_5(n)), \\x_6(n+1) &= x_6(n) + \min(\alpha_6 z_6(n), \beta_6 w_6(n)),\end{aligned}\quad (6)$$

где $\alpha_i, \beta_i, \gamma_i$ — константы.

Ограничения на выбор z_i и w_i имеют вид

$$z_i, w_i \geq 0, \quad (7a)$$

$$z_1 + z_2 + z_4 + z_5 + z_6 \leq x_3, \quad (7б)$$

$$w_2 + w_4 + w_5 \leq x_6. \quad (7в)$$

Если ограничения, связанные с производственными мощностями, являются определяющими, то уравнения (6) становятся линейными. Действительно, в этом случае

$$\alpha_2 z_2 = \beta_2 w_2, \quad (8a)$$

$$\alpha_4 z_4 = \beta_4 w_4, \quad (8б)$$

$$\alpha_6 z_6 = \beta_6 w_6. \quad (8в)$$

Исключая с помощью этих соотношений w_i , получаем линейные уравнения

$$\begin{aligned}x_1(n+1) &= x_1(n) + \alpha_1 z_1(n), \quad x_1(0) = c_1; \\x_2(n+1) &= x_2(n) + \alpha_2 z_2(n), \quad x_2(0) = c_2; \\x_3(n+1) &= x_3(n) - z_1(n) - z_2(n) - z_3(n) - z_4(n) - z_5(n) - \\&\quad - z_6(n) + \gamma_3 x_4(n), \quad x_3(0) = c_3; \\x_4(n+1) &= x_4(n) + \alpha_4 z_4(n), \quad x_4(0) = c_4;\end{aligned}\quad (9)$$

$$\begin{aligned}x_5(n+1) &= x_5(n) - e_2 z_2(n) - e_4 z_4(n) - e_6 z_6(n) + \alpha_5 z_5(n), \\e_i &= \alpha_i / \beta_i, \quad x_5(0) = c_5;\end{aligned}$$

$$x_6(n+1) = x_6(n) + \alpha_6 z_6(n), \quad x_6(0) = c_6.$$

Ограничения на выбор z_i принимают вид

$$z_i \geq 0, \quad (10a)$$

$$z_1 + z_2 + z_4 + z_5 + z_6 \leq x_3, \quad (10б)$$

$$\gamma_2 z_2 + \gamma_4 z_4 + \gamma_6 z_6 \leq x_3, \quad (10в)$$

$$z_1 \leq f_1 x_3, \quad (10г)$$

$$z_5 \leq f_5 x_3. \quad (10д)$$

Предположим, что эти условия выполняются при

$$z_2 = e_2 x_2, \quad z_4 = e_4 x_4, \quad z_6 = e_6 x_6, \quad (11)$$

где скаляры e_2 , e_4 и e_6 выбраны с учетом условий (10) и

$$z_1 = f_1 x_1, \quad z_3 = f_3 x_3. \quad (12)$$

Пусть, кроме того, выполнены условия (10).

Тогда уравнения (9) принимают вид

$$x_i(n+1) = (1 - a_{ii})x_i(n) + \sum_{j \neq i} a_{ij}x_j(n), \quad i = 1, 2, \dots, 6, \quad (13)$$

где

$$a_{ij} \geq 0, \quad i, j = 1, \dots, 6. \quad (14)$$

Для непрерывного случая эти уравнения таковы:

$$\frac{dx_i}{dt} = -a_{ii}x_i + \sum_{j \neq i} a_{ij}x_j, \quad i = 1, 2, \dots, 6. \quad (15)$$

Наше детальное рассмотрение экономической модели иллюстрирует ту важную роль, которую играют матрицы специального вида, называемые иногда «входными—выходными». К сожалению, в наиболее интересных и практически важных случаях мы не можем заменить условие (14) более ограничительным условием положительности. Результатом этого является тот факт, что изучение асимптотического поведения решений уравнений вида (13) очень сложно. Ссылки на довольно обширные исследования в этой области приводятся в конце главы.

§ 15. Матрицы Минковского — Леонтьева. Рассмотрим теперь специальный класс неотрицательных матриц, определяемый условиями

$$0 \leq a_{ij}, \quad (1a)$$

$$\sum_{i=1}^N a_{ij} \leq 1. \quad (16)$$

Эти матрицы возникают в связи с решением линейных уравнений вида

$$x_i = \sum_{j=1}^N a_{ij}x_j + y_i, \quad i = 1, 2, \dots, N, \quad (2)$$

встречающихся при рассмотрении моделей, очень похожих на предыдущую, для межотраслевых производственных процессов.

Докажем следующую теорему:

Теорема 6. Если $0 \leq a_{ij}$ и

$$\sum_{i=1}^N a_{ij} < 1, \quad j = 1, 2, \dots, N, \quad (3)$$

то система (2) имеет единственное решение, которое положительно, если все y_i положительны.

Если $a_{ij} > 0$, то матрица $(I - A)^{-1}$ также положительна*).

Доказательство. Чтобы показать, что матрица $(I - A)^{-1}$ является положительной при условии, что $a_{ij} > 0$, достаточно показать, что транспонированная к ней матрица положительна, т. е. $(I - A')^{-1} \geq 0$. Рассмотрим сопряженную систему уравнений

$$z_i = \sum_{j=1}^N a_{ji} z_j + w_i \quad i = 1, 2, \dots, N. \quad (4)$$

Легко видеть, что в силу условия (3) прямой итерационный процесс дает решение уравнения (4), которое будет положительным, если $w_i > 0$, $i = 1, 2, \dots, N$.

Из условия $a_{ij} > 0$ также следует, что z_i положительны всякий раз, когда w является нетривиальным неотрицательным вектором. Таким образом, доказав, что $(I - A)^{-1}$ существует и положительна, мы доказали вторую часть теоремы.

§ 16. Положительность определителя $|I - A|$. Поскольку при указанных условиях система (15.2) имеет единственное решение при всех y_i , то $|I - A| \neq 0$. Чтобы показать, что $|I - A| > 0$, воспользуемся соображениями непрерывности, подобно тому, как мы сделали это в § 4 гл. 4. Если $\lambda \geq 0$, то матрица λA удовлетворяет тем же условиям что и A . Поэтому определитель $|I - \lambda A|$ отличен от нуля при всех λ таких, что $0 \leq \lambda \leq 1$. Поскольку этот определитель положителен при $\lambda = 0$ и непрерывен по λ , он положителен и при $\lambda = 1$.

§ 17. Усиление теоремы 6. Ясно, что если $\sum_{i=1}^N a_{ji} = 1$ при всех j , то 1 является характеристическим числом матрицы A , откуда следует, что $|I - A| = 0$. С другой стороны, можно ожидать, что условие $\sum a_{ij} < 1$ может быть ослаблено.

Достаточными оказываются следующие предположения относительно элементов матрицы A :

$$1 > a_{ij} > 0, \quad (1a)$$

$$\sum_j a_{ij} < 1 \quad \text{хотя бы для одного } j, \quad (1б)$$

$$\sum_j a_{ij} \leq 1 \quad \text{для всех } j. \quad (1в)$$

Это легко доказать, убедившись в том, что элементы матрицы A^k удовлетворяют условиям (1) при всех j .

*) В условиях теоремы $(I - A)^{-1} = \sum_{k=0}^{\infty} A^k > 0$. (Прим. ред.)

§ 18. **Линейное программирование.** В первой части этой книги мы рассматривали задачу максимизации квадратичной формы при квадратичных ограничениях, а также задачу максимизации квадратичной формы при линейных ограничениях. Однако мы тщательно избегали всех вопросов, связанных с максимизацией линейных форм при линейных ограничениях.

Типичная задача этого класса формулируется следующим образом: требуется максимизировать функцию

$$L(x) = \sum_{i=1}^N c_i x_i \quad (1)$$

при ограничениях на аргументы вида

$$\sum_{j=1}^N a_{ij} x_j \leq b_i, \quad i = 1, 2, \dots, M, \quad (2a)$$

$$x_i \geq 0. \quad (2б)$$

Читатель очень быстро обнаружит, что ни один из методов, рассмотренных в первой части книги, посвященной квадратичным формам, не пригоден здесь. Вопросы этого типа рассматриваются в недавно развитой теории линейных неравенств. Этой теории будет посвящен один из последующих томов этой серии. Теорема Перрона — как сам результат, так и метод доказательства — тесно связана с различными частями теории *линейных неравенств*.

Кроме результатов, связанных с вопросами существования и характером решений сформулированной выше задачи, очень большое значение имеет построение алгоритмов численного решения этой задачи. Эта часть общей теории линейных неравенств называется *линейным программированием*.

