Гантмахер Ф.Р.

Теория матриц

МОСКВА
ФИЗМАТЛИТ ®

УДК 512.83
ББК 517.1
Г 19

Гантмахер Ф. Р. **Теория матриц.** — 5-е изд., — М.: ФИЗМАТЛИТ, 2010. — 560 с. — ISBN 978-5-9221-0524-8.

Книга посвящена матричному исчислению. В ней наряду с собственно теорией матриц содержится изложение ряда математических проблем, решение которых достигается применением развитой матричной техники. Большое внимание уделяется вопросам интегрирования и проблеме устойчивости систем дифференциальных уравнений.

Четвертое издание — 1988 г.

Для студентов старших курсов и аспирантов (математиков, механиков, физиков и др.), а также для математиков, программистов, механиков, физиков и инженеров, использующих матричный математический аппарат.

Ответственный редактор:

В. Б. Лидский

Ил. 11. Библиогр. 302 назв.

ISBN 978-5-9221-0524-8

ОГЛАВЛЕНИЕ

Предисловие автора к первому изданию 7
Предисловие редактора ко второму изданию 10

ЧАСТЬ ПЕРВАЯ
ОСНОВЫ ТЕОРИИ

ГЛАВА I. **МАТРИЦЫ И ДЕЙСТВИЯ НАД НИМИ**

§ 1. Матрицы. Основные обозначения 11
§ 2. Сложение и умножение прямоугольных матриц 13
§ 3. Квадратные матрицы 22
§ 4. Ассоциированные матрицы. Миноры обратной матрицы 27
§ 5. Обращение прямоугольных матриц. Псевдообратная матрица 30

ГЛАВА II. **АЛГОРИТМ ГАУССА И НЕКОТОРЫЕ ЕГО ПРИМЕНЕНИЯ**

§ 1. Метод исключения Гаусса 39
§ 2. Механическая интерпретация алгоритма Гаусса 43
§ 3. Детерминантное тождество Сильвестра 45
§ 4. Разложение квадратной матрицы на треугольные множители 47
§ 5. Разбиение матрицы на блоки. Техника оперирования с блочными матрицами. Обобщенный алгоритм Гаусса 53

ГЛАВА III. **ЛИНЕЙНЫЕ ОПЕРАТОРЫ В n-МЕРНОМ ВЕКТОРНОМ ПРОСТРАНСТВЕ**

§ 1. Векторное пространство 63
§ 2. Линейный оператор, отображающий n-мерное пространство в m-мерное . . . 67
§ 3. Сложение и умножение линейных операторов 69
§ 4. Преобразование координат 71
§ 5. Эквивалентные матрицы. Ранг оператора. Неравенства Сильвестра 72
§ 6. Линейные операторы, отображающие n-мерное пространство само в себя . . . 76
§ 7. Характеристические числа и собственные векторы линейного оператора . . . 79
§ 8. Линейные операторы простой структуры 81

ГЛАВА IV. **ХАРАКТЕРИСТИЧЕСКИЙ И МИНИМАЛЬНЫЙ МНОГОЧЛЕНЫ МАТРИЦЫ**

§ 1. Сложение и умножение матричных многочленов 84
§ 2. Правое и левое деления матричных многочленов. Обобщенная теорема Безу . 86
§ 3. Характеристический многочлен матрицы. Присоединенная матрица 89
§ 4. Метод Д.К. Фаддеева одновременного вычисления коэффициентов характеристического многочлена и присоединенной матрицы 93
§ 5. Минимальный многочлен матрицы 95

ГЛАВА V. **ФУНКЦИИ МАТРИЦЫ**

§ 1. Определение функции матрицы 99
§ 2. Интерполяционный многочлен Лагранжа–Сильвестра 103
§ 3. Другие формы определения $f(A)$. Компоненты матрицы A 106
§ 4. Представление функций матриц рядами 111
§ 5. Некоторые свойства функций от матриц 114

§ 6. Применение функций от матрицы к интегрированию системы линейных дифференциальных уравнений с постоянными коэффициентами 119
§ 7. Устойчивость движения в случае линейной системы 125

ГЛАВА VI. **ЭКВИВАЛЕНТНЫЕ ПРЕОБРАЗОВАНИЯ МНОГОЧЛЕННЫХ МАТРИЦ. АНАЛИТИЧЕСКАЯ ТЕОРИЯ ЭЛЕМЕНТАРНЫХ ДЕЛИТЕЛЕЙ**

§ 1. Элементарные преобразования многочленной матрицы 130
§ 2. Канонический вид λ-матрицы 133
§ 3. Инвариантные многочлены и элементарные делители многочленной матрицы 137
§ 4. Эквивалентность линейных двучленов 142
§ 5. Критерий подобия матриц 144
§ 6. Нормальные формы матрицы 145
§ 7. Элементарные делители матрицы $f(A)$ 149
§ 8. Общий метод построения преобразующей матрицы 152
§ 9. Второй метод построения преобразующей матрицы 156

ГЛАВА VII. **СТРУКТУРА ЛИНЕЙНОГО ОПЕРАТОРА В n-МЕРНОМ ПРОСТРАНСТВЕ (ГЕОМЕТРИЧЕСКАЯ ТЕОРИЯ ЭЛЕМЕНТАРНЫХ ДЕЛИТЕЛЕЙ)**

§ 1. Минимальный многочлен вектора, пространства (относительно заданного линейного оператора) 165
§ 2. Расщепление на инвариантные подпространства с взаимно простыми минимальными многочленами 167
§ 3. Сравнения. Надпространство 169
§ 4. Расщепление пространства на циклические инвариантные подпространства . 171
§ 5. Нормальная форма матрицы 175
§ 6. Инвариантные многочлены. Элементарные делители 178
§ 7. Нормальная жорданова форма матрицы 181
§ 8. Метод А.Н. Крылова преобразования векового уравнения 183

ГЛАВА VIII. **МАТРИЧНЫЕ УРАВНЕНИЯ**

§ 1. Уравнение $AX = XB$ 193
§ 2. Частный случай: $A = B$. Перестановочные матрицы 197
§ 3. Уравнение $AX - XB = C$ 200
§ 4. Скалярное уравнение $f(X) = 0$ 201
§ 5. Матричное многочленное уравнение 202
§ 6. Извлечение корня m-й степени из невырожденной матрицы 205
§ 7. Извлечение корня m-й степени из вырожденной матрицы 208
§ 8. Логарифм матрицы 212

ГЛАВА IX. **ЛИНЕЙНЫЕ ОПЕРАТОРЫ В УНИТАРНОМ ПРОСТРАНСТВЕ**

§ 1. Общие соображения 215
§ 2. Метризация пространства 215
§ 3. Критерий Грама линейной зависимости векторов 218
§ 4. Ортогональное проектирование 220
§ 5. Геометрический смысл определителя Грама и некоторые неравенства 222
§ 6. Ортогонализация ряда векторов 225
§ 7. Ортонормированный базис 230
§ 8. Сопряженный оператор 232
§ 9. Нормальные операторы в унитарном пространстве 235
§ 10. Спектр нормальных, эрмитовых, унитарных операторов 237
§ 11. Неотрицательные и положительно определенные эрмитовы операторы 240
§ 12. Полярное разложение линейного оператора в унитарном пространстве. Формулы Кэли 242
§ 13. Линейные операторы в евклидовом пространстве 246
§ 14. Полярное разложение оператора и формулы Кэли в евклидовом пространстве 252

§ 15. Коммутирующие нормальные операторы . . . 255
§ 16. Псевдообратный оператор . . . 257

ГЛАВА X. КВАДРАТИЧНЫЕ И ЭРМИТОВЫ ФОРМЫ

§ 1. Преобразование переменных в квадратичной форме . . . 259
§ 2. Приведение квадратичной формы к сумме квадратов. Закон инерции . . . 261
§ 3. Метод Лагранжа приведения квадратичной формы к сумме квадратов. Формула Якоби . . . 263
§ 4. Положительные квадратичные формы . . . 268
§ 5. Приведение квадратичной формы к главным осям . . . 271
§ 6. Пучок квадратичных форм . . . 272
§ 7. Экстремальные свойства характеристических чисел регулярного пучка форм 277
§ 8. Малые колебания системы с n степенями свободы . . . 284
§ 9. Эрмитовы формы . . . 288
§ 10. Ганкелевы формы . . . 293

ЧАСТЬ ВТОРАЯ

СПЕЦИАЛЬНЫЕ ВОПРОСЫ И ПРИЛОЖЕНИЯ

ГЛАВА XI. КОМПЛЕКСНЫЕ СИММЕТРИЧЕСКИЕ, КОСОСИММЕТРИЧЕСКИЕ И ОРТОГОНАЛЬНЫЕ МАТРИЦЫ

§ 1. Некоторые формулы для комплексных ортогональных и унитарных матриц . 301
§ 2. Полярное разложение комплексной матрицы . . . 305
§ 3. Нормальная форма комплексной симметрической матрицы . . . 307
§ 4. Нормальная форма комплексной кососимметрической матрицы . . . 309
§ 5. Нормальная форма комплексной ортогональной матрицы . . . 314

ГЛАВА XII. СИНГУЛЯРНЫЕ ПУЧКИ МАТРИЦ

§ 1. Введение . . . 318
§ 2. Регулярный пучок матриц . . . 319
§ 3. Сингулярные пучки. Теорема о приведении . . . 321
§ 4. Каноническая форма сингулярного пучка матриц . . . 326
§ 5. Минимальные индексы пучка. Критерий строгой эквивалентности пучков . . 328
§ 6. Сингулярные пучки квадратичных форм . . . 330
§ 7. Приложения к дифференциальным уравнениям . . . 334

ГЛАВА XIII. МАТРИЦЫ С НЕОТРИЦАТЕЛЬНЫМИ ЭЛЕМЕНТАМИ

§ 1. Общие свойства . . . 337
§ 2. Спектральные свойства неразложимых неотрицательных матриц . . . 339
§ 3. Разложимые матрицы . . . 349
§ 4. Нормальная форма разложимой матрицы . . . 356
§ 5. Примитивные и импримитивные матрицы . . . 360
§ 6. Стохастические матрицы . . . 364
§ 7. Предельные вероятности для однородной цепи Маркова с конечным числом состояний . . . 368
§ 8. Вполне неотрицательные матрицы . . . 376
§ 9. Осцилляционные матрицы . . . 380

ГЛАВА XIV. РАЗЛИЧНЫЕ КРИТЕРИИ РЕГУЛЯРНОСТИ И ЛОКАЛИЗАЦИЯ СОБСТВЕННЫХ ЗНАЧЕНИЙ

§ 1. Критерий регулярности Адамара и его обобщения . . . 387
§ 2. Норма матрицы . . . 390
§ 3. Распространение критерия Адамара на блочные матрицы . . . 392
§ 4. Критерий регулярности Фидлера . . . 394
§ 5. Круги Гершгорина и другие области локализации . . . 395

ГЛАВА XV. **ПРИЛОЖЕНИЯ ТЕОРИИ МАТРИЦ К ИССЛЕДОВАНИЮ СИСТЕМ ЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ**

§ 1. Системы линейных дифференциальных уравнений с переменными коэффициентами. Общие понятия 399
§ 2. Преобразование Ляпунова 402
§ 3. Приводимые системы 403
§ 4. Каноническая форма приводимой системы. Теорема Еругина 405
§ 5. Матрицант 408
§ 6. Мультипликативный интеграл. Инфинитезимальное исчисление Вольтерра . . 412
§ 7. Дифференциальные системы в комплексной области. Общие свойства 416
§ 8. Мультипликативный интеграл в комплексной области 418
§ 9. Изолированная особая точка 422
§ 10. Регулярная особая точка 427
§ 11. Приводимые аналитические системы 439
§ 12. Аналитические функции многих матриц и их применение к исследованию дифференциальных систем. Работы И. А. Лаппо-Данилевского 442

ГЛАВА XVI. **ПРОБЛЕМА РАУСА–ГУРВИЦА И СМЕЖНЫЕ ВОПРОСЫ**

§ 1. Введение 445
§ 2. Индексы Коши 446
§ 3. Алгоритм Рауса 449
§ 4. Особые случаи. Примеры 452
§ 5. Теорема Ляпунова 455
§ 6. Теорема Рауса-Гурвица 459
§ 7. Формула Орландо 464
§ 8. Особые случаи в теореме Рауса–Гурвица 466
§ 9. Метод квадратичных форм. Определение числа различных вещественных корней многочлена 469
§ 10. Бесконечные ганкелевы матрицы конечного ранга 471
§ 11. Определение индекса произвольной рациональной дроби через коэффициенты числители и знаменателя 473
§ 12. Второе доказательство теоремы Рауса–Гурвица 480
§ 13. Некоторые дополнения к теореме Рауса–Гурвица. Критерий устойчивости Льенара и Шипара 483
§ 14. Некоторые свойства многочлена Гурвица. Теорема Стилтьеса. Представление многочленов Гурвица при помощи непрерывных дробей 487
§ 15. Область устойчивости. Параметры Маркова 493
§ 16. Связь с проблемой моментов 496
§ 17. Связь между определителями Гурвица и определителями Маркова 499
§ 18. Теоремы Маркова и Чебышева 501
§ 19. Обобщенная задача Рауса–Гурвица 507

ДОБАВЛЕНИЕ. **НЕРАВЕНСТВА ДЛЯ СОБСТВЕННЫХ И СИНГУЛЯРНЫХ ЧИСЕЛ** (*В. Б. Лидский*)

§ 1. Мажорирующие последовательности 509
§ 2. Неравенства Неймана–Хорна 512
§ 3. Неравенства Вейля 516
§ 4. Максимально-минимальные свойства сумм и произведений собственных чисел эрмитовых операторов 518
§ 5. Неравенства для собственных и сингулярных чисел сумм и произведений операторов 524
§ 6. Другая постановка задачи о спектре суммы и произведения эрмитовых операторов 527

Примечания 533

Список литературы 539

Предметный указатель 555

ПРЕДИСЛОВИЕ АВТОРА К ПЕРВОМУ ИЗДАНИЮ

В настоящее время матричное исчисление широко применяется в различных областях математики, механики, теоретической физики, теоретической электротехники и т. д. В то же время ни в советской, ни в иностранной литературе нет книги, которая достаточно полно освещала бы как вопросы теории матриц, так и разнообразные ее приложения. Данная книга представляет собой попытку восполнить этот пробел в математической литературе. В основе книги лежат курсы лекций по теории матриц и ее приложениям, читанные автором в разное время на протяжении последних 17 лет в Московском Государственном университете им. М. В. Ломоносова, в Тбилисском Государственном университете и в Московском физико-техническом институте.

Книга рассчитана не только на математиков (студентов, аспирантов, научных работников), но и на специалистов в смежных областях (физиков, инженеров-исследователей), интересующихся математикой и ее приложениями. Поэтому автор стремился сделать изложение материала возможно более доступным, предполагая у читателя только знакомство с теорией определителей и курсом высшей математики в объеме программы втуза. Лишь отдельные параграфы в последних главах книги требуют дополнительных математических знаний у читателя. Кроме того, автор старался сделать изложение отдельных глав возможно более независимым друг от друга. Так, например, глава V “Функции от матрицы” не опирается на материал, помещенный в главах II и III. В тех же местах главы V, где впервые используются основные понятия, введенные в главе IV, имеются соответствующие ссылки. Таким образом, читатель, уже знакомый с элементами теории матриц, имеет возможность непосредственно приступить к чтению интересующих его глав книги.

Книга состоит из двух частей, содержащих 16 глав.

В главах I и III приводятся первоначальные сведения о матрицах и линейных операторах и устанавливается связь между операторами и матрицами.

В главе II излагаются теоретические основы метода исключения Гаусса и связанных с ним эффективных методов решения системы n линейных уравнений при большом n. В этой же главе читатель знакомится с техникой оперирования с матрицами, разбитыми на прямоугольные “клетки” или “блоки”.

В главе VI вводятся имеющие фундаментальное значение “характеристический” и “минимальный” многочлены квадратной матрицы, “присоединенная” и “приведенная присоединенная” матрицы.

В главе V, посвященной функциям от матрицы, даются самое общее определение и конкретные способы вычисления $f(A)$, где $f(\lambda)$ — функция скалярного аргумента λ, а A — квадратная матрица. Понятие функции от матрицы используется в § 5 и § 6 этой главы для нахождения и полного исследования решения системы линейных дифференциальных уравнений первого порядка с постоянными коэффициентами. Как понятие о функции от матрицы, так и связанное с ним исследование системы линейных дифференциальных уравнений с постоянными коэффициентами первого порядка опираются только на понятие о минимальном многочлене матрицы и не используют (в отличие от обычного изложения) так называемой “теории элементарных делителей”, которая излагается в последующих главах VI и VII.

Первые пять глав охватывают некоторый цикл сведений о матрицах и их применениях. Более глубокие вопросы теории матриц связаны с приведением матрицы к нормальной форме. Это приведение проводится на основе теории элементарных делителей Вейерштрасса. Ввиду важности этой теории в книге даны два ее изложения: аналитическое — в главе VI и геометрическое — в главе VII. Обращаем внимание читателя на § 7 и § 8 главы VI, в которых рассматриваются эффективные методы нахождения матрицы, преобразующей данную матрицу к нормальной форме. В § 8 главы VII подробно исследуется метод акад. А. Н. Крылова для практического вычисления коэффициентов характеристического многочлена.

В главе VIII решаются матричные уравнения некоторых типов. Здесь же рассматривается задача об определении всех матриц, перестановочных с данной, и детально изучаются многозначные функции от матрицы $\sqrt[m]{A}$, $\ln A$.

Главы IX и X посвящены теории линейных операторов в унитарном пространстве и теории квадратичных и эрмитовых форм. Эти главы не опираются на теорию элементарных делителей Вейерштрасса и используют из предыдущего материала лишь основные сведения о матрицах и линейных операторах, изложенные в первых трех главах книги. В § 9 главы X дается приложение теории форм к исследованию главных колебаний системы с n степенями свободы. В § 10 этой же главы приведены тонкие исследования Фробениуса по теории ганкелевых форм. Эти исследования применяются в дальнейшем в главе XV при рассмотрении особых случаев в проблеме Рауса–Гурвица.

Последние пять глав составляют вторую часть книги. В главе XI определяются нормальные формы для комплексных симметрических, кососимметрических и ортогональных матриц и устанавливаются интересные связи этих матриц с вещественными матрицами тех же классов и с унитарными матрицами.

В главе XII излагается общая теория пучков матриц вида $A + \lambda B$, где A и B — произвольные прямоугольные матрицы одних и тех же размеров. Подобно тому как исследование регулярных пучков матриц $A + \lambda B$ проводится на основе теории элементарных делителей Вейерштрасса, изучение сингулярных пучков опирается на теорию минимальных индексов Кронекера, которая является как бы дальнейшим развитием теории элементарных делителей Вейерштрасса. С помощью теории Кронекера (автору кажется, что ему удалось упростить изложение этой теории) в главе XII устанавливается каноническая форма пучка матриц $A + \lambda B$ в самом общем случае. Полученные результаты применяются к исследованию системы линейных дифференциальных уравнений с постоянными коэффициентами.

В главе XIII излагаются замечательные спектральные свойства матриц с неотрицательными элементами и рассматриваются две важные области применений матриц этого класса: 1) однородные цепи Маркова в теории вероятностей и 2) осцилляционные свойства упругих колебаний в механике. Матричный метод исследования однородных цепей Маркова получил свое развитие в работах В. И. Романовского [29] и опирается на тот факт, что матрица переходных вероятностей в однородной цепи Маркова с конечным числом состояний является матрицей с неотрицательными элементами специального типа (“стохастическая матрица”).

Осцилляционные свойства упругих колебаний связаны с другим важным классом неотрицательных матриц — с “осцилляционными матрицами”. Эти матрицы и их приложения были исследованы М. Г. Крейном совместно с автором настоящей книги. В главе XIII изложены только некоторые основные результаты из этой области. Подробное же изложение всего этого материала читатель найдет в монографии [7].

В главе XV собраны приложения теории матриц к системам дифференциальных уравнений с переменными коэффициентами. В этой главе центральное место (§ 5–9) занимают теория мультипликативного интеграла и связанное с ним инфинитезимальное исчисление Вольтерра. Эти вопросы почти совсем не освещены в советской математической литературе. В первых параграфах и в § 11 изучаются приводимые (по Ляпунову) системы в связи с задачей об устойчивости движения и приводятся некоторые результаты Н. П. Еругина. §§ 9–11 относятся к аналитической теории систем дифференциальных уравнений. Здесь выясняется ошибочность основной теоремы Биркгоффа, которую обычно используют для исследования решения системы дифференциальных уравнений в окрестности особой точки, и устанавливается канонический вид решения в случае регулярной особой точки.

В § 12 главы XV в обзорном порядке излагаются некоторые результаты фундаментальных исследований И. А. Лаппо-Данилевского по аналитическим функциям от многих матриц и их применениям к дифференциальным системам.

Последняя глава (XVI) посвящена применениям теории квадратичных форм (и, в частности, ганкелевых форм) к проблеме Рауса–Гурвица об определении числа корней многочлена, лежащих в правой полуплоскости ($\Re z > 0$). В первых параграфах этой главы приводится классическая трактовка вопроса. В § 5 дана теорема А. М. Ляпунова, в которой устанавливается критерий устойчивости, эквивалентный критерию Рауса–Гурвица. Наряду с критерием устойчивости Рауса–Гурвица в § 11 этой главы выводится сравнительно мало известный критерий Льенара и Шипара, в котором число детерминантных неравенств примерно вдвое меньше, нежели в критерии Рауса–Гурвица.

В конце главы XVI показана тесная связь с задачами устойчивости двух замечательных теорем А. А. Маркова и П. Л. Чебышева, которые были получены знаменитыми авторами на основе теории разложения в ряд по убывающим степеням аргумента некоторых непрерывных дробей специального типа. Здесь же дается матричное доказательство этих теорем.

Таков краткий перечень содержания настоящей книги.

В заключение автор приносит свою искреннюю благодарность Д. К. Фаддееву, В. П. Потапову и Д. М. Котелянскому, прочитавшим рукопись книги и сделавшим много существенных замечаний, которые были учтены автором при подготовке книги к печати. Автор выражает также свою благодарность М. Г. Крейну и А. И. Узкову за ценные советы, использованные автором при написании книги.

ПРЕДИСЛОВИЕ РЕДАКТОРА КО ВТОРОМУ ИЗДАНИЮ

Среди существующей литературы по теории матриц монография Ф. Р. Гантмахера занимает общепризнанно одно из лучших мест. Это объясняется систематичностью, широтой рассмотренных вопросов и четкостью изложения. Первое издание этой книги, вышедшее в 1953–1954 гг., было затем переведено на немецкий и английский языки.

В последние годы своей жизни Ф. Р. Гантмахер очень много времени уделил пересмотру и расширению этой книги. Изменения, сделанные им, частично касаются стиля (приведение некоторых терминов в соответствие с новыми традициями, улучшение отдельных доказательств и т. д.). Однако помимо этого было добавлено много нового материала, главным образом во второй, специальной, части книги. Отдельная новая глава XIV ("Различные критерии регулярности и локализация собственных значений") посвящена различным методам приближенного отыскания собственных значений. Добавлены также § 5 гл. V ("Некоторые свойства функций от матриц"), § 17 гл. XVI ("Связь между определителями Гурвица и определителями Маркова") и два параграфа (§ 5 гл. I, § 16 гл. IX) о псевдообратных операторах и матрицах.

Известно, что автор был намерен включить в свою книгу ряд недавно разработанных вопросов, связанных с комбинаторикой собственных значений в алгебре матриц. К этим вопросам относится, в частности, задача о распределении собственных значений суммы и произведения двух матриц, а также известные неравенства Вейля и их обобщения. В настоящем издании соответствующее добавление было написано В. Б. Лидским, которому принадлежит одна из первых работ в этом направлении. В. Б. Лидский также принимал участие в подготовке и редактировании второго издания этой книги.

Можно надеяться, что некоторое увеличение объема книги не затруднит ее чтения, но, напротив, доставит читателям много интересной и ценной информации.

Д. П. Желобенко

ПРЕДИСЛОВИЕ РЕДАКТОРА К ЧЕТВЕРТОМУ ИЗДАНИЮ

Настоящее четвертое издание монографии Ф. Р. Гантмахера "Теория матриц" полностью совпадает со вторым изданием (1966 г.). Несмотря на интенсивное развитие теории матриц и появление в последние десятилетия новых замечательных книг по ее основам и специальным направлениям, монография Ф. Р. Гантмахера и поныне не утратила своего выдающегося значения. Причина этого не только в богатстве идей, но и в плодотворном использовании матричных методов в задачах механики, при интегрировании сингулярных дифференциальных уравнений, в проблеме устойчивости и других важных разделах математики.

Главным образом для пополнения списка литературы в конце книги сделан обзор новых достижений по отдельным вопросам.

В. Б. Лидский

ЧАСТЬ ПЕРВАЯ

ОСНОВЫ ТЕОРИИ

ГЛАВА I

МАТРИЦЫ И ДЕЙСТВИЯ НАД НИМИ

§ 1. Матрицы. Основные обозначения

1. Пусть дано некоторое числовое поле K [1]).

Определение 1. Прямоугольную таблицу чисел из поля K

$$\left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{matrix} \right\| \tag{1}$$

будем называть *матрицей*. Если $m = n$, то матрица называется *квадратной*, а число m, равное n, — ее *порядком*. В общем же случае матрица называется *прямоугольной* (размера $m \times n$) или $m \times n$-матрицей. Числа, составляющие матрицу, называются ее *элементами*.

Обозначение. При двухиндексном обозначении элементов первый индекс всегда указывает номер строки, а второй индекс — номер столбца, на пересечении которых стоит данный элемент.

Наряду с обозначениями матрицы (1) будем употреблять и сокращенное обозначение

$$\|a_{ik}\| \quad (i = 1, 2, ..., m; \quad k = 1, 2, ..., n).$$

Часто матрицу (1) будем обозначать также одной буквой, например, матрица A. Если A — квадратная матрица порядка n, то будем писать $A = \|a_{ik}\|_1^n$. Определитель квадратной матрицы $A = \|a_{ik}\|_1^n$ будем обозначать $|a_{ik}|_1^n$ или $|A|$.

Введем сокращенные обозначения для определителей, составленных из элементов данной матрицы:

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \begin{vmatrix} a_{i_1k_1} & a_{i_1k_2} & \dots & a_{i_1k_p} \\ a_{i_2k_1} & a_{i_2k_2} & \dots & a_{i_2k_p} \\ \dots & \dots & \dots & \dots \\ a_{i_pk_1} & a_{i_pk_2} & \dots & a_{i_pk_p} \end{vmatrix}. \tag{2}$$

Определитель (2) называется *минором* p-го порядка матрицы A, если $1 \leqslant i_1 < < i_2 < ... < i_p \leqslant m$ и $1 \leqslant k_1 < k_2 < ... < k_p \leqslant n$.

$m \times n$-матрица $A = \|a_{ik}\|$ имеет $C_m^p C_n^p$ миноров p-го порядка

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \tag{2'}$$

$$(1 \leqslant i_1 < i_2 < ... < i_p \leqslant m, \quad 1 \leqslant k_1 < k_2 < ... < k_p \leqslant n, \quad p \leqslant m, n).$$

Миноры (2′), у которых $i_1 = k_1$, $i_2 = k_2$, ..., $i_p = k_p$, называются *главными*.

---

[1]) Под *числовым полем* понимают любую совокупность чисел, в пределах которой всегда выполнимы и однозначны четыре операции: сложение, вычитание, умножение и деление на число, отличное от нуля.

Примерами числовых полей могут служить совокупность всех рациональных чисел, совокупность всех действительных чисел или совокупность всех комплексных чисел.

Предполагается, что все встречающиеся в дальнейшем числа принадлежат данному исходному числовому полю.

В обозначениях (2) определитель квадратной матрицы $A = \|a_{ik}\|_1^n$ запишется так:

$$|A| = A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Наибольший из порядков отличных от нуля миноров, порождаемых матрицей, называется *рангом* матрицы. Если r — ранг прямоугольной матрицы A размера $m \times n$, то, очевидно, $r \leqslant m, n$.

Прямоугольную матрицу, состоящую из одного столбца

$$\left\| \begin{matrix} x_1 \\ x_2 \\ \dots \\ x_n \end{matrix} \right\|,$$

мы будем называть *столбцевой* и обозначать так: $(x_1, x_2, \dots, x_n)$.