Приведем теперь очень простой пример того, как такие задачи возникают в математической экономике. Предположим, что мы располагаем M видами ресурсов в количествах соответственно x_1, x_2, \dots, x_M . Этими ресурсами могут быть люди, машины, деньги и т. д. Пусть эти ресурсы используются в N различных производственных процессах, таких, как, например, бурение скважин, производство автомобилей и т. д.

Пусть

x_{ij} — количество i -го ресурса, используемое
в j -м производственном процессе, (3)

и при этом

$$\sum_{j=1}^N x_{ij} = x_i, \quad i = 1, 2, \dots, M, \quad (4a)$$

$$x_{ij} \geq 0. \quad (4б)$$

Предположим — и это является обычно довольно грубым приближением, — что полезность вложенного ресурса x_{ij} определяется простой пропорциональностью, т. е. полезность равна $a_{ij}x_{ij}$. Сделаем дальнейшее предположение относительно аддитивности полезностей, мы приходим к задаче максимизации линейной формы

$$l(\mathbf{x}) = \sum_{i,j} a_{ij}x_{ij} \quad (5)$$

при ограничениях вида (4).

В качестве другого примера того, как возникают линейные вариационные задачи, можно рассмотреть модель, приведенную в § 14. Вместо того, чтобы заранее задавать значения величин z_i и w_i , как мы это делали, можно искать значения этих величин, максимизирующие суммарный выпуск стали за N циклов процесса.

§ 19. Теория игр. Предположим, что два игрока A и B принимают участие в игре следующего простого типа. У первого игрока имеется M ходов, а у второго N . Если A делает i -й ход, а B делает j -й ход, то A получает сумму a_{ij} , а B получает $-a_{ij}$.

Матрица

$$A = |a_{ij}| \quad (1)$$

называется *матрицей платежей*.

Игроки делают свои ходы одновременно, не зная о решении противника. Затем арбитр в соответствии со сформулированными выше правилами определяет суммы, подлежащие платежу.

Предположим, что эта ситуация повторяется снова и снова. Вообще говоря, ни одному из игроков не выгодно делать один и тот же ход в каждой партии. Каждый игрок с помощью некоторого случайного процесса выбирает свои ходы так, чтобы не дать своему противнику извлечь выгоду из его стратегии.

Пусть

$$\begin{aligned} x_i & \text{— вероятность того, что } A \text{ сделает } i\text{-й ход,} \\ & \quad i = 1, 2, \dots, M, \\ y_j & \text{— вероятность того, что } B \text{ сделает } j\text{-й ход,} \\ & \quad j = 1, 2, \dots, N. \end{aligned} \quad (2)$$

Тогда средний выигрыш, получаемый за игру игроком A , равен

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i,j} a_{ij}x_i y_j. \quad (3)$$

Задача состоит в том, чтобы определить x_i и y_j .

Игрок A может рассуждать так: «Предположим, что B знает мой выбор. Тогда он должен выбрать y_j , чтобы минимизировать

$f(x, y)$. В таком случае я выберу x_i , чтобы максимизировать выигрыш. Придерживаясь такой стратегии, игрок A получит средний выигрыш

$$v_A = \max_x \min_y f(x, y), \quad (4)$$

где область изменения x и y задается соотношениями

$$x_i, y_i \geq 0, \quad (5a)$$

$$\sum_i x_i = \sum_j y_j = 1. \quad (5b)$$

Точно так же игрок B может гарантировать, что не проиграет больше, чем

$$v_B = \min_y \max_x f(x, y). \quad (6)$$

Основной результат теории игр состоит в том, что $v_A = v_B$. Это *минимакс-теорема* фон Неймана.

Отметим еще один замечательный факт. Можно показать, что упомянутая теоретико-игровая задача и задача из теории линейных неравенств, описанная в § 18, математически эквивалентны. Эта эквивалентность, оказывается, является следствием дуальности, присущей N -мерной евклидовой геометрии.

§ 20. Марковские процессы принятия решений. Рассмотрим теперь некоторые задачи, возникающие в теории динамического программирования.

Представим себе физическую систему S , которая в каждый момент времени $t = 0, 1, \dots$ находится в одном из состояний S_1, S_2, \dots, S_N . Пусть

$x_i(n)$ — вероятность того, что система S находится в состоянии S_i в момент n ; $n = 1, 2, \dots$ (1)

Пусть для каждого q , являющегося векторной переменной,

$$M(q) = \|m_{ij}(q)\| \quad (2)$$

представляет собой марковскую матрицу.

Вместо того чтобы считать, что при фиксированном q $M(q) \cdot x(0)$ задает вероятностное распределение в момент времени n , как для обычного марковского процесса, мы предположим, что q может изменяться от такта к такту. Конкретно, мы будем считать, что осуществляется управление этим процессом с целью увеличения вероятности того, что система будет находиться в состоянии I в любой момент времени.

Вместо обычных уравнений мы получаем соотношения

$$x_i(n+1) = \max_q \sum_{j=1}^N m_{ij}(q) x_j(n),$$

$$x_i(n+1) = \sum_{j=1}^N m_{ij}(q^*) x_j(n),$$
(3)

где q^* — значение вектора q , при котором достигается максимум суммы в правой части выражения в первой строке.

Мы оставляем читателю исследование очень интересного вопроса, связанного с получением простых условий, при которых существует «стационарное», или равновесное, решение. В § 21 мы рассмотрим более общую систему.

§ 21. Экономическая модель. Пусть мы располагаем N различными видами ресурсов. Пусть

$$x_i(n) — \text{количество } i\text{-го ресурса в момент } n. \quad (1)$$

Предположим, что имеют место следующие линейные соотношения:

$$x_i(n+1) = \sum_{j=1}^N a_{ij}(q) x_j(n), \quad (2)$$

где q , как и ранее, — некоторый векторный параметр. Если $A(q)$ — положительная матрица, то асимптотическое поведение системы описывается теоремой Перрона.

Предположим, однако, как и в § 20, что процесс управляется с целью увеличения количества каждого ресурса в каждый момент времени. В этом случае уравнения (2) принимают вид

$$x_i(n+1) = \max_q \sum_{j=1}^N a_{ij}(q) x_j(n), \quad i = 1, 2, \dots, N. \quad (3)$$

Имеет место следующее обобщение теоремы Перрона:

Теорема 7. Если q принимает значения из конечного множества (q_1, q_2, \dots, q_M) такого, что

$$\text{максимум в (3) достигается,} \quad (4a)$$

$$0 < m_1 \leq a_{ij}(q) \leq m_2 < \infty, \quad (4b)$$

$\max_q \lambda(A(q))$ существует и достигается при некотором q_i , (4в)

то существует, и притом единственное, положительное число λ такое, что однородная система

$$\lambda y_i = \max_q \sum_{j=1}^N a_{ij}(q) y_j \quad (5)$$

имеет положительное решение $y_i > 0$. Это решение единственно с точностью до множителя и

$$\lambda = \max_q \lambda(A(q)). \quad (6)$$

Кроме того, при $n \rightarrow \infty$

$$x_i(n) \sim a_i y_i \lambda^n, \quad (7)$$

где a_i зависит от начальных условий.

Доказательство этой теоремы можно найти в работах, цитируемых в конце этой главы.

Упражнения к гл. 16

1. Пусть $A = \|a_{ij}\|$ — неотрицательная матрица. Необходимым и достаточным условием того, чтобы все характеристические числа матрицы A лежали внутри единичного круга, является положительность всех главных миноров матрицы $I - A$. (Мецлер.)

2. Пусть D — диагональная матрица; пусть диагональные элементы матрицы A отрицательны, а ее внедиагональные элементы неотрицательны. Пусть, кроме того, собственные значения матриц A и DA имеют отрицательные действительные части. Тогда матрица D положительна. (Эрроу и Энтховен.)

3. Рассмотрим матрицу $Z(t) = \|z_{ij}(t)\|$, элементы которой обладают следующими свойствами:

$$(a) z_{ij} > 0,$$

$$(b) \int_0^{\infty} z_{ii} dt > 1 \text{ при некотором } i,$$

$$(в) \int_0^{\infty} z_{ij} e^{-at} dt < \infty \text{ при некотором } a > 0.$$

Показать, что существует положительный вектор x и положительное число s_0 такие, что

$$\left(\int_0^{\infty} Z e^{-s_0 t} dt \right) x = x.$$

Более того, s_0 является простым корнем уравнения $\left| I - \int_0^{\infty} Z e^{-s t} dt \right| = 0$, обладающим наибольшей действительной частью. (Боненблат¹⁾.)