Прямоугольную матрицу, состоящую из одной строки

$$\|z_1 \; z_2 \; \dots \; z_n\|,$$

мы будем называть *строчной* и обозначать так: $[z_1, z_2, \dots, z_n]$.

Квадратную матрицу, у которой все элементы, расположенные вне главной диагонали, равны нулю,

$$\left\| \begin{matrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{matrix} \right\|$$

мы будем называть *диагональной* и обозначать так: $\|d_i \delta_{ik}\|_1^n$ [2]) или

$$\{d_1, d_2, \dots, d_n\}.$$

Введем еще специальные обозначения для строк и столбцов $m \times n$-матрицы $A = \|a_{ik}\|$. Будем обозначать i-ю строку матрицы A через $a_{i\,\cdot}$, а j-й столбец — через $a_{\cdot\,j}$:

$$a_{i\,\cdot} = [a_{i1}, a_{i2}, \dots, a_{in}], \quad a_{\cdot\,j} = [a_{1j}, a_{2j}, \dots, a_{mj}] \\ (i = 1, \dots, m;\; j = 1, \dots, n). \tag{3}$$

Пусть m величин $y_1, y_2, \dots, y_m$ выражаются линейно и однородно через n других величин $x_1, x_2, \dots, x_n$:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n, \end{aligned} \tag{4}$$

или, в сокращенной записи,

$$y_i = \sum_{k=1}^{n} a_{ik}x_k \quad (i = 1, 2, \dots, m). \tag{4$'$}$$

Преобразование величин $x_1, x_2, \dots, x_n$ в величины $y_1, y_2, \dots, y_m$ при помощи формул (4) называется *линейным преобразованием.* Коэффициенты этого преобразования образуют $m \times n$-матрицу (1). Задание линейного преобразования (4) однозначно определяет матрицу (1) и наоборот.

В следующем параграфе, исходя из свойств линейных преобразований (4), мы определим основные операции над прямоугольными матрицами.

---

[2]) Здесь δ_{ik} — символ Кронекера: $\delta_{ik} = \begin{cases} 1 & (i = k), \\ 0 & (i \neq k). \end{cases}$

§ 2. Сложение и умножение прямоугольных матриц

Определим основные операции над матрицами: сложение матриц, умножение матрицы на число и умножение матриц.

1. Пусть величины $y_1, y_2, ..., y_m$ выражаются через величины $x_1, x_2, ..., x_n$ при помощи линейного преобразования

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \quad (i = 1, 2, ..., m), \tag{5}$$

а величины $z_1, z_2, ..., z_m$ — через те же величины $x_1, x_2, ..., x_n$ при помощи преобразования

$$z_i = \sum_{k=1}^{n} b_{ik} x_k \quad (i = 1, 2, ..., m). \tag{6}$$

Тогда

$$y_i + z_i = \sum_{k=1}^{n} (a_{ik} + b_{ik}) x_k \quad (i = 1, 2, ..., m). \tag{7}$$

В соответствии с этим мы устанавливаем

Определение 2. *Суммой двух прямоугольных матриц* $A = \|a_{ik}\|$ и $B = \|b_{ik}\|$ одинаковых размеров $m \times n$ называется матрица $C = \|c_{ik}\|$ того же размера, элементы которой равны суммам соответствующих элементов данной матрицы:

$$C = A + B,$$

если

$$c_{ik} = a_{ik} + b_{ik} \quad (i = 1, 2, ..., m; \quad k = 1, 2, ..., n).$$

Операция нахождения суммы данных матриц называется *сложением матриц*.

Пример.

$$\left\| \begin{matrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{matrix} \right\| + \left\| \begin{matrix} c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{matrix} \right\| = \left\| \begin{matrix} a_1 + c_1 & a_2 + c_2 & a_3 + c_3 \\ b_1 + d_1 & b_2 + d_2 & b_3 + d_3 \end{matrix} \right\|.$$

Согласно определению 2, складывать можно только прямоугольные матрицы одинаковых размеров.

В силу этого же определения матрица коэффициентов в преобразовании (7) есть сумма матриц коэффициентов в преобразованиях (5) и (6).

Из определения сложения матриц непосредственно следует, что эта операция обладает переместительным и сочетательным свойствами:

1°) $A + B = B + A$;

2°) $(A + B) + C = A + (B + C)$.

Здесь A, B, C — произвольные прямоугольные матрицы одинаковых размеров.

Операция сложения матриц естественным образом распространяется на случай любого числа слагаемых.

2. Умножим в преобразовании (5) величины $y_1, y_2, ..., y_m$ на некоторое число α из K. Тогда

$$\alpha y_i = \sum_{k=1}^{n} (\alpha a_{ik}) x_k \quad (i = 1, 2, ..., m).$$

В соответствии с этим имеет место

Определение 3. *Произведением матрицы* $A = \|a_{ik}\|$ $(i = 1,2,...,m;\ k = 1,2,...,n)$ *на число* α из K называется матрица $C = \|c_{ik}\|$ $(i = 1,2,...,m;\ k = 1,2,...,n)$, элементы которой получаются из соответствующих элементов матрицы A умножением на число α:

$$C = \alpha A,$$

если

$$c_{ik} = \alpha a_{ik} \quad (i = 1,2,...,m; \quad k = 1,2,...,n).$$

Операция нахождения произведения матрицы на число называется *умножением матрицы на число.*

Пример.

$$\alpha \begin{Vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{Vmatrix} = \begin{Vmatrix} \alpha a_1 & \alpha a_2 & \alpha a_3 \\ \alpha b_1 & \alpha b_2 & \alpha b_3 \end{Vmatrix}.$$

Легко видеть, что:

1°) $\alpha(A + B) = \alpha A + \alpha B$;

2°) $(\alpha + \beta)A = \alpha A + \beta A$;

3°) $(\alpha\beta)A = \alpha(\beta A)$.

Здесь A, B — прямоугольные матрицы одинаковых размеров, α, β — числа из поля K.

Разность $A - B$ двух прямоугольных матриц одинаковых размеров определяется равенством

$$A - B = A + (-1)B.$$

Если A — квадратная матрица порядка n, а α — число из K, то [3])

$$|\alpha A| = \alpha^n |A|.$$

3. Пусть величины $z_1, z_2, ..., z_m$ выражаются через величины y_1, $y_2, ..., y_n$ при помощи преобразования

$$z_i = \sum_{k=1}^{n} a_{ik} y_k \quad (i = 1,2,...,m), \tag{8}$$

а величины $y_1, y_2, ..., y_n$ выражаются через величины $x_1, x_2, ..., x_q$ при помощи формул

$$y_k = \sum_{j=1}^{q} b_{kj} x_j \quad (k = 1,2,...,n). \tag{9}$$

Тогда, подставляя эти выражения для y_k $(k = 1,2,...,n)$ в формулы (8), мы выразим $z_1, z_2, ..., z_m$ через $x_1, x_2, ..., x_q$ при помощи "составного" преобразования:

$$z_i = \sum_{k=1}^{n} a_{ik} \sum_{j=1}^{q} b_{kj} x_j = \sum_{j=1}^{q} \Big(\sum_{k=1}^{n} a_{ik} b_{kj} \Big) x_j \quad (i = 1,2,...,m). \tag{10}$$

В соответствии с этим имеет место

Определение 4. *Произведением двух прямоугольных матриц*

$$A = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}, \quad B = \begin{Vmatrix} b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & b_{22} & \dots & b_{2q} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nq} \end{Vmatrix}$$

[3]) Здесь символы $|A|$ и $|\alpha A|$ обозначают определители матриц A и αA (см. с. 11).

называется матрица

$$C = \left\| \begin{array}{cccc} c_{11} & c_{12} & \dots & c_{1q} \\ c_{21} & c_{22} & \dots & c_{2q} \\ \multicolumn{4}{c}{\dotfill} \\ c_{m1} & c_{m2} & \dots & c_{mq} \end{array} \right\|,$$

у которой элемент c_{ij}, стоящий на пересечении i-й строки и j-го столбца, равен “произведению” i-й строки первой матрицы A на j-й столбец второй матрицы B [4]):

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \quad (i = 1, 2, ..., m; \quad j = 1, 2, ..., q). \tag{11}$$

Операция нахождения произведения данных матриц называется *умножением матриц*.

Пример.

$$\left\| \begin{array}{ccc} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{array} \right\| \left\| \begin{array}{cccc} c_1 & d_1 & e_1 & f_1 \\ c_2 & d_2 & e_2 & f_2 \\ c_3 & d_3 & e_3 & f_3 \end{array} \right\| =$$

$$= \left\| \begin{array}{cccc} a_1c_1 + a_2c_2 + & a_1d_1 + a_2d_2 + & a_1e_1 + a_2e_2 + & a_1f_1 + a_2f_2 + \\ + a_3c_3 & + a_3d_3 & + a_3e_3 & + a_3f_3 \\ b_1c_1 + b_2c_2 + & b_1d_1 + b_2d_2 + & b_1e_1 + b_2e_2 + & b_1f_1 + b_2f_2 + \\ + b_3c_3 & + b_3d_3 & + b_3e_3 & + b_3f_3 \end{array} \right\|.$$

По определению 4 матрица коэффициентов в преобразовании (10) равна произведению матрицы коэффициентов в (8) на матрицу коэффициентов в (9).

Заметим, что *операция умножения двух прямоугольных матриц выполнима лишь в том случае, когда число столбцов в первом сомножителе равно числу строк во втором*. В частности, умножение всегда выполнимо, если оба сомножителя — квадратные матрицы одного и того же порядка. Обратим внимание читателя и на то, что даже в этом частном случае умножение матриц не обладает переместительным свойством. Так, например,

$$\left\| \begin{array}{cc} 1 & 2 \\ 3 & 4 \end{array} \right\| \left\| \begin{array}{cc} 2 & 0 \\ 3 & -1 \end{array} \right\| = \left\| \begin{array}{cc} 8 & -2 \\ 18 & -4 \end{array} \right\|, \quad \left\| \begin{array}{cc} 2 & 0 \\ 3 & -1 \end{array} \right\| \left\| \begin{array}{cc} 1 & 2 \\ 3 & 4 \end{array} \right\| = \left\| \begin{array}{cc} 2 & 4 \\ 0 & 2 \end{array} \right\|.$$

Если $AB = BA$, то матрицы A и B называются *перестановочными* или *коммутирующими* между собой.

Пример. Матрицы

$$A = \left\| \begin{array}{cc} 1 & 2 \\ -2 & 0 \end{array} \right\|, \quad B = \left\| \begin{array}{cc} -3 & 2 \\ -2 & -4 \end{array} \right\|$$

перестановочны между собой, так как

$$AB = \left\| \begin{array}{cc} -7 & -6 \\ 6 & -4 \end{array} \right\|, \quad BA = \left\| \begin{array}{cc} -7 & -6 \\ 6 & -4 \end{array} \right\|.$$

Легко проверяется сочетательное свойство умножения матриц, а также распределительное свойство умножения относительно сложения:

---

[4]) Под произведением двух рядов чисел $a_1, a_2, ..., a_n$ и $b_1, b_2, ..., b_n$ мы понимаем сумму произведений соответствующих чисел этих рядов: $\sum\limits_{i=1}^{n} a_i b_i$.

$$
\begin{aligned}
&1^\circ)\ (AB)C = A(BC);\\
&2^\circ)\ (A+B)C = AC + BC; \qquad (12)\\
&3^\circ)\ A(B+C) = AB + AC.
\end{aligned}
$$

Операция умножения матриц естественным образом распространяется на случай нескольких сомножителей.

4. Если воспользоваться произведением прямоугольных матриц, то линейное преобразование

$$
\begin{aligned}
y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n,\\
y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n,
\end{aligned} \tag{13}
$$

можно записать одним матричным равенством

$$
\begin{Vmatrix} y_1 \\ y_2 \\ \ldots \\ y_m \end{Vmatrix} = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix} \begin{Vmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{Vmatrix},
$$

или в сокращенной записи:

$$
y = Ax. \tag{13$'$}
$$

Здесь $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_n)$ — столбцевые матрицы, $A = \|a_{ik}\|$ — прямоугольная матрица размера $m \times n$.

Равенства (13) выражают собой тот факт, что столбец y является линейной комбинацией столбцов матрицы A с коэффициентами $x_1, x_2, ..., x_n$:

$$
y = x_1 a_{\cdot 1} + x_2 a_{\cdot 2} + ... + x_n a_{\cdot n} = \sum_{k=1}^{n} x_k a_{\cdot k}. \tag{13$''$}
$$

Вернемся теперь к равенствам (11), которые эквивалентны одному матричному равенству

$$
C = AB. \tag{14}
$$

Эти равенства могут быть записаны в виде

$$
c_{\cdot j} = \sum_{k=1}^{n} b_{kj} a_{\cdot k} \quad (j = 1, 2, ..., q) \tag{14$'$}
$$

или в виде

$$
c_{i\,\cdot} = \sum_{k=1}^{n} a_{ik} b_{k\,\cdot} \quad (i = 1, 2, ..., m). \tag{14$''$}
$$

Таким образом, любой j-й столбец матрицы-произведения $C = AB$ является линейной комбинацией столбцов первого сомножителя, т. е. матрицы A, причем коэффициенты этой линейной зависимости образуют j-й столбец во втором сомножителе B. Аналогично, любая i-я строка в матрице C является линейной комбинацией строк матрицы B, а коэффициентами этой линейной зависимости являются элементы i-й строки матрицы A[5]).

[5]) Следовательно, матричные уравнения $AX = C$, где A и C — заданные соответственно $m \times n$-матрица и $m \times q$-матрица, а X — искомая $n \times q$-матрица, имеет решение в том и только том случае, когда столбцы матрицы C являются линейными комбинациями столбцов матрицы A. Для уравнения $XB = C$ необходимое и достаточное условие существования решения X состоит в том, что строки матрицы C должны быть линейными комбинациями строк матрицы B.

Остановимся еще на том частном случае, когда в произведении $C = AB$ второй сомножитель является квадратной и притом диагональной матрицей $B = \{d_1, d_2, ..., d_n\}$. Тогда из формул (11) следует

$$c_{ij} = a_{ij}d_j \quad (i = 1, 2, ..., m; \quad j = 1, 2, ..., n),$$

т. е.

$$\left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{matrix} \right\| \left\| \begin{matrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{matrix} \right\| = \left\| \begin{matrix} a_{11}d_1 & a_{12}d_2 & \dots & a_{1n}d_n \\ a_{21}d_1 & a_{22}d_2 & \dots & a_{2n}d_n \\ \dots & \dots & \dots & \dots \\ a_{m1}d_1 & a_{m2}d_2 & \dots & a_{mn}d_n \end{matrix} \right\|.$$

Аналогично

$$\left\| \begin{matrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_m \end{matrix} \right\| \left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{matrix} \right\| = \left\| \begin{matrix} d_1a_{11} & d_1a_{12} & \dots & d_1a_{1n} \\ d_2a_{21} & d_2a_{22} & \dots & d_2a_{2n} \\ \dots & \dots & \dots & \dots \\ d_ma_{m1} & d_ma_{m2} & \dots & d_ma_{mn} \end{matrix} \right\|.$$

Таким образом, при *умножении прямоугольной матрицы A справа* (*слева*) *на диагональную матрицу* $\{d_1, d_2, ...\}$ *все столбцы* (*соответственно строки*) *матрицы A умножаются на числа* $d_1, d_2, ...$

5. Пусть *квадратная* матрица $C = \|c_{ij}\|_1^m$ является произведением двух прямоугольных матриц $A = \|a_{ik}\|$ и $B = \|b_{kj}\|$ соответственно размеров $m \times n$ и $n \times m$:

$$\left\| \begin{matrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{matrix} \right\| = \left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{matrix} \right\| \left\| \begin{matrix} b_{11} & \dots & b_{1m} \\ b_{21} & \dots & b_{2m} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nm} \end{matrix} \right\|, \tag{15}$$

т. е.

$$c_{ij} = \sum_{\alpha=1}^{n} a_{i\alpha}b_{\alpha j} \quad (i, j = 1, 2, ..., m). \tag{15$'$}$$

Установим важную ф о р м у л у Б и н е – К о ш и, выражающую определитель $|C|$ через миноры матриц A и B:

$$\left| \begin{matrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{matrix} \right| = \sum_{1 \leqslant k_1 < k_2 < \ldots < k_m \leqslant n} \left| \begin{matrix} a_{1k_1} & \dots & a_{1k_m} \\ \dots & \dots & \dots \\ a_{mk_1} & \dots & a_{mk_m} \end{matrix} \right| \left| \begin{matrix} b_{k_1 1} & \dots & b_{k_m 1} \\ \dots & \dots & \dots \\ b_{k_1 m} & \dots & b_{k_m m} \end{matrix} \right|, \tag{16}$$

или в специальных обозначениях (см. с. 11):

$$C \begin{pmatrix} 1 & 2 & \dots & m \\ 1 & 2 & \dots & m \end{pmatrix} = \sum_{1 \leqslant k_1 < k_2 < \ldots < k_n \leqslant n} A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}. \tag{16$'$}$$

Согласно этой формуле *определитель матрицы C равен сумме произведений всевозможных миноров максимального (m-го) порядка*[6]) *матрицы A на соответствующие миноры того же порядка матрицы B.*

Вывод формулы Бине–Коши. На основании формулы (15′) определитель матрицы C можно представить в виде

$$\begin{vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} = \begin{vmatrix} \sum\limits_{\alpha_1=1}^{n} a_{1\alpha_1} b_{\alpha_1 1} & \dots & \sum\limits_{\alpha_m=1}^{n} a_{1\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots \\ \sum\limits_{\alpha_1=1}^{n} a_{m\alpha_1} b_{\alpha_1 1} & \dots & \sum\limits_{\alpha_m=1}^{n} a_{m\alpha_m} b_{\alpha_m m} \end{vmatrix} =$$

$$= \sum_{\alpha_1,\dots,\alpha_m=1}^{n} \begin{vmatrix} a_{1\alpha_1} b_{\alpha_1 1} & \dots & a_{1\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots \\ a_{m\alpha_1} b_{\alpha_1 1} & \dots & a_{m\alpha_m} b_{\alpha_m m} \end{vmatrix} =$$

$$= \sum_{\alpha_1,\dots,\alpha_m=1}^{n} A \begin{pmatrix} 1 & 2 & \dots & m \\ \alpha_1 & \alpha_2 & \dots & \alpha_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m}. \quad (16'')$$

Если $m > n$, то среди чисел $\alpha_1, \alpha_2, \dots, \alpha_m$ всегда найдутся равные между собой числа и, следовательно, каждое слагаемое в правой части равенства (16″) будет равно нулю. Значит, в этом случае $|C| = 0$.

Пусть теперь $m \leqslant n$. Тогда в сумме, стоящей в правой части равенства (16″), будут равны нулю те слагаемые, у которых хотя бы два из индексов $\alpha_1, \alpha_2, \dots, \alpha_m$ равны между собой. Все же остальные слагаемые этой суммы можно разбить на группы по $m!$ слагаемых в каждой, объединяя в одну группу те слагаемые, которые отличаются друг от друга только порядком индексов $\alpha_1, \alpha_2, \dots, \alpha_m$ (индексы $\alpha_1, \alpha_2, \dots, \alpha_m$ в пределах каждой группы слагаемых имеют одну и ту же совокупность значений). Тогда в пределах одной такой группы сумма соответствующих слагаемых будет равна[7])

$$\sum \varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m} =$$

$$= A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} \sum \varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m)\, b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m} =$$

$$= A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}.$$

Поэтому из (16″) получаем (16′).

Пример 1.

$$\left\| \begin{matrix} a_1c_1 + a_2c_2 + \dots & a_1d_1 + a_2d_2 + \dots \\ \dots + a_nc_n & \dots + a_nd_n \\ b_1c_1 + b_2c_2 + \dots & b_1d_1 + b_2d_2 + \dots \\ \dots + b_nc_n & \dots + b_nd_n \end{matrix} \right\| = \left\| \begin{matrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \end{matrix} \right\| \left\| \begin{matrix} c_1 & d_1 \\ c_d & d_2 \\ \vdots & \vdots \\ c_n & d_n \end{matrix} \right\|.$$

---

[6]) Если $m > n$, то матрицы A и B не имеют миноров m-го порядка. В этом случае правые части формул (16) и (16′) следует заменить нулями.

[7]) Здесь $k_1 < k_2 < \dots < k_m$ — нормальное расположение индексов $\alpha_1, \alpha_2, \dots, \alpha_m$, а $\varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) = (-1)^N$, где N — число транспозиций индексов, необходимых для преобразования перестановки $\alpha_1, \alpha_2, \dots, \alpha_m$ к нормальному расположению $k_1 < k_2 < \dots < k_m$.

Поэтому формула (16) дает так называемое тождество Коши

$$\left\| \begin{matrix} a_1c_1 + a_2c_2 + \ldots & a_1d_1 + a_2d_2 + \ldots \\ \ldots + a_nc_n & \ldots + a_nd_n \\ b_1c_1 + b_2c_2 + \ldots & b_1d_1 + b_2d_2 + \ldots \\ \ldots + b_nc_n & + b_nd_n \end{matrix} \right\| = \sum_{1 \leqslant i < k \leqslant n} \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix} \begin{vmatrix} c_i & d_i \\ c_k & d_k \end{vmatrix}.$$

Полагая в этом тождестве $a_i = c_i$, $b_i = d_i$ $(i = 1, 2, \ldots, n)$, получим

$$\begin{vmatrix} a_1^2 + a_2^2 + \ldots + a_n^2 & a_1b_1 + a_2b_2 + \ldots + a_nb_n \\ a_1b_1 + a_2b_2 + \ldots + a_nb_n & b_1^2 + b_2^2 + \ldots + b_n^2 \end{vmatrix} = \sum_{1 \leqslant i < k \leqslant n} \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix}^2.$$

В случае, когда a_i и b_i $(i = 1, 2, \ldots, n)$ — вещественные числа, отсюда следует известное неравенство

$$(a_1b_1 + a_2b_2 + \ldots + a_nb_n)^2 \leqslant (a_1^2 + a_2^2 + \ldots + a_n^2)(b_1^2 + b_2^2 + \ldots + b_n^2).$$

При этом знак равенства имеет место тогда и только тогда, когда все числа a_i пропорциональны соответствующим числам b_i $(i = 1, 2, \ldots, n)$.

Пример 2.

$$\left\| \begin{matrix} a_1c_1 + b_1d_1 & \ldots & a_1c_n + b_1d_n \\ \ldots & \ldots & \ldots \\ a_nc_1 + b_nd_1 & \ldots & a_nc_n + b_nd_n \end{matrix} \right\| = \left\| \begin{matrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{matrix} \right\| \left\| \begin{matrix} c_1 & \ldots & c_n \\ d_1 & \ldots & d_n \end{matrix} \right\|.$$

Поэтому [8]) при $n > 2$

$$\left\| \begin{matrix} a_1c_1 + b_1d_1 & \ldots & a_1c_n + b_1d_n \\ \ldots & \ldots & \ldots \\ a_nc_1 + b_nd_1 & \ldots & a_nc_n + b_nd_n \end{matrix} \right\| = 0.$$

Рассмотрим частный случай, когда A и B — квадратные матрицы одного и того же порядка n, и положим в $(16')$ $m = n$. Тогда приходим к известной теореме об умножении определителей

$$C \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} = A \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} B \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix},$$

или, в других обозначениях,

$$|C| = |AB| = |A| \cdot |B|. \tag{17}$$

Таким образом, *определитель произведения двух квадратных матриц равен произведению определителей перемножаемых матриц.*

6. Формула Бине–Коши дает возможность в самом общем случае выразить миноры произведения двух прямоугольных матриц через миноры сомножителей. Пусть

$$A = \|a_{ik}\|, \quad B = \|b_{kj}\|, \quad C = \|c_{ij}\|$$

$$(i = 1, 2, \ldots, m; \quad k = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, q)$$

и

$$C = AB.$$

---

[8]) См. сноску на с. 18.

Рассмотрим произвольный минор матрицы C:

$$C\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix}$$

$$(1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant m, \quad 1 \leqslant j_1 < j_2 < \ldots < j_p \leqslant q; \quad p \leqslant m, q).$$

Матрица, составленная из элементов этого минора, представляет собой произведение двух прямоугольных матриц

$$\left\| \begin{matrix} a_{i_1 1} & a_{i_1 2} & \dots & a_{i_1 n} \\ \dots & \dots & \dots & \dots \\ a_{i_p 1} & a_{i_p 2} & \dots & a_{i_p n} \end{matrix} \right\|, \quad \left\| \begin{matrix} b_{1 j_1} & \dots & b_{1 j_p} \\ b_{2 j_1} & \dots & b_{2 j_p} \\ \dots & \dots & \dots \\ b_{n j_1} & \dots & b_{n j_p} \end{matrix} \right\|.$$

Поэтому, применяя формулу Бине–Коши, получаем [9]

$$C\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} = \\ = \sum_{1 \leqslant k_1 < k_2 < \ldots < k_p \leqslant n} A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \dots & k_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix}. \quad (18)$$

При $p = 1$ формула (18) переходит в формулу (11). При $p > 1$ формула (18) является естественным обобщением формулы (11).

Отметим еще одно следствие из формулы (18).

Ранг произведения двух прямоугольных матриц не превосходит ранга любого из сомножителей.

Если $C = AB$ и r_A, r_B, r_C — ранги матриц A, B, C, то

$$r_C \leqslant r_A, r_B.$$

7. Если X — решение матричного уравнения $AX = C$ (размеры матриц A, X и C соответственно $m \times n$, $n \times q$ и $m \times q$), то $r_X \leqslant r_C$. Покажем, что *среди решений матричного уравнения $AX = C$ существует решение X_0 минимального ранга, для которого $r_{X_0} = r_C$.*

Действительно, пусть $r = r_C$. Тогда среди столбцов матрицы C имеется r линейно независимых [10]. Пусть для конкретности первые r столбцов $C_{\cdot 1}, \ldots, C_{\cdot r}$ линейно независимы, а остальные столбцы $C_{\cdot r+1}, \ldots, C_{\cdot q}$ являются линейными комбинациями первых r:

$$C_{\cdot j} = \sum_{k=1}^{r} \alpha_{jk} C_{\cdot k} \quad (j = r+1, \ldots, q). \quad (19)$$

Пусть X — произвольное решение уравнения $AX = C$. Тогда (см. с. 16)

$$AX_{\cdot k} = C_{\cdot k} \quad (k = 1, \ldots, r). \quad (20)$$

Определим столбцы $\widetilde{X}_{\cdot r+1}, \ldots, \widetilde{X}_{\cdot q}$ равенствами

$$\widetilde{X}_{\cdot j} = \sum_{k=1}^{r} \alpha_{jk} X_{\cdot k} \quad (j = r+1, \ldots, q).$$

---

[9]) Из той же формулы Бине–Коши следует, что миноры p-го порядка матрицы C при $p > n$ (если миноры таких порядков имеются) все равны нулю. В этом случае правую часть формулы (18) следует заменить нулем. См. сноску на с. 18.

[10]) Мы ссылаемся на хорошо известное положение: ранг матрицы равен числу линейно независимых столбцов (строк) матрицы. Доказательство этого положения приведено в гл. III, с. 66.

Умножая эти равенства слева почленно на A, в силу равенств (19) и (20) находим

$$A\widetilde{X}_{.j} = C_{.j} \quad (j = r+1, ..., q). \tag{20'}$$

Система из q равенств (20) и (20′) эквивалентна одному матричному равенству

$$AX_0 = C,$$

где $X_0 = (X_{.1}, ..., x_{.r}, \widetilde{X}_{.r+1}, ..., \widetilde{X}_{.q})$ — матрица ранга r [11]).

Решение X_0 минимального ранга r_C матричного уравнения $AX = C$ всегда представимо в виде

$$X_0 = VC,$$

где V — некоторая $n \times m$-матрица.

Действительно, из равенства $AX_0 = C$ следует, что строки матрицы C являются линейными комбинациями строк матрицы X_0. Поскольку как среди строк матрицы C, так и среди строк матрицы X_0 имеется одно и то же число r_C линейно независимых [12]), то и, обратно, строки матрицы X_0 являются линейными комбинациями строк матрицы C, а отсюда уже следует равенство $X_0 = VC$.