4. Пусть $A = \|a_{ij}\|$ и $s_i = |a_{ii}| - \sum_{j \neq i} |a_{ij}|$. Если $s_i > 0$, $1 \leq i \leq N$, то (как мы уже знаем) $A^{-1} = \|b_{ij}\|$ существует. Показать, что $|b_{ij}| < 1/s_j$. (Фань Цзы.)

¹⁾ См. Беллман (R. Bellman), A Survey of the Mathematical Theory of Time-lag, Retarded Control, and Hereditary Processes, with the assistance of J. M. Danskin, Jr., The Rand Corporation, Rept. R-256, March 1, 1954.

5. Говорят, что действительная матрица $A = \|a_{ij}\|$ увеличивает максимум, если

$$\max_{1 \leq i \leq N} x_i < \max_{1 \leq i \leq N} \sum_{j=1}^N a_{ij} x_j$$

для любого действительного вектора x . Доказать, что матрица A является увеличивающей максимум тогда и только тогда, когда ее обратная матрица $A^{-1} = \|b_{ij}\|$ обладает следующими свойствами

(а) $b_{ij} > 0$,

(б) $\sum_{j=1}^N b_{ij} = 1, \quad 1 \leq i \leq N.$ (Фань Цзы.)

6. Пусть x_i — положительные числа, а $f(x_1, x_2, \dots, x_N)$ и $g(x_1, x_2, \dots, x_N)$ обозначают соответственно наименьшую и наибольшую из $N+1$ величин $x_1, x_2 + 1/x_1, x_3 + 1/x_2, \dots, x_N + 1/x_{N-1}, 1/x_N$. Доказать, что

$$\max_{x_i > 0} f(x_1, x_2, \dots, x_N) = \min_{x_i > 0} g(x_1, x_2, \dots, x_N) = 2 \cos(\pi/(N+2)).$$

7. Пусть $a > 0$. Для любого разбиения $0 = t_0 < t_1 < \dots < t_N = 1$ отрезка $[0, 1]$ сумма Римана $\sum_{i=1}^N (t_i - t_{i-1})/(a + t_i)$ для интеграла $\int_0^1 dt/(a+t)$ всегда

содержит член $\geq [1 - (a/a+1)^{1/N}]$, а также член, не превосходящий этой величины (Фань Цзы.)

8. Пусть элементы матрицы A таковы, что $a_{ij} < 0$ при $i \neq j$ и $a_{jj} + \sum_{i \neq j} a_{ij} = 0, \quad j = 1, 2, \dots, N$. Тогда если λ — собственное значение матрицы A , то либо $\lambda = 0$, либо $\text{Re}(\lambda) > 0$. (А. Брауэр — О. Таусски.)

9. Если $u_i \geq 0, \quad a_{ij} \geq 0, \quad i \neq j$, и $\sum_{j=1}^N a_{ij} = a_{ij}, \quad \text{то}$

$$D_N(u) = \begin{vmatrix} a_{11} + u_1 & -a_{12} & \dots & -a_{1N} \\ -a_{21} & a_{22} + u_2 & \dots & -a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{N1} & -a_{N2} & \dots & a_{NN} + u_N \end{vmatrix} \geq 0. \quad (\text{Минковский.})$$

10. Если $a_{ij} \geq 0$ при $i \neq j$ и $b_j \geq \sum_{i \neq j} a_{ij}$, то все алгебраические дополнения элементов определителя

$$D_N = \begin{vmatrix} b_1 & -a_{12} & \dots & -a_{1N} \\ -a_{21} & b_2 & \dots & -a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{N1} & -a_{N2} & \dots & b_N \end{vmatrix}$$

неотрицательны (Ледерман.)

11. Действительная матрица X , у которой $x_{ii} > \sum_{j \neq i} |x_{ij}|$, называется матрицей Адамара. Пусть A и B — адамаровы матрицы. Тогда $|A+B| \geq |A|+|B|$. (Хейнсворт¹⁾.) См также работы А. Островского²⁾.

Матрицы Адамара играют важную роль в численном анализе. Эти матрицы и матрицы более общего типа, *доминантные матрицы*, находят применение в исследовании многополюсников. См. работы

Слепян и Вайнберг (P. Slepian and L. Weinberg), *Synthesis Applications of Paramount and Dominant Matrices*, Hughes Research Laboratories, 1958;

Берингтон (R. S. Burington), On the Equivalence of Quadrics in m -affine n -space and Its Relation to the Equivalence of $2m$ -pole Networks, *Trans. Amer. Math. Soc.* 38 (1935), 163—176;

Берингтон (R. S. Burington), R -matrices and Equivalent Networks, *J. Math. and Phys.* 16 (1937), 85—103.

12. Пусть $|a_{ij}| \leq m|a_{ii}|$, $j < i$, $i = 1, 2, \dots, N$, и $|a_{ij}| \leq M|a_{ii}|$, $j > i$, $i = 1, 2, \dots, N-1$, тогда матрица A вырожденная, если $m/(1+m)^N < M(1+M)^N$ и $m < M$. Если $m = M$, то для невырожденности достаточно, чтобы $m < 1/(N-1)$. (Островский³⁾.)

13. M -матрицей называется такая действительная матрица A , что $a_{ij} \geq 0$, $i \neq j$, и имеет место одно из следующих трех эквивалентных свойств:

(а) существует N положительных чисел x_j таких, что

$$\sum_{j=1}^N a_{ij}x_j > 0, \quad i = 1, 2, \dots, N;$$

(б) матрица A невырожденная и все элементы матрицы A^{-1} неотрицательны;

(в) все главные миноры матрицы A положительны

M -матрицы были введены Островским⁴⁾. См. также Хейнсворт (E. V. Haynsworth), Note on Bounds for Certain Determinants, *Duke Math. J.* 24 (1957), 312—320;

Хейнсворт (E. V. Haynsworth), Bounds for Determinants with Dominant Main Diagonal, *Duke Math. J.* 20 (1953), 199—209, где имеются дальнейшие ссылки.

14. Пусть A и B — положительные квадратные матрицы порядка N такие, что $AB \neq BA$, а c — данный положительный N -мерный вектор. Рассмотрим вектор $x_N = Z_N Z_{N-1} \dots Z_2 Z_1 c$, где каждая матрица Z_i равна либо A , либо B . Допустим, что Z_i надлежит выбрать таким образом, чтобы максимизировать скалярное произведение (x_N, b) , где b — фиксированный вектор. Определим функцию

$$f_N(c) = \max_{\{Z_i\}} (x_N, b)$$

¹⁾ E. V. Haynsworth, Bounds for Determinants with Dominant Main Diagonal, *Duke Math. J.* 20 (1953), 199—209.

²⁾ A. Ostrowski, Über die Determinanten mit überwiegender Hauptdiagonale, *Comment. Math. Helveticii* 10 (1937—1938), 69—96; A. Ostrowski, Note on Bounds for Some Determinants, *Duke Math. J.* 22 (1955), 95—102.

³⁾ A. Ostrowski, On Nearly Triangular Matrices, *J. Research Nat. Bur. Standards* 52 (1954), 319—345.

⁴⁾ A. Ostrowski, *Comment. Math. Helveticii* 30 (1956), 175—210; *ibid.* 10 (1937), 69—96.

Доказать, что

$$f_1(c) = \max((Ac, b), (Bc, b)),$$

.....

$$f_N(c) = \max(f_{N-1}(Ac), f_{N-1}(Bc)).$$

15. Существует ли такой скаляр λ , что $f_N(c) \leq \lambda^N g(c)$?

16. Рассмотреть соответствующую задачу для случая, когда максимизируется перроново характеристическое число матрицы $Z_N Z_{N-1} \dots Z_2 Z_1$.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

§ 1. Очень много интересных и важных аналитических задач возникает при исследовании различных математических моделей экономического взаимодействия. См. обзорную статью

Герштейн (I. N. Herstein), Some Mathematical Methods and Techniques in Economics, Quart. Appl. Math. 11 (1953), 249—261.

Укажем еще две книги, к которым может обратиться читатель, интересующийся этими вопросами:

Купманс (T. Koopmans), Activity Analysis of Production and Allocation, John Wiley & Sons, Inc., New York, 1951;

Моргенштерн (O. Morgenstern) (ed.), Economic Activity Analysis, John Wiley & Sons, Inc., New York, 1954.

Приложения теории положительных матриц к исследованию динамической устойчивости больших систем сбыта описаны в статьях

Эрроу и Нерлав (K. J. Arrow and M. Nerlove), A Note on Expectations and Stability, TR No. 41, Dept. of Economics, Stanford University, 1957;

Эрроу и Энтховен (K. J. Arrow and A. Enthoven), A Theorem on Expectations and Stability of Expectations, Econometrica 24 (1956), 288—293.