Докажем теперь следующее предложение [13]).

Матричное уравнение

$$AXB = C, \tag{21}$$

где A, B — заданные матрицы, а X — искомая прямоугольная матрица [14]), *имеет решение в том и только том случае, когда одновременно имеют решения матричные уравнения*

$$AY = C, \quad ZB = C, \tag{22}$$

т. е. когда столбцы матрицы C являются линейными комбинациями столбцов матрицы A, а строки матрицы C являются линейными комбинациями строк матрицы B.

В самом деле, если матрица X — решение уравнения (21), то матрицы $Y = XB$ и $Z = AX$ являются решениями уравнений (22).

Обратно, пусть существуют решения Y, Z уравнений (22). Тогда первое из этих уравнений имеет решение Y_0 минимального ранга r_C, которое по доказанному представимо в виде

$$Y_0 = VC.$$

Поэтому

$$C = AY_0 = AVC = AVZB.$$

Тогда матрица $X = VZ$ будет решением уравнения (21).

---

[11]) В матрице X_0 последние $n - r$ столбцов являются линейными комбинациями первых r столбцов; первые же r столбцов $X_{.1}, ..., X_{.r}$ линейно независимы, так как линейная зависимость между этими столбцами в силу равенств (20) повлекла бы линейную зависимость между столбцами $C_{.1}, ..., C_{.r}$.

[12]) См. сноску на с. 16.

[13]) См. [231, 199].

[14]) Предполагается, что размеры матриц A, X, B, C таковы, что произведение AXB имеет смысл и имеет размер матрицы C.

§ 3. Квадратные матрицы

1. Квадратную матрицу n-го порядка, у которой на главной диагонали стоят единицы, а все остальные элементы равны нулю, будем называть *единичной матрицей* и обозначать через $E^{(n)}$ или просто E. Название "единичная матрица" связано со следующим свойством матрицы E: для любой прямоугольной матрицы

$$A = \|a_{ik}\| \quad (i = 1, 2, ..., m; \quad k = 1, 2, ..., n)$$

имеют место равенства

$$E^{(m)}A = AE^{(n)} = A.$$

Очевидно,

$$E^{(n)} = \|\delta_{ik}\|_1^n.$$

Пусть $A = \|a_{ik}\|_1^n$ — квадратная матрица. Тогда *степень матрицы* определяется обычным образом:

$$A^p = \underbrace{A\ A\ ...\ A}_{p \text{ раз}} \quad (p = 1, 2, ...), \qquad A^0 = E.$$

Из сочетательного свойства умножения матриц следует

$$A^p A^q = A^{p+q}.$$

Здесь p, q — произвольные целые неотрицательные числа.

Рассмотрим многочлен (целую рациональную функцию) с коэффициентами из поля K

$$f(t) = \alpha_0 t^m + \alpha_1 t^{m-1} + ... + \alpha_m.$$

Тогда под $f(A)$ будем понимать матрицу

$$f(A) = \alpha A^m + \alpha_1 A^{m-1} + ... + \alpha_m E.$$

Так определяется *многочлен от матрицы.*

Пусть многочлен $f(t)$ равен произведению многочленов $h(t)$ и $g(t)$:

$$f(t) = h(t)g(t).$$

Многочлен $f(t)$ получается из $h(t)$ и $g(t)$ путем почленного перемножения и приведения подобных членов. При этом используется правило перемножения степеней: $t^p t^q = t^{p+q}$. Так как все эти действия правомерны и при замене скалярной величины t на матрицу A, то

$$f(A) = h(A)g(A).$$

Отсюда, в частности [15]),

$$h(A)g(A) = g(A)h(A),$$

т. е. *два многочлена от одной и той же матрицы всегда перестановочны между собой.*

П р и м е р ы. Условимся *p-й наддиагональю* (*поддиагональю*) в прямоугольной матрице $A = \|a_{ik}\|$ называть ряд элементов a_{ik}, у которых $k - i = p$ (соответственно $i - k = p$). Обозначим через H квадратную матрицу n-го порядка, у кото-

[15]) Так как каждое из этих произведений равно одному и тому же $f(A)$, поскольку и $g(t)h(t) = f(t)$. Следует отметить, что подстановка матриц в алгебраическое тождество с н е с к о л ь к и м и переменными неправомерна. Впрочем, перестановочные между собой матрицы можно подставлять и в этом случае.

рой элементы первой наддиагонали равны единице, а все остальные элементы равны нулю. Тогда

$$H=\begin{Vmatrix} 0 & 1 & 0 & & 0\\ 0 & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ 0 & 0 & 0 & \dots & 0\end{Vmatrix},\quad H^2=\begin{Vmatrix} 0 & 0 & 1 & & 0\\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1\\ & & & \ddots & 0\\ 0 & 0 & 0 & & 0\end{Vmatrix},\quad \ldots,$$

$$H^p=0\quad (p\geqslant n).$$

В силу этих равенств, если

$$f(t)=a_0+a_1t+a_2t^2+\ldots+a_{n-1}t^{n-1}+\ldots$$

— многочлен относительно t, то

$$f(H)=a_0E+a_1H+a_2H^2+\ldots=\begin{Vmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1}\\ 0 & a_0 & a_1 & \ddots & \vdots\\ & & \ddots & \ddots & a_2\\ & & & \ddots & a_1\\ 0 & 0 & 0 & & a_0\end{Vmatrix}.$$

Аналогично, если F — квадратная матрица n-го порядка, у которой все элементы первой поддиагонали равны единице, а все остальные элементы равны нулю, то

$$f(F)=a_0E+a_1F+a_2F^2+\ldots=\begin{Vmatrix} a_0 & 0 & \dots & 0\\ a_1 & a_0 & \ddots & \\ & \ddots & \ddots & 0\\ a_{n-1} & & a_1 & a_0\end{Vmatrix}.$$

Предлагаем читателю проверить следующие свойства матриц H и F.

1°. *В результате умножения произвольной $m\times n$-матрицы A слева на матрицу H (матрицу F) m-го порядка все строки матрицы A поднимаются (опускаются) на одно место вверх (вниз), первая (последняя) строка матрицы A исчезает, а последняя (первая) строка произведения заполняется нулями.* Так, например,

$$\begin{Vmatrix} 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{Vmatrix}\begin{Vmatrix} a_1 & a_2 & a_3 & a_4\\ b_1 & b_2 & b_3 & b_4\\ c_1 & c_3 & c_3 & c_4\end{Vmatrix}=\begin{Vmatrix} b_1 & b_2 & b_3 & b_4\\ c_1 & c_2 & c_3 & c_4\\ 0 & 0 & 0 & 0\end{Vmatrix},$$

$$\begin{Vmatrix} 0 & 0 & 0\\ 1 & 0 & 0\\ 0 & 1 & 0\end{Vmatrix}\begin{Vmatrix} a_1 & a_2 & a_3 & a_4\\ b_1 & b_2 & b_3 & b_4\\ c_1 & c_2 & c_3 & c_4\end{Vmatrix}=\begin{Vmatrix} 0 & 0 & 0 & 0\\ a_1 & a_2 & a_3 & a_4\\ b_1 & b_2 & b_3 & b_4\end{Vmatrix}.$$

2°. *В результате умножения произвольной $m\times n$-матрицы A справа на матрицу $H(F)$ n-го порядка все столбцы матрицы A сдвигаются вправо (влево) на одно место, при этом последний (первый) столбец матрицы A исчезает, а первый (последний) столбец произведения заполняется нулями.* Так, например,

$$\begin{Vmatrix} a_1 & a_2 & a_3 & a_4\\ b_1 & b_2 & b_3 & b_4\\ c_1 & c_2 & c_3 & c_4\end{Vmatrix}\begin{Vmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{Vmatrix}=\begin{Vmatrix} 0 & a_1 & a_2 & a_3\\ 0 & b_1 & b_2 & b_3\\ 0 & c_1 & c_2 & c_3\end{Vmatrix},$$

$$\begin{Vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{Vmatrix} \begin{Vmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{Vmatrix} = \begin{Vmatrix} a_2 & a_3 & a_4 & 0 \\ b_2 & b_3 & b_4 & 0 \\ c_2 & c_3 & c_4 & 0 \end{Vmatrix}.$$

2. Квадратную матрицу будем называть *вырожденной*, если $|A| = 0$. В противном случае квадратная матрица A называется *невырожденной*.

Пусть $A = \|a_{ik}\|_1^n$ — невырожденная матрица ($|A| \neq 0$). Рассмотрим линейное преобразование с матрицей коэффициентов A

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \quad (i = 1, 2, ..., n). \tag{23}$$

Рассматривая равенства (23) как уравнения относительно $x_1,\ x_2, ..., x_n$ и замечая, что определитель системы уравнений (23) по условию отличен от нуля, мы можем однозначно по известным формулам выразить $x_1, x_2, ..., x_n$ через $y_1, y_2, ..., y_n$:

$$x_i = \frac{1}{|A|} \begin{Vmatrix} a_{11} & \dots & a_{1,i-1} & y_1 & a_{1,i+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2,i-1} & y_2 & a_{2,i+1} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,i-1} & y_n & a_{n,i+1} & \dots & a_{nn} \end{Vmatrix} \equiv \sum_{k=1}^{n} a_{ik}^{(-1)} y_k \quad (i = 1, 2, ..., n). \tag{24}$$

Мы получили "обратное" преобразование для (23). Матрицу коэффициентов этого преобразования

$$A^{-1} = \|a_{ik}^{(-1)}\|_1^n$$

мы назовем *обратной матрицей* для матрицы A. Из (24) легко усмотреть, что

$$a_{ik}^{(-1)} = \frac{A_{ki}}{|A|} \quad (i, k = 1, 2, ..., n), \tag{25}$$

где A_{ki} — алгебраическое дополнение элемента a_{ki} в определителе $|A|$ $(i, k = 1, 2, ..., n)$.

Так, например, если

$$A = \begin{Vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{Vmatrix}, \quad |A| \neq 0,$$

то

$$A^{-1} = \frac{1}{|A|} \begin{Vmatrix} b_2c_3 - b_3c_2 & a_3c_2 - a_2c_3 & a_2b_3 - a_3b_2 \\ b_3c_1 - b_1c_3 & a_1c_3 - a_3c_1 & a_3b_1 - a_1b_3 \\ b_1c_2 - b_2c_1 & a_2c_1 - a_1c_2 & a_1b_2 - a_2b_1 \end{Vmatrix}.$$

Образуя составное преобразование из данного преобразования (23) и обратного (24) в одном и в другом порядке, мы в обоих случаях получаем тождественное преобразование (с единичной матрицей коэффициентов); поэтому

$$AA^{-1} = A^{-1}A = E. \tag{26}$$

В справедливости равенств (26) можно убедиться и непосредственным перемножением матриц A и A^{-1}. Действительно, в силу (25) [16])

$$[AA^{-1}]_{ij} = \sum_{k=1}^{n} a_{ik} a_{kj}^{(-1)} = \frac{1}{|A|} \sum_{k=1}^{n} a_{ik} A_{jk} = \delta_{ij} \quad (i, j = 1, 2, ..., n).$$

[16]) Здесь мы используем известное свойство определителя, согласно которому сумма произведений элементов какого-либо столбца на их алгебраические дополнения равна величине определителя, а сумма произведений элементов столбца на алгебраические дополнения соответствующих элементов другого столбца равна нулю.

Аналогично

$$[A^{-1}A]_{ij} = \sum_{k=1}^{n} a_{ij}^{(-1)} a_{kj} = \frac{1}{|A|} \sum_{k=1}^{n} A_{ki} a_{kj} = \delta_{ij} \quad (i, j = 1, 2, ..., n).$$

Нетрудно видеть, что матричные уравнения

$$AX = E, \quad XA = E \quad (|A| \neq 0) \tag{27}$$

никаких других решений, кроме решения $X = A^{-1}$, не имеют. Действительно, умножая обе части первого уравнения слева, а второго справа на A^{-1} и используя сочетательное свойство произведения матриц, а также равенства (26), мы в обоих случаях получим [17])

$$X = A^{-1}.$$

Этим же способом доказывается, что каждое из матричных уравнений

$$AX = B, \quad XA = B \quad (|A| \neq 0), \tag{28}$$

где X и B — прямоугольные матрицы равных размеров, A — квадратная матрица соответствующего размера, имеет одно и только одно решение:

$$X = A^{-1}B \quad \text{и соответственно} \quad X = BA^{-1}. \tag{29}$$

Матрицы (29) являются как бы "левым" и "правым" частными от "деления" матрицы B на матрицу A. Из (28) и (29) следует соответственно (см. с. 21) $r_B \leqslant r_X$ и $r_X \leqslant r_B$, т. е. $r_X = r_B$. Сопоставляя с (28), имеем:

при умножении прямоугольной матрицы слева или справа на невырожденную матрицу ранг исходной матрицы не изменяется.

Заметим еще, что из (26) вытекает $|A||A^{-1}| = 1$, т. е.

$$|A^{-1}| = \frac{1}{|A|}.$$

Для произведения двух неособенных матриц имеем

$$(AB)^{-1} = B^{-1}A^{-1}. \tag{30}$$

3. Все матрицы n-го порядка образуют кольцо [18]) с единичным элементом E. Поскольку в этом кольце определена операция умножения на число из поля K и существует базис из n^2 линейно независимых матриц, через которые линейно выражаются все матрицы n-го порядка [19]), то кольцо матриц n-го порядка является *алгеброй* [20]).

[17]) Если A — вырожденная матрица, то уравнения (27) не имеют решений. Действительно, если бы какое-либо из этих уравнений имело решение $X = \|x_{ik}\|_1^n$, то тогда было бы в силу теоремы об умножении определителей [см. формулу (17)] $|A||X| = |E| = 1$, что невозможно при $|A| = 0$.

[18]) *Кольцом* называется совокупность элементов, в которой определены и всегда однозначно выполнимы две операции: "сложение" двух элементов (с переместительным и сочетательным свойствами) и "умножение" двух элементов (с сочетательным и распределительным относительно сложения свойствами), причем сложение обратимо. См., например, [20, с. 270, 271] или [39, с. 333].

[19]) Действительно, произвольная матрица $A = \|a_{ik}\|_1^n$ с элементами из K представима в виде $A = \sum_{i,k=1}^{n} a_{ik} E_{ik}$, где E_{ik} — матрица n-го порядка, у которой на пересечении i-й строки и k-го столбца стоит 1, а все остальные элементы равны 0.

[20]) См., например, [20, с. 101].

Все квадратные матрицы n-го порядка образуют коммутативную группу относительно операции сложения [21]). Все невырожденные матрицы n-го порядка образуют (некоммутативную) группу относительно операции умножения.

Квадратная матрица $A = \|a_{ik}\|_1^n$ называется *верхней треугольной* (*нижней треугольной*), если равны нулю все элементы матрицы, расположенные под главной диагональю (над главной диагональю):

$$A = \begin{Vmatrix} a_{11} & a_{12} & & a_{1n} \\ 0 & a_{22} & & a_{2n} \\ & & \ddots & \\ 0 & 0 & & a_{nn} \end{Vmatrix}, \qquad A = \begin{Vmatrix} a_{11} & 0 & & 0 \\ a_{21} & a_{22} & & 0 \\ & & \ddots & \\ a_{n1} & a_{n2} & & a_{nn} \end{Vmatrix}.$$
(1) (2)

Диагональная матрица является частным случаем как верхней, так и нижней треугольной матрицы.

Так как определитель треугольной матрицы равен произведению ее диагональных элементов, то треугольная (и, в частности, диагональная) матрица является невырожденной только тогда, когда все ее диагональные элементы отличны от нуля.

Легко проверить, что сумма и произведение двух диагональных (верхних треугольных, нижних треугольных) матриц есть диагональная (соответственно верхняя треугольная, нижняя треугольная) матрица и что обратная матрица для невырожденной диагональной (верхней треугольной, нижней треугольной) матрицы является матрицей того же типа. Поэтому:

1°) все диагональные, все верхние треугольные, все нижние треугольные матрицы n-го порядка образуют три коммутативные группы относительно операции сложения;

2°) все невырожденные диагональные матрицы образуют коммутативную группу относительно умножения;

3°) все невырожденные верхние (нижние) треугольные матрицы составляют группу (некоммутативную) относительно умножения.

4. В заключение этого параграфа укажем на две важные операции над матрицами — *транспонирование* матрицы и переход к *сопряженной* матрице.

Если $A = \|a_{ik}\|$ $(i = 1, 2, ..., m;\ k = 1, 2, ..., n)$, то транспонированная матрица A' определяется равенством $A' = \|a'_{ik}\|$, где $a'_{ki} = a_{ik}$ $(i = 1, 2, ..., m;\ k = 1, 2, ..., n)$. Сопряженная же матрица $A^* = \|a^*_{ik}\|$, где $a^*_{ki} = \overline{a}'_{ki} = \overline{a}_{ik}$ $(i = 1, 2, ..., m;$ $k = 1, 2, ..., n)$ [22]). Если матрица A имеет размер $m \times n$, то матрицы A' и A^* имеют размер $n \times m$.

Легко проверяются следующие свойства [23]):

---

[21]) *Группой* называется всякая совокупность объектов, в которой установлена операция, относящая любым двум элементам a и b совокупности определенный третий элемент $a * b$ той же совокупности, если: 1) операция обладает сочетательным свойством $[(a * b) * c = a * (b * c)]$; 2) существует в совокупности единичный элемент e $(a * e = e * a = a)$; 3) для любого элемента a совокупности существует обратный элемент a^{-1} $(a * a^{-1} = a^{-1} * a = e)$. Группа называется *коммутативной* или *абелевой*, если групповая операция обладает переместительным свойством. Относительно понятия группы см., например, [20, с. 392–398].

[22]) Чертой обозначается переход к комплексно сопряженной величине.

[23]) В формулах 1°), 2°), 3°) и 5°) на с. 27 A, B — произвольные прямоугольные матрицы, для которых соответствующие операции выполнимы, а α — произвольное комплексное число. В формуле 4°) A — произвольная квадратная невырожденная матрица.

1°) $(A+B)' = A' + B', \qquad (A+B)^* = A^* + B^*;$

2°) $(\alpha A)' = \alpha A', \qquad (\alpha A)^* = \overline{\alpha} A^*;$

3°) $(AB)' = B'A', \qquad (AB)^* = B^*A^*;$

4°) $(A^{-1})' = (A')^{-1}, \qquad (A^{-1})^* = (A^*)^{-1};$

5°) $(A')' = A, \qquad (A^*)^* = A.$

Если квадратная матрица $S = \|s_{ik}\|_1^n$ совпадает со своей транспонированной $(S' = S)$, то такая матрица называется *симметрической*. Если же квадратная матрица $H = \|h_{ik}\|$ совпадает со своей сопряженной $(H^* = H)$, то она называется *эрмитовой*. В симметрической матрице элементы, симметрично расположенные относительно главной диагонали, равны, а в эрмитовой они комплексно сопряжены между собой. Диагональные элементы эрмитовой матрицы всегда вещественны. Заметим, что произведение двух симметрических (эрмитовых) матриц, вообще говоря, не является симметрической (эрмитовой) матрицей. В силу 3°) это имеет место только в том случае, когда данные две симметрические или эрмитовы матрицы перестановочны между собой.

Если A — *вещественная* матрица, т. е. матрица с вещественными элементами, то $A^* = A'$. Эрмитова вещественная матрица всегда является симметрической.

С каждой прямоугольной матрицей $A = \|a_{ik}\|$ размера $m \times n$ связаны две эрмитовы матрицы, AA^* и A^*A, размеров $m \times m$ и $n \times n$. Любое из равенств $AA^* = 0$ или $A^*A = 0$ влечет за собой [24]) равенство $A = 0$.

Если квадратная матрица $K = \|k_{ij}\|_1^n$ отличается множителем -1 от своей транспонированной $(K' = -K)$, то такая матрица называется *кососимметрической*. В кососимметрической матрице любые два элемента, расположенные симметрично относительно главной диагонали, отличаются друг от друга множителем -1, а диагональные элементы равны нулю. Из 3°) следует, что произведение двух перестановочных между собой кососимметрических матриц является симметрической матрицей [25]).

§ 4. Ассоциированные матрицы. Миноры обратной матрицы

Пусть дана матрица $A = \|a_{ik}\|_1^n$. Рассмотрим всевозможные миноры p-го $(1 \leqslant p \leqslant n)$ порядка матрицы A:

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \qquad (31)$$
$$(1 \leqslant i_1 < i_2 < \dots < i_p \leqslant n, \quad 1 \leqslant k_1 < k_2 < \dots < k_p \leqslant n).$$

Число этих миноров равно N^2, где $N = C_n^p$ — число сочетаний из n по p. Для того чтобы расположить миноры (31) в квадратную таблицу, занумеруем в определенном (например, лексикографическом) порядке все сочетания по p из n индексов $1, 2, \dots, n$.

Если сочетания индексов $i_1 < i_2 < \dots < i_p$ и $k_1 < k_2 < \dots < k_p$ при этой нумера-

---

[24]) Это следует из того, что сумма диагональных элементов в каждой из матриц AA^* и A^*A равна $\sum_{i=1}^{m}\sum_{k=1}^{n} |a_{ik}|^2$.

[25]) Относительно представления квадратной матрицы A в виде произведения двух симметрических $(A = S_1S_2)$ либо двух кососимметрических матриц $(A = K_1K_2)$ см. [246].

ции будут иметь номера α и β, то минор (31) будем обозначать и так:

$$\mathfrak{a}_{\alpha\beta} = A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}.$$

Давая α и β независимо друг от друга все значения от 1 до N, мы охватим все миноры p-го порядка матрицы $A = \|a_{ik}\|_1^n$.

Квадратная матрица N-го порядка

$$\mathfrak{A}_p = \|\mathfrak{a}_{\alpha\beta}\|_1^N$$

называется *p-й ассоциированной матрицей* для матрицы $A = \|a_{ik}\|_1^n$; p может принимать значения $1, 2, ..., n$. При этом $\mathfrak{A}_1 = A$, а матрица $\mathfrak{A}_n$ состоит из одного элемента, равного $|A|$.

Замечание. Порядок нумерации сочетаний индексов фиксируется раз навсегда и не связан с выбором матрицы A.

Пример. Пусть

$$A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{Vmatrix}.$$

Перенумеруем все сочетания из четырех индексов $1, 2, 3, 4$ по два, расположив их в следующем порядке:

$$(1\ 2)\ (1\ 3)\ (1\ 4)\ (2\ 3)\ (2\ 4)\ (3\ 4).$$

Тогда

$$\mathfrak{A}_2 = \begin{Vmatrix}
A\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 1 & 4 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 3 & 4 \end{pmatrix}
\end{Vmatrix}.$$

Отметим некоторые свойства ассоциированных матриц.

1°. *Из $C = AB$ следует $\mathfrak{C}_p = \mathfrak{A}_p\mathfrak{B}_p$ $(p = 1, 2, ..., n)$.*

Действительно, выражая миноры p-го порядка $(1 \leqslant p \leqslant n)$ матрицы-произведения C через миноры того же порядка матриц-сомножителей по формуле (18), будем иметь

$$C\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} = \sum_{1 \leqslant l_1 < l_2 < \ldots < l_p \leqslant n} A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ l_1 & l_2 & \ldots & l_p \end{pmatrix}\begin{pmatrix} l_1 & l_2 & \ldots & l_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} \tag{32}$$

$$(1 \leqslant i_1 < i_2 < ... < i_p \leqslant n, \quad 1 \leqslant k_1 < k_2 < ... < k_p \leqslant n).$$

Очевидно, в обозначениях этого параграфа равенство (32) может быть записано так:

$$\mathfrak{c}_{\alpha\beta} = \sum_{\lambda=1}^{N} \mathfrak{a}_{\alpha\lambda}\mathfrak{b}_{\lambda\beta} \quad (\alpha, \beta = 1, 2, ..., N)$$

(здесь α, β, λ — номера сочетаний индексов $i_1 < i_2 < ... < i_p$, $k_1 < k_2 < ... < k_p$, $l_1 < l_2 < ... < l_p$). Отсюда

$$\mathfrak{C}_p = \mathfrak{A}_p \mathfrak{B}_p \quad (p = 1, 2, ..., n).$$

2°. *Из* $B = A^{-1}$ *следует* $\mathfrak{B}_p = \mathfrak{A}_p^{-1}$ $(p = 1, 2, ..., n)$.

Это предложение непосредственно вытекает из предыдущего, если там положить $C = E$ и обратить внимание на то, что $\mathfrak{C}_p$ — единичная матрица порядка $N = C_n^p$.

Из предложения 2° вытекает важная формула, выражающая миноры обратной матрицы через миноры данной матрицы:

если $B = A^{-1}$, то при любых $1 \leqslant i_1 < i_2 < ... < i_p \leqslant n$, $1 \leqslant k_1 < k_2 < ... < k_p \leqslant n$

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \frac{(-1)^{\sum_{\nu=1}^p i_\nu + \sum_{\nu=1}^p k_\nu} A\begin{pmatrix} k'_1 & k'_2 & \dots & k'_{n-p} \\ i'_1 & i'_2 & \dots & i'_{n-p} \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}}, \qquad (33)$$

где $i_1 < i_2 < ... < i_p$ вместе с $i'_1 < i'_2 < ... < i'_{n-p}$, а $k_1 < k_2 < ... < k_p$ вместе с $k'_1 < k'_2 < ... < k'_{n-p}$ составляют полную систему индексов $1, 2, ..., n$.

Действительно, из $AB = E$ следует

$$\mathfrak{A}_p \mathfrak{B}_p = \mathfrak{C}_p,$$

или, в более подробной записи,

$$\sum_{\alpha=1}^{N} \mathfrak{a}_{\gamma\alpha} \mathfrak{b}_{\alpha\beta} = \delta_{\gamma\beta} = \begin{cases} 1, & \gamma = \beta, \\ 0, & \gamma \neq \beta. \end{cases} \qquad (34)$$

Равенства (34) могут быть записаны еще так:

$$\sum_{1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n} A\begin{pmatrix} j_1 & j_2 & \dots & j_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} B\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} =$$

$$= \begin{cases} 1, & \text{если} \quad \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 = 0, \\ 0, & \text{если} \quad \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 > 0 \end{cases} \qquad (34')$$

$$(1 \leqslant j_1 < j_2 < ... < j_p \leqslant n, \quad 1 \leqslant k_1 < k_2 < ... < k_p \leqslant n).$$

С другой стороны, применяя к определителю $|A|$ известные разложения Лапласа, получаем

$$\sum_{1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n} A\begin{pmatrix} j_1 & j_2 & \dots & j_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} \cdot (-1)^{\sum_{\nu=1}^p i_\nu + \sum_{\nu=1}^p k_\nu} A\begin{pmatrix} k'_1 & k'_2 & \dots & k'_{n-p} \\ i'_1 & i'_2 & \dots & i'_{n-p} \end{pmatrix} =$$

$$= \begin{cases} |A|, & \text{если} \quad \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 = 0, \\ 0, & \text{если} \quad \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 > 0, \end{cases} \qquad (35)$$

где $i'_1 < i'_2 < \ldots < i'_{n-p}$ вместе с $i_1 < i_2 < \ldots < i_p$, а $k'_1 < k'_2 < \ldots < k'_{n-p}$ вместе с $k_1 < k_2 < \ldots < k_p$ образуют полную систему индексов $1, 2, \ldots, n$. Сопоставление (35) с (34′) и (34) показывает, что равенства (34) удовлетворятся, если вместо $\mathfrak{b}_{\alpha\beta}$ взять не $B\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}$, а

$$\frac{(-1)^{\sum_{\nu=1}^{p} i_\nu + \sum_{\nu=1}^{p} k_\nu} A\begin{pmatrix} k'_1 & k'_2 & \ldots & k'_{n-p} \\ i'_1 & i'_2 & \ldots & i'_{n-p} \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}}.$$

Так как из системы (34) элементы $\mathfrak{b}_{\alpha\beta}$ обратной матрицы для $\mathfrak{A}_p$ определяются однозначно, то имеет место равенство (33).

§ 5. Обращение прямоугольных матриц. Псевдообратная матрица

Если A — квадратная и невырожденная матрица, то для нее существует обратная матрица A^{-1}. Если же A — не квадратная, а прямоугольная $m \times n$-матрица $(m \neq n)$ или квадратная, но вырожденная, то матрица A не имеет обратной и символ A^{-1} не имеет смысла. Однако, как будет показано далее, для произвольной прямоугольной матрицы A существует “псевдообратная” матрица A^+, которая обладает некоторыми свойствами обратной матрицы и имеет важные применения при решении системы линейных уравнений. В случае, когда A — квадратная неенная матрица, псевдообратная матрица A^+ совпадает с обратной A^{-1} [26]).