где также приведена литература по этим вопросам.

По поводу терминологической неразберихи можно добавить, что помимо положительно определенных матриц, с которыми мы имели дело в первой части книги, и положительных матриц Перрона, важную роль играют также вполне положительные матрицы*).

§ 2. Изложение математических моделей процессов роста этого типа содержится в монографии

Харрис (T. E. Harris), The Theory of Branching Processes, Ergeb. Math., 1960 [русский перевод: Теория ветвящихся случайных процессов, «Мир», 1966],

где можно найти много дальнейших ссылок.

*) См. Ф. Р. Гаитмахер и М. Г. Крейн, Осцилляционные матрицы, Гостехиздат, 1950. (Прим. ред.)

§ 3. Понятие положительного оператора является одним из основных в анализе. Развернутое изложение теории и приложений этих операторов имеется в работе

М. Г. Крейн и М. А. Рутман, Линейные операторы, оставляющие инвариантным конус в пространстве Банаха, УМН 3, № 1 (23) (1948), 3—95

§ 4. См.

Перрон (O. Perron), Zur Theorie der Matrizen, Math. Ann. 64 (1907), 248—263

Доказательство, данное самим Перроном, отличается от приводимого нами и очень интересно само по себе. Приведенное же доказательство важно тем, что оно очень легко обобщается на случай бесконечномерных операторов. Это доказательство было дано Боненбластом в связи с одной задачей из теории многомерных ветвящихся процессов, поставленной Беллманом и Харрисом. См. цитированную выше монографию Харриса и докторскую диссертацию

Сноу (R. Snow), Multi-Dimensional Branching Processes, UCLA, Dept. of Mathematics, 1960.

Результат Перрона был существенно обобщен Фробениусом в статье

Фробениус (G. Frobenius), Über Matrizen aus nicht-negativen Elementen, Sitzber. der Kgl. preuss. Acad. Wiss., 1912, 456—477,

где исследован более общий случай, когда на a_{ij} накладывается только ограничение неотрицательности. В течение многих лет основополагающая работа Перрона была забыта и теорему неправильно приписывали Фробениусу.

Об одном подходе к этой теореме с позиций теории линейных дифференциальных уравнений и некоторых обобщениях можно прочитать в работе

Хартман и Уинтнер (P. Hartman and A. Wintner), Linear Differential Equations and Difference Equations with Monotone Solutions, Amer. J. Math. 1 (1953), 731—743

Последние результаты можно найти в статье

Биркгоф и Варга (G. Birkhoff and R. S. Varga), Reactor Criticality and Nonnegative Matrices, J. Ind. and Appl. Math. 6 (1958), 354—377.

Как и в случае марковских матриц, задача изучения поведения собственных значений при более слабом ограничении неотрицательности порождает различные частные случаи. Условия существования единственного собственного значения с наибольшей абсолютной величиной наиболее четко выражаются в вероятностных, экономических и топологических терминах. Детальное

рассмотрение этих вопросов и ссылки на другие источники читатель найдет в работах

Уонг (Y. K. Wong), *Some Mathematical Concepts for Linear Economic Models, Economic Activity Analysis*, O. Morgenstern (ed.), John Wiley and Sons, Inc., New York, 1954;

Уонг (Y. K. Wong), *An Elementary Treatment of an Input-Output System*, *Naval Research Logistics Quarterly* 1 (1954), 321—326;

Вудбери (M. A. Woodbury), *Properties of Leontieff-Type Input-Output Matrices, Economic Activity Analysis*, O. Morgenstern (ed.), John Wiley & Sons, Inc., New York, 1954;

Вудбери (M. A. Woodbury), *Characteristic Roots of Input-Output Matrices, Economic Activity Analysis*, O. Morgenstern (ed.), John Wiley and Sons, Inc., New York, 1954;

Нобл (S. B. Noble), *Measures of the Structure of Some Static Linear Economic Models*, George Washington University, 1958.

Первое доказательство теоремы Перрона на основе теорем о неподвижной точке дано в книге

Александров и Хопф (P. Alexandroff and H. Hopf), *Topologie I*, Berlin, 1935.

Подробное рассмотрение этого метода со многими обобщениями и приложениями читатель найдет в работе

Фань Цзы (Ku Fan), *Topological Proofs for Certain Theorems on Matrices with Non-negative Elements*, *Monatsh. Math.* 62 (1958), 219—237.

§ 5. Другие доказательства имеются в работах

Брауэр (A. Brauer), *A New Proof of Theorems of Perron and Frobenius on Non-negative Matrices, I: Positive Matrices*, *Duke Math. J.* 24 (1957), 367—378;

Дебре и Герштейн (G. Debreu and I. N. Herstein), *Non-negative Square Matrices*, *Econometrica* 21 (1953), 597—607;

Сеймелсон (H. Samuelson), *On the Perron-Frobenius Theorem*, *Michigan Math. J.* 4 (1957), 57—59;

Ульман (J. L. Ullman), *On a Theorem of Frobenius*, *Michigan Math J.* 1 (1952) 189—193

Дальнейшие ссылки имеются в работе

Путнам (C. R. Putnam), *Commutators on a Hilbert Space—On Bounded Matrices with Non-negative Elements* Purdue University, PRF-1421, May 1958,

где намечен вывод теоремы Перрона как следствия теоремы Виванти—Принсгейма о степенных рядах. Этот результат был независимо открыт Уинтнером, Боненбластом—Карлином, Маллиkenом—Сноу и Беллманом.

Были предложены различные итеративные схемы для нахождения наибольшего собственного значения. См. цитированную выше работу Брауэра, а также работы

Брауэр (A. Brauer), A Method for the Computation for the Greatest Root of a Positive Matrix, University of North Carolina, December, 1957;

Беллман (R. Bellman), An Iterative Procedure for Obtaining the Perron Root of a Positive Matrix, Proc. Amer. Math. Soc. 6 (1955), 719—725

§ 6. Этот метод был также использован в работе

Беллман и Данскин (R. Bellman and J. M. Danskin), A Survey of the Mathematical Theory of Time-lag, Retarded Control, and Hereditary Processes, The RAND Corporation, Rept. R-256, March. 1. 1954

для исследования задачи, сформулированной в виде упражнения (упражнение 3 к гл. 16).

§ 8. Такое представление для $\lambda_N(A)$, по-видимому, впервые появилось в статье

Коллатц (L. Collatz), Einschliessungssatz für die charakteristischen Zahlen von Matrizen, Math. Z. 48 (1946), 221—226.

Строго этот результат был впервые изложен в работе

Виландт (H. Wielandt), Unzerlegbare, nicht negative Matrizen, Math. Z. 52 (1950), 642—648

Этот результат был несколько раз переоткрыт, в частности, Боненбластом, а также фон Нейманом.

§ 13. См. цитированную выше монографию Харриса, а также статьи

Беллман, Калаба, Уинг (R. Bellman, R. E. Kalaba and G. M. Wing), On the Principle of Invariant Imbedding and One-dimensional Neutron Multiplication, Proc. Nat. Acad. Sci. U. S. A. 43 (1957), 517—520;

Беллман, Калаба и Уинг (R. Bellman, R. E. Kalaba and G. M. Wing), On the Principle of Invariant Imbedding and Neutron Transport Theory, I: One-dimensional Case, J. Math. Mech. 7 (1958), 149—162.

§ 14. Эта модель детально исследована методами динамического программирования в книге

Беллман (R. Bellman), Dynamic Programming, Princeton University Press, Princeton, N. J., 1957. [Русский перевод: Динамическое программирование, ИЛ, 1960.]

§ 15. См. цитированные в комментарии к § 4 статьи Уонга и Вудбери.

§ 18. См. цитированную выше работу Купманса, а также другие работы по линейному программированию, опубликованные в Annals of Mathematics.

Блестящее изложение теории линейных неравенств читатель найдет в работе

Гуд (R. A. Good), Systems of Linear Relations, SIAM Review 1 (1959) 1—31.

Где имеются дальнейшие указания на литературу.

§ 19. Монография

фон Нейман и Morgenstern (J. von Neumann and O. Morgenstern), *The Theory of Games and Economic Behavior*, Princeton University Press, Princeton, N. J., 1948

представляет собой классическую работу по теории игр. Книга

Вильямс (J. D. Williams), *The Compleat Strategyst*, McGraw-Hill Book Company, Inc., New York, 1955 [русский перевод: Совершенный стратег, «Советское радио», 1960]

может служить очень забавным введением в теорию игр.