1. Скелетное разложение матрицы. В дальнейшем мы будем пользоваться представлением произвольной прямоугольной $m \times n$-матрицы $A = \|a_{ik}\|$ ранга r в виде произведения двух матриц B и C, имеющих соответственно размеры $m \times r$ и $r \times n$:

$$A = BC = \left\| \begin{matrix} b_{11} & \ldots & b_{1r} \\ b_{21} & \ldots & b_{2r} \\ \ldots & \ldots & \ldots \\ b_{m1} & \ldots & b_{mr} \end{matrix} \right\| \left\| \begin{matrix} c_{11} & c_{12} & \ldots & c_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ c_{r1} & c_{r2} & \ldots & c_{rn} \end{matrix} \right\| \quad (r = r_A). \tag{36}$$

Здесь ранги сомножителей B и C обязательно равны рангу произведения A: $r_B = r_C = r$. Действительно (см. с. 21), $r \leqslant r_B, r_C$. Но ранги r_B и r_C не могут превосходить r, так как r — один из размеров матриц B и C. Поэтому $r_B = r_C = r$.

Для того чтобы получить разложение (36), достаточно в качестве столбцов матрицы B взять любые r линейно независимых столбцов матрицы A либо любые r линейно независимых столбцов, через которые линейно выражаются столбцы матрицы A [27]). Тогда произвольный j-й столбец матрицы A будет линейной комбинацией столбцов матрицы B с коэффициентами $c_{1j}, c_{2j}, \ldots, c_{rj}$; эти коэффи-

[26]) Приведенное далее в этом параграфе определение псевдообратной матрицы было дано в 1920 г. Муром [214], указавшим на важные применения этого понятия. Позже независимо от Мура в несколько иной форме псевдообратная матрица определялась и исследовалась в работах Бьерхаммара [159], Пенроуза [226,a] и других авторов.

[27]) Мы исходим из известного положения: в матрице A ранга r имеется r линейно независимых столбцов, через которые линейно (т. е. в виде линейных комбинаций с числовыми коэффициентами из данного поля) выражаются все остальные столбцы. Аналогичное утверждение имеет место и для строк. Подробнее об этом см. гл. III, § 1.

циенты и образуют j-й столбец матрицы C ($j = 1, ..., n$, см. с. 16)[28]).

Поскольку матрицы B и C имеют максимально возможный ранг r, то квадратные матрицы B^*B и CC^* являются невырожденными:

$$|B^*B| \neq 0, \quad |CC^*| \neq 0. \tag{37}$$

Действительно, пусть столбец x — произвольное решение уравнения

$$B^*Bx = 0. \tag{38}$$

Умножим это уравнение слева на строку x^*. Тогда $x^*B^*Bx = (Bx)^*Bx = 0$. Отсюда[29]) следует $Bx = 0$ и (поскольку Bx — линейная комбинация линейно независимых столбцов матрицы B; ср. с формулой ($13''$)) $x = 0$. Из того, что уравнение (38) имеет только нулевое решение $x = 0$, вытекает, что $|B^*B| \neq 0$. Аналогично устанавливается второе неравенство (37)[30]).

Разложение (36) будем называть *скелетным* разложением матрицы A.

2. Существование и единственность псевдообратной матрицы. Рассмотрим матричное уравнение

$$AXA = A. \tag{39}$$

Если A — квадратная невырожденная матрица, то это уравнение имеет единственное решение $X = A^{-1}$. Если же A — произвольная прямоугольная $m \times n$-матрица, то искомое решение X имеет размеры $n \times m$, но не определяется однозначно. В общем случае уравнение (39) имеет бесчисленное множество решений. Ниже будет показано, что среди этих решений имеется только одно, обладающее тем свойством, что его строки и столбцы являются линейными комбинациями соответственно строк и столбцов сопряженной матрицы A^*. Именно это решение мы будем называть псевдообратной матрицей для A и обозначать через A^+.

Определение 5. Матрица A^+ размера $n \times m$ называется *псевдообратной* для $m \times n$-матрицы A, если выполняются равенства[31])

$$AA^+A = A, \tag{40}$$

$$A^+ = UA^* = A^*V, \tag{41}$$

где U и V — некоторые матрицы.

Докажем сначала, что для данной матрицы A не может существовать двух различных псевдообратных матриц A_1^+ и A_2^+. Действительно, из равенств

$$AA_1^+A = AA_2^+A = A, \quad A_1^+ = U_1A^* = A^*V_1, \quad A_2^+ = U_2A^* = A^*V_2,$$

полагая $D = A_2^+ - A_1^+$, $U = U_2 - U_1$, $V = V_2 - V_1$, найдем

$$ADA = 0, \quad D = UA^* = A^*V.$$

[28]) Совершенно так же строками матрицы C могут быть любые r строк, через которые выражаются в виде линейных комбинаций все строки матрицы A. Тогда коэффициенты этих линейных комбинаций образуют строки матрицы B.

[29]) См. конец § 3.

[30]) Неравенства (37) также непосредственно следуют из формулы Бине–Коши. Согласно этой формуле определитель $|B^*B|$ ($|CC^*|$) равен сумме квадратов модулей всех миноров r-го порядка матрицы B (соответственно C).

[31]) Условия (41) означают, что строки (столбцы) матрицы A^+ являются линейными комбинациями строк (столбцов) матрицы A^* (см. сноску на с. 16). Условия (41) могут быть заменены одним условием $A^+ = A^*WA^*$, где W — некоторая матрица (см. конец § 2).

Отсюда

$$(DA)^*DA = A^*D^*DA = A^*V^*ADA = 0$$

и, следовательно (см. конец § 3),

$$DA = 0.$$

Но тогда $DD^* = DAU^* = 0$, т. е. $D = A_2^+ - A_1^+ = 0$.

Для того чтобы установить существование матрицы A^+, мы воспользуемся скелетным разложением (36) и будем искать сначала псевдообратные матрицы B^+ и C^+ [32]). Так как по определению должны иметь место равенства

$$BB^+B = B, \quad B^+ = \widehat{U}B^*, \tag{42}$$

где $\widehat{U}$ — некоторая матрица, то

$$B\,\widehat{U}\,B^*B = B.$$

Умножая слева на B^* и замечая, что B^*B — невырожденная квадратная матрица, найдем

$$\widehat{U} = (B^*B)^{-1}.$$

Но тогда второе из равенств (42) дает искомое выражение для B^+:

$$B^+ = (B^*B)^{-1}B^*. \tag{43}$$

Совершенно аналогично найдем

$$C^+ = C^*(CC^*)^{-1}. \tag{44}$$

Покажем теперь, что матрица

$$A^+ = C^+B^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^* \tag{45}$$

удовлетворяет условиям (40), (41) и, следовательно, является псевдообратной матрицей для A.

В самом деле,

$$AA^+A = BCC^*(CC^*)^{-1}(B^*B)^{-1}B^*BC = BC = A.$$

С другой стороны, из равенств (43)–(45) с учетом равенства $A^* = C^*B^*$, полагая $K = (CC^*)^{-1}(B^*B)^{-1}$, находим

$$A^+ = C^*KB^* = C^*K(CC^*)^{-1}CC^*B^* = UC^*B^* = UA^*,$$
$$A^+ = C^*KB^* = C^*B^*B(B^*B)^{-1}KB^* = C^*B^*V = A^*V,$$

где

$$U = C^*K(CC^*)^{-1}C, \quad V = B(B^*V)^{-1}KB^*.$$

Таким образом, доказано, что для произвольной прямоугольной матрицы A существует одна и только одна псевдообратная матрица A^+, которая определяется формулой (45), где B и C — сомножители в скелетном разложении $A = BC$ матрицы A [33]). Из самого определения псевдообратной матрицы непосредственно следует, что в случае квадратной невырожденной матрицы A псевдообратная матрица A^+ совпадает с обратной A^{-1}.

---

[32]) Из определения 5 сразу следует, что если $A = 0$, то и $A^+ = 0$. Поэтому в дальнейшем предполагается, что $A \neq 0$, и потому $r = r_A > 0$.

[33]) Разложение (36) не определяет однозначно сомножителей BC. Однако поскольку, как было доказано, существует только одна псевдообратная матрица A^+, формула (45) при всех скелетных разложениях матрицы A дает одно и то же значение для A^+. В гл. II, § 5 будет изложен другой метод вычисления псевдообратной матрицы, использующий разбиение исходной матрицы на блоки (см. формулу (101) на с. 62).

Пример. Пусть

$$A = \begin{Vmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{Vmatrix}.$$

Здесь $r = 2$. Примем в качестве столбцов матрицы B первые два столбца матрицы A. Тогда

$$A = BC = \begin{Vmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{Vmatrix} \begin{Vmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{Vmatrix}$$

и

$$B^*B = \begin{Vmatrix} 2 & -3 \\ -3 & 6 \end{Vmatrix}, \quad (B^*B)^{-1} = \begin{Vmatrix} 2 & 1 \\ 1 & 2/3 \end{Vmatrix},$$

$$CC^* = \begin{Vmatrix} 3 & 0 \\ 0 & 3 \end{Vmatrix}, \quad (CC^*)^{-1} = \begin{Vmatrix} 1/3 & 0 \\ 0 & 1/3 \end{Vmatrix} = \frac{1}{3}\, E.$$

Поэтому согласно формуле (45)

$$A^+ = \frac{1}{3} \begin{Vmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 1 & 1 \end{Vmatrix} \begin{Vmatrix} 2 & 1 \\ 1 & 2/3 \end{Vmatrix} \begin{Vmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \end{Vmatrix} = \begin{Vmatrix} 1/3 & 0 & 1/3 \\ 1/9 & 1/9 & 2/9 \\ 2/9 & -1/9 & 1/9 \\ 4/9 & 1/9 & 5/9 \end{Vmatrix}.$$

3. Свойства псевдообратной матрицы. Отметим следующие свойства псевдообратной матрицы:

1°) $(A^*)^+ = (A^+)^*$;

2°) $(A^+)^+ = A$;

3°) $(AA^+)^* = AA^+,\ (AA^+)^2 = AA^+$;

4°) $(A^+A)^* = A^+A,\ (A^+A)^2 = A^+A$.

Первое свойство означает, что операции перехода к сопряженной и к псевдообратной матрице перестановочны между собой. Равенство 2°) выражает собой взаимность понятия псевдообратной матрицы, так как, согласно 2°), псевдообратной матрицей для A^+ является исходная матрица A. Согласно равенствам 3°) и 4°) матрицы AA^+ и A^+A являются эрмитовыми и *инволютивными* (квадрат каждой из этих матриц равен самой матрице).

Для вывода равенства 1°) воспользуемся скелетным разложением (36): $A = BC$. Тогда равенство $A^* = C^*B^*$ дает скелетное разложение матрицы A^*. Поэтому, заменяя в формуле (45) матрицу B на C^*, а матрицу C на B^*, получим

$$(A^*)^+ = B(B^*B)^{-1}(CC^*)^{-1}C = [C^*(CC^*)^{-1}(B^*B)^{-1}B^*]^* = (A^+)^*.$$

Равенства $A^+ = C^+B^+$, $B^+ = (B^*B)^{-1}B^*$, $C^+ = C^*(CC^*)^{-1}$ являются скелетными разложениями. Следовательно,

$$(A^+)^+ = (B^+)^+(C^+)^+ = (B^*)^+B^*BCC^*(C^*)^+.$$

Используя свойство 1°), а также выражения для B^+ и C^+, найдем

$$(A^+)^+ = B(B^*B)^{-1}B^*BCC^*(CC^*)^{-1}C = BC = A.$$

Справедливость равенств 3°) и 4°) проверяется непосредственно путем подстановки в эти равенства вместо A^+ соответствующего выражения из формулы (45).

Заметим, что в общем случае, когда разложение $A = BC$ не является скелетным, не всегда имеет место равенство $A^+ = C^+B^+$. Так, например,

$$A = \|1\| = \|0\ 1\| \left\| \begin{matrix} 1 \\ 1 \end{matrix} \right\| = BC.$$

Здесь

$$A^+ = A^{-1} = \|1\|, \quad B^+ = \|0\ 1\|^+ = (\|1\| \cdot \|0\ 1\|)^+ = \left\| \begin{matrix} 0 \\ 1 \end{matrix} \right\| \|1\| = \left\| \begin{matrix} 0 \\ 1 \end{matrix} \right\|,$$

$$C^+ = \left\| \begin{matrix} 1 \\ 1 \end{matrix} \right\|^+ = \left(\left\| \begin{matrix} 1 \\ 1 \end{matrix} \right\| \cdot \|1\| \right)^+ = \|1\| \cdot \|2\|^{-1} \|1\ 1\| = \|1/2\ \ 1/2\|.$$

Поэтому

$$C^+B^+ = \|1/2\ \ 1/2\| \left\| \begin{matrix} 0 \\ 1 \end{matrix} \right\| = \|1/2\| \neq A^+.$$

4. Наилучшее приближенное решение (по методу наименьших квадратов). Рассмотрим произвольную систему линейных уравнений

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = y_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = y_m, \end{array} \tag{46}$$

или, в матричной записи,

$$Ax = y. \tag{46$'$}$$

Здесь $y_1, y_2, ..., y_m$ — заданные числа, а $x_1, x_2, ..., x_n$ — искомые. В общем случае система (46) может быть и несовместной.

Столбец

$$x^0 = (x_1^0, x_2^0, ..., x_n^0) \tag{47}$$

называется *наилучшим приближенным решением* системы (46), если при значениях $x_1 = x_1^0,\ x_2 = x_2^0,\ ...,\ x_n = x_n^0$ "квадратичное отклонение"

$$|y - Ax|^2 = \sum_{i=1}^{m} \left| y_i - \sum_{k=1}^{n} a_{ik}x_k \right|^2 \tag{48}$$

достигает своего наименьшего значения и среди всех столбцов x, для которых это отклонение имеет минимальное значение, столбец x^0 имеет наименьшую "длину", т. е. для этого столбца величина

$$|x|^2 = x^*x = \sum_{i=1}^{n} |x_i|^2 \tag{49}$$

имеет наименьшее значение.