§ 20. См. гл. 11 книги Беллмана, приведенной в комментарии к § 14, а также статью

Беллман (R. Bellman), *A Markovian Decision Processes*, J. Math. Mech. 6 (1957), 679—684.

§ 21. Другая мотивировка дана в работе

Беллман (R. Bellman), *On a Quasi-Linear Equation*, Canad. J. Math. 6 (1956), 198—202.

Интересное обобщение положительности с удивительными разветвлениями представляет собой понятие *областей положительности* данной матрицы A . Мы говорим, что R является областью положительности матрицы A , если для любой пары $x, y \in R$ имеет место неравенство $(x, Ay) > 0$. См. по этому поводу работы

Кохер (M. Koecher), *Die Geodatischen von Positivitätsbereichen*, Math. Ann. 135 (1958), 192—202;

Кохер (M. Koecher), *Positivitätsbereiche im R^m* , Amer. J. Math. 79 (1957), 575—596;

Бохнер (S. Bochner), *Bessel Functions and Modular Relations of Higher Type and Hyperbolic Differential Equations* Comm. Sem. Math. Univ. Lund, Suppl., 1952, 12—20;

Бохнер (S. Bochner), *Gamma Factors in Functional Equations*, Proc. Nat. Acad. U. S. A. 42 (1956), 86—89

Другим интересным и важным обобщением понятия положительного преобразования являются *преобразования, уменьшающие вариацию*. Такие преобразования, естественно, возникают в самых различных разделах анализа. См. работы

Шёнберг (I. J. Schoenberg), *On Smoothing Functions and Their Generating Functions*, Bull. Amer. Math. Soc. 59 (1953), 199—230;

Гантмахер Ф. Р. и Крейн М. Г., *Sur les matrices complement non-negatives et oscillatoires*, Comp. math. 4 (1937), 445—476;

Поля и Шёнберг (G. Polya and F. J. Schoenberg), *Remarks on de la Vallée Poussin Means and Convex Maps of the Circle*, Pacific J. Math. 8 (1958), 201—212;

Карлин (S. Karlin), *Polya-type Distributions*, IV, Ann. Math. Stat. 29 (1958), 1—21;

Карлин и Мак-Грегор (S. Karlin and J. L. MacGregor), *The Differential Equations of Birth-and-Death Processes and the Stieltjes Moment Problem*, Trans. Amer. Math. Soc. 85 (1957), 489—546.

См. также работу

Слепян и Вайнберг (P. Slepian and L. Weinberg), *Synthesis Applications of Paramount and Dominant Matrices*, Hughes Research Laboratories, Culver City, California.

Упомянем, наконец, работу

Карпилевич Ф. И., О характеристических корнях матриц с неотрицательными элементами, *ИАН СССР, сер. матем.* 15 (1951), 361—383,

посвященную исследованию областей локализации характеристических чисел неотрицательных матриц.

ЛИНЕЙНЫЕ УРАВНЕНИЯ И РАНГ

§ 1. Введение. Мы хотим напомнить некоторые элементарные результаты, касающиеся определителей и их роли в решении систем линейных уравнений вида

$$\sum_{j=1}^N a_{ij}x_j = c_i, \quad i = 1, 2, \dots, M. \quad (1)$$

Коэффициенты a_{ij} и c_i предполагаются комплексными.

Мы получим также условия существования решений у однородной системы линейных уравнений вида

$$\sum_{j=1}^N a_{ij}x_j = 0, \quad i = 1, 2, \dots, M, \quad (2)$$

отличных от $x_1 = x_2 = \dots = x_N = 0$. Это последнее решение называют *тривиальным*.

Для нас наиболее важен случай $M = N$. Однако мы будем изучать его в рамках общей системы.

В заключение мы введем функцию от коэффициентов системы (2), называемую *рангом*, и изучим некоторые ее свойства.

§ 2. Определители. Предполагается, что читатель знаком с элементарными свойствами определителей, а также с той ролью, которую они играют в решении линейных систем уравнений. В частности, нам потребуются следующие два результата.

Лемма 1. *Определитель, имеющий две одинаковые строки или два одинаковых столбца, равен нулю.*

Этот результат является следствием леммы 2.

Лемма 2. *Определитель*

$$|a_{ij}| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{vmatrix} \quad (1)$$

может быть разложен по элементам первой строки:

$$|a_{ij}| = a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1N}A_{1N}, \quad (2)$$

где A_{ij} — определители порядка $N-1$, получаемые вычеркиванием i -й строки и j -го столбца и умножением на $(-1)^{i+j}$.

Числа A_{ij} называются алгебраическими дополнениями a_{ij} .

§ 3. Свойство алгебраических дополнений. Комбинируя утверждения лемм 1 и 2, мы приходим к заключению, что алгебраические дополнения A_{ij} , $j=1, 2, \dots, N$, удовлетворяют следующей системе линейных уравнений:

$$\sum_{j=1}^N a_{ij}A_{ij} = 0, \quad i=2, \dots, N. \quad (1)$$

Действительно, равенства (1) означают, что определитель, у которого первая и i -я строки совпадают, равен нулю.

§ 4. Правило Крамера. Если $|a_{ij}| \neq 0$, то решение системы (1.1) можно представить в виде

$$x_i = \frac{\begin{vmatrix} c_1 & a_{12} & \dots & a_{1N} \\ c_2 & a_{22} & \dots & a_{2N} \\ \dots & \dots & \dots & \dots \\ c_N & a_{N2} & \dots & a_{NN} \end{vmatrix}}{|a_{ij}|}, \quad (1)$$

а для x_i справедливы аналогичные выражения, получаемые заменой i -го столбца $|a_{ij}|$ столбцом коэффициентов c_1, c_2, \dots, c_N .

§ 5. Однородные системы. В дальнейшем мы воспользуемся следующим фактом:

Теорема 1. Система линейных уравнений

$$\sum_{j=1}^N a_{ij}x_j = 0, \quad i=1, 2, \dots, M, \quad (1)$$

имеет нетривиальное решение, если $N > M$.

Если $N \leq M$, то необходимым и достаточным условием существования нетривиального решения является равенство нулю всех определителей $|a_{ij}|$ порядка N .

Кроме того, можно найти нетривиальное решение, в котором каждый из x_i представляет собой полином относительно a_{ij} .

Доказательство проведем с помощью элементарной, хотя и громоздкой, индукции. Начнем с первого утверждения. В случае $M=1$ утверждение очевидно. Далее рассмотрим случай $M=2$.

Если $N = 2$, то мы имеем уравнения

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0, \\ a_{21}x_1 + a_{22}x_2 &= 0. \end{aligned} \quad (2)$$

Если $|a_{ij}| = 0$, то набор

$$x_1 = a_{22}, \quad x_2 = -a_{11} \quad (3)$$

является решением системы (2). Если это решение тривиально, т. е. $x_1 = x_2 = 0$, то, воспользовавшись алгебраическими дополнениями коэффициентов второго из уравнений (2), мы получим решение

$$x_1 = a_{11}, \quad x_2 = -a_{12}. \quad (4)$$

Если и это решение тривиально, то все коэффициенты a_{ij} равны нулю, и тогда любой набор значений x_1, x_2 дает решение.

Если $N > 2$, то мы имеем набор уравнений

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1N}x_N &= 0, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2N}x_N &= 0. \end{aligned} \quad (5)$$

Разрешим эти уравнения относительно x_1 и x_2 , выразив их через остальные x_i . Это легко сделать, если определитель

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad (6)$$

отличен от нуля. Если этот определитель равен нулю, то мы попытаемся найти такую пару переменных x_i, x_j , у которой определитель

$$\begin{vmatrix} a_{1i} & a_{1j} \\ a_{2i} & a_{2j} \end{vmatrix} \quad (7)$$

отличен от нуля. Затем мы выразим переменные x_i и x_j через остальные переменные.

Предположим, однако, что все определители вида (7) равны нулю. Мы утверждаем, что в этом случае одно из уравнений (5) является избыточным в том смысле, что любой набор значений x_i , удовлетворяющий одному из этих уравнений, с необходимостью удовлетворяет и другому. Чтобы доказать это, мы покажем, что существуют такие два параметра λ_1 и λ_2 , по крайней мере один из которых не равен нулю, что

$$\begin{aligned} \lambda_1 (a_{11}x_1 + a_{12}x_2 + \dots + a_{1N}x_N) + \\ + \lambda_2 (a_{21}x_1 + a_{22}x_2 + \dots + a_{2N}x_N) = 0 \end{aligned} \quad (8)$$

тождественно по x_i . Если при этом окажется, что $\lambda_1 \neq 0$, то, очевидно, достаточно удовлетворить второму из уравнений (5), чтобы удовлетворялось и первое. Таким образом, мы сведем

приводится к сумме r квадратов:

$$(\mathbf{x}, A\mathbf{x}) = \sum_{i=1}^r b_i y_i^2, \quad (1)$$

где все b_i отличны от нуля, то ранг A , очевидно, равен r .

В дополнение к этому имеет место следующая

Теорема 3. Если действительная квадратичная форма $(\mathbf{x}, A\mathbf{x})$ ранга r парой невырожденных преобразований $\mathbf{x} = T\mathbf{y}$ и $\mathbf{x} = S\mathbf{z}$ приводится к сумме r квадратов вида

$$(\mathbf{x}, A\mathbf{x}) = b_1 y_1^2 + b_2 y_2^2 + \dots + b_r y_r^2 = c_1 z_1^2 + c_2 z_2^2 + \dots + c_r z_r^2, \quad (2)$$

то число положительных коэффициентов b_i равно числу положительных коэффициентов c_i .

§ 9. Сигнатура. Пусть $p(A)$ и $n(A)$ представляют собой соответственно число положительных и отрицательных квадратов в представлении формы $(\mathbf{x}, A\mathbf{x})$. Разность $s(A)$ между числом положительных и отрицательных квадратов $s(A) = p(A) - n(A)$ называется *сигнатурой* формы $(\mathbf{x}, A\mathbf{x})$. Из предыдущего следует, что

$$s(A) = s(T'AT), \quad (1)$$

если матрица T невырожденная.

Упражнение

1. Показать, что $r(AB) + r(BC) \leq r(B) + r(ABC)$ всякий раз, когда существует произведение прямоугольных матриц ABC . (Фробениус.)

Упражнения к приложению А

1. Для любой $N \times N$ -матрицы $A = \|a_{ij}\|$ ранга r имеют место соотношения

$$\sum_{i=1}^N \frac{|a_{ii}|}{\sum_{j=1}^N |a_{ij}|} \leq r, \quad \sum_{i=1}^N \frac{|a_{ii}|^2}{\sum_{j=1}^N |a_{ij}|^2} \leq r,$$

причем неопределенности вида $0/0$, если таковые встречаются, следует полагать равными нулю. (Фань Цзы, Гоффман.)

В статье Фань Цзы и Гоффмана¹⁾ рассматривается задача нахождения нижних границ для ранга матрицы A непосредственно по ее элементам a_{ij} , а также связь между этой задачей и задачей локализации характеристических чисел.

¹⁾ Ky Fan and A. J. Hoffman, Lower Bounds for the Rank and Location of the Eigenvalues of a Matrix, Contributions to the Solution of Systems of Linear Equations and the Determination of Eigenvalues, Nat. Bur. Standards Applied Math., ser. 39, 1954.

2. Пусть $A = B + iC$, где $B = (A + \bar{A})/2$ и $C = -(A - \bar{A})/2i$. Тогда если A — эрмитова неотрицательно определенная матрица, то ранг матрицы B не меньше ранга матрицы C . (Брудер—Смит.)

3. Если матрица A эрмитова и неотрицательно определенная, то ранг матрицы B не меньше ранга матрицы A . (Брудер—Смит.)

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

Блестящее изложение многих увлекательных вопросов, возникающих в такой на первый взгляд «исчерпанной» задаче, как решение линейных систем, имеется в статье

Форсайт (G. E. Forsythe), Solving Linear Equations Can Be Interesting, Bull. Amer. Math. Soc. 59 (1953), 299—329*.

См. также интересную статью

Ланцош (C. Lanczos), Linear Systems in Self-adjoint Form, Amer. Math. Monthly 65 (1958), 665—679.

Связь между рангом и топологическими инвариантами рассматривается в работе

Бурген (D. G. Bourgin), Quadratic Forms, Bull. Amer. Math. Soc. 51 (1945), 907—908.

Ряд других результатов, связанных с понятием ранга, можно найти в работе

Ольденбургер (R. Oldenburger), Expansion of Quadratic Forms, Bull. Amer. Math. Soc. 49 (1943), 136—141.

Обобщение понятия линейной зависимости читатель найдет в работах

Тэтчер (H. C. Thatcher), Generalization of Concepts Related to Linear Dependence, J. Ind. Appl. Math. 6 (1958), 288—300;

Уэйплс (G. Whaples), A Note on Degree- n Independence, *ibid.*, 300—301.

*) Изложение вопросов, затронутых в этом добавлении, см. в книгах: А. Г. Курош, Курс высшей алгебры, «Наука», 1975; Г. Е. Шолов, Математический анализ (конечномерные линейные пространства), «Наука», 1969. (Прим. ред.)

МЕТОД ЭРМИТА

Рассмотрим многочлен $p(x)$ с действительными коэффициентами и попарно различными действительными корнями r_1, r_2, \dots, r_N . Наша задача состоит в определении числа корней, больших, чем данное действительное число a_1 .

Пусть

$$y_i = x_1 + r_i x_2 + r_i^2 x_3 + \dots + r_i^{N-1} x_N, \quad (1)$$

где x_i — действительные числа, и пусть $Q(x)$ обозначает квадратичную форму

$$Q(x) = \sum_{i=1}^N y_i^2 / (r_i - a_1). \quad (2)$$

В силу закона инерции приведение квадратичной формы $Q(x)$ к сумме квадратов любым другим способом с помощью действительного преобразования сохранит число положительных квадратов, которое, очевидно, равно числу действительных корней многочлена $p(x)$, больших, чем a_1 .

Этот результат на первый взгляд кажется неконструктивным, поскольку коэффициенты формы $Q(x)$ зависят от неизвестных корней многочлена $p(x)$. Заметим, однако, что $Q(x)$ представляет собой рациональную симметрическую функцию корней многочлена и, следовательно, рационально выражается через его коэффициенты.

Более интересен случай, когда $p(x)$ имеет комплексные корни. Эрмит рассматривал этот случай следующим образом. Пусть r_1 и r_2 — комплексно сопряженные корни:

$$r_1 = \rho (\cos \alpha + i \sin \alpha), \quad r_2 = \rho (\cos \alpha - i \sin \alpha),$$

и пусть

$$y_1 = u + iv, \quad y_2 = u - iv, \quad (3)$$

где u и v — вещественные линейные формы от x_i . Пусть далее

$$\frac{1}{r_1 - a_1} = r (\cos \psi + i \sin \psi), \quad \frac{1}{r_2 - a_1} = r (\cos \psi - i \sin \psi). \quad (4)$$

Нетрудно проверить, что

$$\frac{u^2}{r_1 - a_1} + \frac{v^2}{r_2 - a_1} = 2r(u \cos \varphi/2 - v \sin \varphi/2)^2 - 2r(u \sin \varphi/2 + v \cos \varphi/2)^2. \quad (5)$$

Теперь сформулируем результат Эрмита.

Теорема. Пусть многочлен $p(x)$ с действительными коэффициентами имеет только простые корни. Пусть форма

$$Q(x) = \sum_{i=1}^N y_i^2 / (r_i - a_i) \quad (6)$$

приведена к сумме квадратов с помощью действительной невырожденной подстановки. Тогда число положительных квадратов формы равно числу пар не вещественных корней уравнения $p(x) = 0$ плюс число действительных корней, больших a_1 .

Некоторые относящиеся сюда результаты Сильвестра рассматриваются в книге по теории уравнений Бернсайда и Пэнтон, на которую мы уже ссылались. Здесь мы следовали их изложению*).

Обобщение метода Эрмита было дано Гурвицем¹⁾ в его классической работе, где даны необходимые и достаточные условия отрицательности действительных частей корней многочлена.

Другие методы и ссылки на литературу читатель найдет в работе Беллмана, Гликсберга и Гросса²⁾.

В этом приложении мы хотели подчеркнуть эффективность квадратичных форм как «орудия счета». Другую иллюстрацию этой роли квадратичных форм читатель найдет в работе Эрмита³⁾ по приведению квадратичных форм.

* См. также Ф. Г. Гантмахер, Теория матриц, «Наука», 1967, 493. (Прим. ред.)

¹⁾ A. Hurwitz, Über die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt, Math. Ann. 46 (1895) (Werke, vol. 2, 533—545).

²⁾ R. Bellman, I. Glicksberg and O. Gross, Some Aspects of the Mathematical Theory of Control Processes, The RAND Corporation, Rept. R-313, January 16, 1958.

³⁾ C. Hermite, Oeuvres, 284—287; 396—414; J. Crelle, Bd. LII, 1856, 39—51

**МОМЕНТЫ
И КВАДРАТИЧНЫЕ ФОРМЫ**

§ 1. Введение. Мы будем предполагать, что читатель знаком с понятием интеграла Римана—Стилтьеса. Кроме того, чтобы завершить оба приводимые ниже доказательства, нам придется воспользоваться теоремой, принадлежащей Хелли.

§ 2. Обозначения. Рассмотрим последовательность величин $\{a_k\}$, $k=0, 1, \dots$, определенных интегралами

$$a_k = \int_0^1 t^k dg(t), \quad (1)$$

где $g(t)$ —монотонно возрастающая на отрезке $[0, 1]$ функция, причем $g(1)=1$, а $g(0)=0$. Эти величины называются моментами весовой функции g (или функции распределения).

Стилтьес указал на то, что квадратичные формы

$$Q_N(x) = \sum_{k, l=0}^N a_{k+l} x_k x_l = \int_0^1 \left(\sum_{k=0}^N x_k t^k \right)^2 dg(t) \quad (2)$$

с необходимостью являются неотрицательно определенными. Будем называть матрицы

$$A_N = \|a_{k+l}\| \quad k, l=0, 1, \dots, N, \quad (3)$$

ганкелевыми матрицами.

Критерии, выведенные нами в § 2 гл. 5, дают в этом случае ряд интересных неравенств вида

$$\left| \begin{array}{cc} \int_0^1 t^k dg & \int_0^1 t^{k+l} dg \\ \int_0^1 t^{k+l} dg & \int_0^1 t^{k+2l} dg \end{array} \right| \geq 0, \quad (4)$$

причем если g имеет достаточно много точек роста, то справедливо строгое неравенство.

Если вместо моментов (1) мы рассмотрим тригонометрические моменты

$$c_k = \int_0^1 e^{2\pi i k t} dg, \quad k = 0, \pm 1, \pm 2, \dots, \quad (5)$$

то полученная по аналогии квадратичная форма

$$\int_0^1 \left| \sum_{k=1}^N x_k e^{2\pi i k t} \right|^2 dg \quad (6)$$

называется *тёплицевой формой*, а соответствующая матрица — *тёплицевой матрицей*.

Очевидно, что необходимое условие того, что последовательность $\{c_k\}$ задается формулами (5) при некоторой монотонно возрастающей функции $g(t)$, удовлетворяющей условиям $g(1) = 1$, $g(0) = 0$, состоит в том, что матрица

$$C = \{c_{k-l}\} \quad (7)$$

является неотрицательно определенной эрмитовой матрицей. Возникает вопрос: является ли это условие достаточным^{*}). Как мы сейчас увидим, в одном случае это так, в другом — нет.

§ 3. Метод Фишера. Рассмотрим бесконечную последовательность действительных чисел $\{a_n\}$, удовлетворяющую следующему условию: квадратичная форма

$$Q_N(x) = \sum_{i, j=0}^N a_{i+j} x_i x_j \quad (1)$$

положительно определена для каждого N .

Теорема 1. Если $2N+1$ действительных чисел a_0, a_1, \dots, a_{2N} таковы, что форма $Q_N(x)$ положительно определена, то эти числа можно представить в виде

$$a_k = b_0 r_0^k + b_1 r_1^k + \dots + b_N r_N^k, \quad k = 0, 1, \dots, 2N, \quad (2)$$

где $b_i > 0$, а числа r_i действительные и различные.

Доказательство. Рассмотрим производную квадратичную форму

$$P_N(x) = \sum_{i, j=0}^N a_{i+j} x_i x_j \quad (3)$$

^{*}) По поводу проблемы моментов см. Н. И. Ахнезер, Классическая проблема моментов, Физматгиз, 1961. (Прим. ред.)

Поскольку по предположению $Q_N(\mathbf{x})$ положительно определена, мы можем одновременно привести формы $P_N(\mathbf{x})$ и $Q_N(\mathbf{x})$ к виду

$$Q_N(\mathbf{x}) = \sum_{i=0}^N (m_{i0}x_0 + \dots + m_{iN}x_N)^2, \quad (4)$$

$$P_N(\mathbf{x}) = \sum_{j=0}^N r_j (m_{j0}x_0 + \dots + m_{jN}x_N)^2,$$

где r_j — действительные числа.

Из представления $Q_N(\mathbf{x})$ имеем

$$a_{i+(j+1)} = m_{0i}m_{0,j+1} + \dots + m_{Ni}m_{N,j+1}, \quad (5)$$

$$i=0, 1, \dots, N, \quad j=0, 1, \dots, N-1.$$

С другой стороны, из представления $P_N(\mathbf{x})$ получаем

$$a_{i+i+j} = r_0 m_{0i}m_{0j} + \dots + r_N m_{Ni}m_{Nj}. \quad (6)$$

Сравнивая (5) и (6), получаем систему линейных однородных уравнений

$$m_{0i}(m_{0,j+1} - r_0 m_{0j}) + \dots + m_{Ni}(m_{N,j+1} - r_N m_{Nj}) = 0, \quad (7)$$

$$i=0, 1, \dots, N.$$

Поскольку определитель $|m_{ij}|$ отличен от нуля, из (7) получаем соотношения

$$m_{k,j+1} = r_k m_{kj}, \quad k=0, \dots, N, \quad (8)$$

или

$$m_{kj} = r_k^j m_{k0}. \quad (9)$$

Следовательно, форму $Q_N(\mathbf{x})$ можно записать в виде

$$Q_N(\mathbf{x}) = \sum_{i=0}^N m_{i0}^2 (x_0 + r_i x_1 + \dots + r_i^N x_N)^2, \quad (10)$$

откуда следует искомое представление (2). Поскольку $Q_N(\mathbf{x})$ положительно определена, то $r_i \neq r_j$ при $i \neq j$ и $m_{i0}^2 > 0$.

§ 4. Моментное представление. Используя теорему 1, легко доказать теорему 2.

Теорема 2. Если действительная последовательность $\{a_n\}$ удовлетворяет условиям:

форма $\sum_{i,j=0}^N a_{i+j} x_i x_j$ положительно определена при всех N , (1a)

форма $\sum_{i,j=0}^N (a_{i+j} \pm a_{i+j+1}) x_i x_j$ положительно определена при всех N , (1б)

то эта последовательность является моментной, т. е. существует такая ограниченная монотонно возрастающая функция $g(t)$, определенная на отрезке $[-1, 1]$, что

$$a_n = \int_{-1}^1 t^n dg(t). \quad (2)$$

Доказательство. Из представления, полученного в теореме 1, и условия (16) следует, что $|r_t| \leq 1$, $t = 0, 1, \dots, N^*$. Следовательно, мы можем найти монотонно возрастающую функцию $g_N(t)$, ограниченную числом a_0 и такую, что

$$a_n = \int_{-1}^1 t^n dg_N(t), \quad n = 0, 1, \dots, 2N. \quad (3)$$

Чтобы завершить доказательство, мы воспользуемся теоремой Хелли, которая гарантирует нам существование монотонно возрастающей функции $g(t)$, ограниченной числом a_0 и обладающей тем свойством, что

$$\lim_{M \rightarrow \infty} \int_{-1}^1 t^n dg_M(t) = \int_{-1}^1 t^n dg(t), \quad (4)$$

где M пробегает бесконечную последовательность значений. При этом при всех n выполняются равенства (2).

§ 5. Результат Герглотца. С помощью таких же рассуждений можно доказать следующий результат, установленный Герглотцем.

Теорема 3. Для того чтобы последовательность комплексных чисел $\{a_n\}$, $n = 0, \pm 1, \dots$, $a_{-n} = \bar{a}_n$, удовлетворяла условию

$$\sum_{h=1}^N \sum_{k=1}^N a_{h-k} x_h \bar{x}_k \geq 0 \quad (1)$$

при всех комплексных $\{x_n\}$ и всех N , необходимо и достаточно существование действительной неубывающей функции $u(y)$ ограниченной вариации такой, что

$$a_n = \int_0^{2\pi} e^{in\psi} du(\psi). \quad (2)$$

Метод Фишера читатель найдет в его статье¹⁾.

¹⁾ Имеем $a_k > a_{k+1}$ в силу положительной определенности формы (16). (Прим. ред.)

¹⁾ E. Fischer, Über die Caratheodorysche Problem, Potenzreihen mit positiven reellen Teil betreffend, Rend. circ. mat. Palermo 32 (1911), 240—256.

БИБЛИОГРАФИЯ И КОММЕНТАРИЙ

Квадратичные формы типа $\sum_{i,j=0}^N a_{i-j} x_i x_j$ называют еще L -формами, подчеркивая тем самым связь между этими формами и рядами Лорана аналитических функций. Эти формы играют чрезвычайно важную роль в анализе. См., например, следующие статьи:

Сас (O. Szasz), Über harmonische Funktionen und L -formen, Math. Z. 1 (1918), 149—162;

Сас (O. Szasz), Über Potenzreihen und Bilinearformen, Math. Z. 4 (1919), 163—176;

Фейер (L. Fejer), Über trigonometrische Polynome, J. Math. 146 (1915), 63—82;

Шур (I. Schur), Bemerkungen zur Theorie der Beschränkten Bilinearformen mit unendlich vielen Veränderlichen, J. Math. 140 (1911), 1—28;

Каратеодори (A. Carathéodory), Über den Zusammenhang der Extremen..., Rend. circ. mat. Palermo 32 (1911), 1—22.

Дальнейшие ссылки можно найти в монографии

Беккенбах и Беллман (E. F. Beckenbach and R. Bellman), Inequalities, Ergeb. Math., 1960 [русский перевод: Неравенства, «Мир», 1965],

а также в обширной библиографии к статье

Кац, Мёрдок и Сеге (M. Kac, L. Murdock and G. Szegő), On the Eigenvalues of Certain Hermitian Forms, J. Rat. Mech. Analysis 2 (1953), 767—800.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Алгебра кронекеровских произведений 256

Алгебраическое дополнение 48, 335

Биортогональность 237

Вариации связанные 98

Вектор 33

— вероятностный 285

— инвариантный 43

—, компоненты 34

— неотрицательный 283, 309

— нулевой 35

— положительный 285, 309

—, размерность 34

— собственный 56

— столбец 34

— строка 34

—, умножение на скаляр 35

Векторы, линейная зависимость 65

—, — независимость 65

—, ортогональность 36

— ортонормированные 36

—, ортонормированный набор 67

—, равенство 34

—, сложение 34

Задача наименьшего отклонения 167

Закон инерции Якоби — Сильвестра 339

Коммутатор 49

Корень из матрицы 116

Критерий Гурвица 273, 277

— Сильвестра 95

Кронекеровская производная от матрицы 266

— сумма 259

Кронекеровский логарифм 258

Линейное программирование 321

Линейные неравенства 321

Линейные однородные уравнения решение 54

— — —, — нетривиальное 54

— — —, — тривиальное 334

— уравнения с периодическими коэффициентами 231

Математическая экономика 316

Матрица 37

— Адамара 327

— ассоциированная 263

— вещественная 38

— вырожденная (особенная) 113

— ганкелева 344

— двойко стохастическая 298

—, диагонализация 217

— диагональная 40, 60

— доминантная 327

— единичная 38

— идемпотентная 90

—, каноническая форма Жордана 220, 221

— квадратная 37

— клеточная 106

— комплексно сопряженная 38

— кососимметрическая 61

— косоэрмитова 63

— Лоренца 91

— марковская 285, 292

— Минковского — Леонтьева 319

— невырожденная (неособенная) 113

— неотрицательная 283

— нормальная 226

— нулевая 38

—, область положительности 332

— обратная 113

— ортогональная 46

— платежей 322

— положительная 283, 308

— положительно определенная 62

— прямоугольная 105

—, размер (порядок) 37

— симметризуемая 90

— симметрическая 45

- Матрица симметрическая, диагонализация 75
 — —, каноническая форма 54
 — —, параметрическое представление элементов 117
 — —, ранг 145
 —, след 90, 118
 — случайная 302
 — —, среднее значение 303
 —, собственное значение 56
 — транспонированная 44, 106
 — треугольная 40
 — — форма 224
 — унитарная 47
 —, условия устойчивости 275, 277
 — устойчивая 271
 —, характеристическое уравнение 56, 83
 —, — число 56
 —, — — кратное 56, 219, 221
 —, — — простое 56
 —, элемент 37
 — эрмитова 46
 — Якоби 50, 128, 173, 175
 Матрицы коммутативные 41
 —, одновременное приведение к диагональной форме 78
 —, равенство 37
 — слабо связанные 177
 Метод Ляпунова 271
 — ортогонализации Грама — Шмидта 65
 — Фишера 345
 — Эйлера 212
 — Эрмита 342
 Многочлен устойчивый 274
 Моменты 344

 Неравенство Адамара 154
 — Бергстрема 154
 — Гельдера 150
 — Коши — Шварца 149
 — треугольника 36
 — Шура 161
 Норма вектора 190
 — матрицы 190
 Нормирование решения 26

 Оператор сопряженный 53
 Определитель Вандермонда 215
 — Вронского 243
 — Грама 67
 — Касорати 244
 — Коши 216

 Отклонение среднеквадратичное 273
 Отношение Релея 135

 Полиномы Лагерра 68
 — Лежандра 68
 — Эрмита 68
 Правило Крамера 77, 335
 Преобразование Гаусса 49
 — дробно-линейное 40
 — индуцированное 45
 — квадратичное 296
 — Кэли 115
 — Лапласа 237
 — — в матричном случае 240
 — линейное однородное 296
 — Лоренца 91
 — «растягивающее» 81
 — сопряженное 45, 53
 —, уменьшающее вариацию 332
 — унитарное 295
 — Фурье конечное 49
 — Шура 297
 Принцип суперпозиции 213
 Проблема Штурма — Лиувилля 170
 Произведение кронекеровское 256
 — матриц, ассоциативность 42
 — —, некоммутативность 41
 — скалярное векторов 35
 — — обобщенное 90
 Процесс марковский дискретный 286
 — — —, асимптотическое поведение 286
 — — — принятия решений 323
 — — стохастический непрерывный 293
 — — простой 283

 Ранг квадратичной формы 339
 — матрицы 339
 Рост, непрерывные процессы 314
 — стационарный 313
 Ряд матричный 191

 Символ Кронекера 39
 Система неустойчивая 269
 — сопряженная 199
 — устойчивая 269
 — —, необходимые и достаточные условия 270
 Скаляр 34
 Скобки Якоби 243
 Спектр 64

 Теорема Гамильтона — Кэли 229, 247
 — — — для симметрических матриц

- Теорема Герглота 347
 — Куранта — Фишера 137
 — об аппроксимации вторая 229
 — — — первая 228
 — о представлении 142
 — отделения Пуанкаре 142, 145
 — — Штурма 140
 — Перрона 309
 — —, аналог 315
 — —, обобщение 334
 — — о приведении к треугольному виду 243
 — Пуанкаре — Ляпунова 280
 — Стильеса 182
 — Финслера 98
 — фон Неймана 323
 — Хелли 347
 — Шура 117, 126, 224
 Теория возмущений 82, 83, 199
 — игр 322
 Тождество Лагранжа 70, 156
 — Якоби 196
 Точка седловая 23
 — стационарная 21
- Умножение вектора на матрицу 38
 — — — скаляр 35
 — матрицы на матрицу 39
 — — — скаляр 38
- Уравнение вариационное 143
- Уравнение вариационное, приближенный метод решения 143
 — Матье 231
 — Риккати 242
 — функциональное Поля 202
- Форма билинейная 50
 — квадратичная 23
 — — неопределенная 29
 — — неотрицательная 29
 — —, одновременное приведение к сумме квадратов 80
 — — отрицательно определенная 29
 — — положительная 29
 — — положительно определенная 29, 62
 — —, представление в виде суммы квадратов 98
 — —, приведение к каноническому виду 60
 — —, сигнатура 340
 — тёплицева 345
- Экспонента матричная 194
 — —, функциональные уравнения 194, 195
- Якобиан 50

Ричард Беллман

ВВЕДЕНИЕ В ТЕОРИЮ МАТРИЦ

Издание 2-е

Перевод с английского
под редакцией *В. В. Лидского*

М., 1976 г., 352 стр. в обл.

Редактор *В. В. Донченко*

Техн редактор *Н. В. Кошелева*

Корректоры *О. А. Бутусова, Н. Б. Румянцева*

Сдано в набор 30/XII 1975 г. Подписано к печати 5/IV 1976 г.

Бумага 60×90^{1/16}.

Физ. печ. л. 22. Условн. печ. л. 22.

Уч.-изд. л. 21,36.

Тираж 28 000 экз.

Цена книги 1 р. 63 к.

Заказ № 3708

Издательство «Наука»

Главная редакция физико-математической литературы

117071, Москва, В-71, Ленинский проспект, 15

Ордена Трудового Красного Знамени

Первая Образцовая типография имени А. А. Жданова

Союзполиграфпрома при Государственном комитете

Совета Министров СССР по делам издательства,

полиграфии и книжной торговли.

Москва, М-54, Валуевая, 28