Покажем, что система (46) всегда имеет одно и только одно наилучшее приближенное решение и это приближенное решение определяется по формуле

$$x^0 = A^+y, \tag{50}$$

где A^+ — псевдообратная матрица для матрицы A.

Для этого рассмотрим произвольный столбец x и положим

$$y - Ax = u + v,$$

где

$$u = y - Ax^0 = y - AA^+y, \quad v = A(x^0 - x). \tag{51}$$

Тогда

$$|y - Ax|^2 = (y - Ax)^*(y - Ax) = (u + v)^*(u + v) = u^*u + v^*u + u^*v + v^*v. \tag{52}$$

Но

$$v^*u = (x^0 - x)^*A^*(y - AA^+y) = (x^0 - x)^*(A^* - A^*AA^+)y. \tag{53}$$

Исходя из разложения (36) и формулы (45) найдем

$$A^*AA^+ = C^*B^*BCC^*(CC^*)^{-1}(B^*B)^{-1}B^* = C^*B^* = A^*.$$

Поэтому из равенства (53) следует

$$v^*u = 0, \tag{54}$$

но тогда и

$$u^*v = (v^*u)^* = 0. \tag{54'}$$

Поэтому из равенства (52) находим

$$|y - Ax|^2 = |u|^2 + |v|^2 = |y - Ax^0|^2 + |A(x^0 - x)|^2, \tag{55}$$

и, следовательно, для любого столбца x

$$|y - Ax| \geqslant |y - Ax^0|. \tag{56}$$

Пусть теперь

$$|y - Ax| = |y - Ax^0|;$$

тогда, согласно равенству (55),

$$Az = 0, \tag{57}$$

где

$$z = x - x^0.$$

С другой стороны,

$$|x|^2 = (x^0 + z)^*(x^0 + z) = |x^0|^2 + |z|^2 + (x^0)^*z + z^*x^0. \tag{58}$$

Вспоминая, что $A^+ = A^*V$ (см. определение 5), получим в силу (57)

$$(x^0)^*z = (A^+y)^*z = (A^*Vy)^*z = y^*V^*Az = 0. \tag{59}$$

Но тогда и

$$z^*x^0 = ((x^0)^*z)^* = 0.$$

Поэтому из равенства (58) находим

$$|x|^2 = |x^0|^2 + |z|^2$$

и, следовательно,

$$|x|^2 \geqslant |x^0|^2,$$

причем знак = имеет место только при $z = 0$, т. е. при $x = x^0$, где $x^0 = A^+y$.

Пример. Найти наилучшее приближенное решение (по методу наименьших квадратов) системы линейных уравнений:

$$\begin{aligned} x_1 - x_2 + 2x_3 \phantom{{}+x_4} &= 3,\\ -x_1 + 2x_2 - 3x_3 + x_4 &= 6,\\ x_2 - x_3 + x_4 &= 0. \end{aligned}$$

Здесь

$$A = \begin{Vmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{Vmatrix}.$$

Но тогда (см. пример на с. 33)

$$A^+ = \begin{Vmatrix} 1/3 & 0 & 1/3 \\ 1/9 & 1/9 & 2/9 \\ 2/9 & -1/9 & 1/9 \\ 4/9 & 1/9 & 5/9 \end{Vmatrix},$$

и потому

$$x^0 = \begin{Vmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \\ x_4^0 \end{Vmatrix} = \begin{Vmatrix} 1/3 & 0 & 1/3 \\ 1/9 & 1/9 & 2/9 \\ 2/9 & -1/9 & 1/9 \\ 4/9 & 1/9 & 5/9 \end{Vmatrix} \begin{Vmatrix} 3 \\ 6 \\ 0 \end{Vmatrix}.$$

Следовательно,

$$x_1^0 = 1, \quad x_2^0 = 1, \quad x_3^0 = 0, \quad x_4^0 = 2.$$

Определим норму $\|A\|$ $m \times n$-матрицы $A = \|a_{ik}\|$ как неотрицательное число, задаваемое формулой

$$\|A\|^2 = \sum_{i,k} |a_{ik}|^2. \tag{61}$$

При этом очевидно, что

$$\|A\|^2 = \sum_{k=1}^{n} |A_{\cdot\, k}|^2 = \sum_{i=1}^{n} |A_{i\, \cdot}|^2. \tag{61$'$}$$

Рассмотрим матричное уравнение

$$AX = Y, \tag{62}$$

где A и Y — заданные $m \times n$- и $m \times p$-матрицы, а X — искомая $n \times p$-матрица.

Определим наилучшее приближенное решение X^0 уравнения (62) из условия

$$\|Y - AX^0\| = \min \|Y - AX\|,$$

причем в случае, когда

$$\|Y - AX\| = \|Y - AX^0\|,$$

требуется, чтобы

$$\|X^0\| \leqslant \|X\|.$$

Из соотношений

$$\|Y - AX\|^2 = \sum_{k=1}^{p} |Y_{\cdot\, k} - AX_{\cdot\, k}|^2, \tag{63}$$

$$\|X\|^2 = \sum_{k=1}^{p} |X_{\cdot\, k}|^2 \tag{64}$$

следует, что k-й столбец искомой матрицы $X^0_{\cdot\, k}$ должен быть наилучшим приближенным решением системы линейных уравнений

$$AX_{\cdot\, k} = Y_{\cdot\, k}.$$

Поэтому

$$X^0_{.k} = A^+ Y_{.k}.$$

Поскольку это равенство справедливо при любом $k = 1, ..., p$, то

$$X^0 = A^+ Y. \tag{65}$$

Таким образом, уравнение (62) всегда имеет одно и только одно наилучшее приближенное решение, определяемое формулой (65).

В частном случае, когда $Y = E$ — единичная матрица m-го порядка, имеем $X^0 = A^+$. Следовательно, *псевдообратная матрица A^+ является наилучшим приближенным решением (по методу наименьших квадратов) матричного уравнения*

$$AX = E.$$

Это свойство псевдообратной матрицы A^+ может быть принято в качестве ее определения.

5. Метод Гревилля последовательного нахождения псевдообратной матрицы состоит в следующем. Пусть a_k — k-й столбец в $m \times n$-матрице A, $A_k = (a_1, ..., a_k)$ — матрица, образованная первыми k столбцами матрицы A, b_k — последняя строка в матрице A_k^+ $(k = 1, ..., n,\ A_1 = a_1,\ A_n = A)$. Тогда[34])

$$A_1^+ = a_1^+ = \frac{a_1^*}{a_1^* a_1}, \tag{66}$$

и для $k > 1$ имеют место рекуррентные формулы

$$A_k^+ = \begin{pmatrix} B_k \\ b_k \end{pmatrix}, \quad B_k = A_{k-1}^+ - d_k b_k, \quad d_k = A_{k-1}^+ a_k. \tag{67}$$

При этом, если $c_k = a_k - A_{k-1} d_k \neq 0$, то

$$b_k = c_k^+ = (a_k - A_{k-1} d_k)^+; \tag{68}$$

если же $c_k = 0$, т. е. $a_k = A_{k-1} d_k$, то

$$b_k = (1 + d_k^* d_k)^{-1} d_k^* A_{k-1}^+. \tag{69}$$

Предлагаем читателю проверить, что матрица $\begin{pmatrix} B_k \\ b_k \end{pmatrix}$ является псевдообратной для матрицы A_k^+, если матрица B_k и строка b_k определяются формулами (61)–(64). Этот метод не требует вычисления детерминантов и может быть использован для вычисления обратной матрицы.

Пример. Пусть

$$A = \begin{Vmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 2 & -3 & -1 \\ 0 & 1 & 1 \end{Vmatrix}.$$

Заметим, что для каждой вещественной матрицы M мы можем писать M' вместо M^*. Тогда

$$A_1^+ = (A_1' A_1)^{-1} A_1' = \frac{1}{6}\ A_1' = \begin{Vmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & 0 \end{Vmatrix},$$

---

[34]) Если $A_1 = a_1 = 0$, то и $A_1^+ = 0$.

$$d_2 = A_1^+ a_2 = -\frac{3}{2}, \quad c_2 - a_2 - A_1 d_2 = \begin{Vmatrix} 1/2 \\ 1/2 \\ 0 \\ 1 \end{Vmatrix},$$

$$b_2 = c_2^+ = (c_2' c_2)^{-1},\ c_2' = \frac{2}{3}\ c_2' = \begin{Vmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{2}{3} \end{Vmatrix},$$

$$B_2 = A_1^+ - d_2 b_2 = \begin{Vmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 \end{Vmatrix}.$$

Таким образом,

$$A_2^+ = \begin{Vmatrix} 2/3 & 1/3 & 1/3 & 1 \\ 1/3 & 1/3 & 0 & 2/3 \end{Vmatrix}.$$

Далее,

$$d_3 = A_2^+ a_3 = \begin{Vmatrix} 1 \\ 1 \end{Vmatrix}, \quad c_3 = a_3 - A_2 d_3 = 0.$$

Поэтому

$$b_3 = (1 + d_3' d_3)^{-1} d_3' A_2^+ = \begin{Vmatrix} \frac{1}{3} & \frac{1}{3} \end{Vmatrix} A_2^+ = \begin{Vmatrix} \frac{1}{3} & \frac{2}{9} & \frac{1}{9} & \frac{5}{9} \end{Vmatrix}$$

и

$$B_3 = A_2^+ - d_3 b_3 =$$

$$= \begin{Vmatrix} 2/3 & 1/3 & 1/3 & 1 \\ 1/3 & 1/3 & 0 & 2/3 \end{Vmatrix} - \begin{Vmatrix} 1/3 & 2/9 & 1/9 & 5/9 \\ 1/3 & 2/3 & 1/9 & 5/9 \end{Vmatrix} = \begin{Vmatrix} 1/3 & 1/9 & 2/9 & 4/9 \\ 0 & 1/9 & -1/9 & 1/9 \end{Vmatrix},$$

$$A^+ = A_3^+ = \begin{Vmatrix} 1/3 & 1/9 & 2/9 & 4/9 \\ 0 & 1/9 & -1/9 & 1/9 \\ 1/3 & 2/9 & 1/9 & 5/9 \end{Vmatrix}.$$

ГЛАВА II

АЛГОРИТМ ГАУССА И НЕКОТОРЫЕ ЕГО ПРИМЕНЕНИЯ

§ 1. Метод исключения Гаусса

1. Пусть дана система n линейных уравнений с n неизвестными $x_1, x_2, ..., x_n$ и правыми частями $y_1, y_2, ..., y_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= y_n. \end{aligned} \tag{1}$$

В матричной форме эта система может быть записана так:

$$Ax = y. \tag{1'}$$

Здесь $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_n)$ — столбцы и $A = \|a_{ik}\|_1^n$ — квадратная матрица коэффициентов.

Если A — невырожденная матрица, то можно записать

$$x = A^{-1}y \tag{2}$$

или, в развернутом виде,

$$x_i = \sum_{k=1}^{n} a_{ik}^{(-1)} y_k \quad (i = 1, 2, ..., n). \tag{2'}$$

Таким образом, задача вычисления элементов обратной матрицы $A^{-1} = \|a_{ik}^{(-1)}\|_1^n$ эквивалентна задаче решения системы уравнений (1) при л ю б ы х правых частях $y_1, y_2, ..., y_n$. Элементы обратной матрицы определяются формулами (25) гл. I. Однако фактическое вычисление элементов матрицы A^{-1} по этим формулам при большом n весьма затруднительно. Поэтому большое практическое значение имеют эффективные методы вычисления элементов обратной матрицы и, следовательно, решения системы линейных уравнений[1]).

В настоящей главе мы изложим теоретические основы некоторых из этих методов, представляющих собой разновидности метода исключения Гаусса, знакомство с которым у читателя началось еще в курсе алгебры средней школы.

2. Пусть в системе уравнений (1) $a_{11} \neq 0$. Мы исключим x_1 из всех уравнений, начиная со второго, для чего ко второму уравнению почленно прибавим первое, умноженное на $-a_{21}/a_{11}$, к третьему почленно прибавим первое, умноженное на $-a_{31}/a_{11}$, и т. д. После этого система уравнений (1) заменится эквивалентной системой

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + \ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n2}^{(1)}x_2 + \ldots + a_{nn}^{(1)}x_n &= y_n^{(1)}. \end{aligned} \tag{3}$$

[1]) Для подробного ознакомления с этими методами мы рекомендуем обратиться к книге Фаддеевых [33], а также к циклу статей, напечатанных в "Успехах математических наук" (1950 г., т. 5, вып. 3).

Коэффициенты при неизвестных и свободные члены в последних $n-1$ уравнениях определяются формулами

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{i1}}{a_{11}}\, a_{1j}, \quad y_i^{(1)} = y_i - \frac{a_{i1}}{a_{11}}\, y_1 \quad (i, j = 2, ..., n). \tag{3'}$$

Пусть $a_{22}^{(1)} \neq 0$. Тогда таким же образом мы исключим x_2 из последних $n-2$ уравнений системы (3) и получим систему уравнений

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= y_3^{(2)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n3}^{(2)}x_3 + \ldots + a_{nn}^{(2)}x_n &= y_n^{(2)}. \end{aligned} \tag{4}$$

При этом новые коэффициенты и правые части связаны с предыдущими формулами

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}}\, a_{2j}^{(1)}, \quad y_i^{(2)} = y_i^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}}\, y_2^{(1)} \quad (i, j = 3, ..., n). \tag{5}$$

Продолжая этот алгоритм далее, мы на $(n-1)$-м этапе приведем исходную систему (1) к треугольной рекуррентной системе

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= y_3^{(2)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{nn}^{(n-1)}x_n &= y_n^{(n-1)}. \end{aligned} \tag{6}$$

Это приведение выполнимо в том и только том случае, когда в процессе приведения все числа $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, ..., a_{n-1,n-1}^{(n-2)}$ оказываются отличными от нуля.

Изложенный нами алгоритм Гаусса состоит из однотипных операций, которые легко выполняются на современных счетных машинах.

3. Выразим коэффициенты и правые части приведенной системы через коэффициенты и правые части исходной системы (1). При этом мы не будем предполагать, что в процессе приведения все числа $a_{11}, a_{22}^{(1)}, ..., a_{n-1,n-1}^{(n-2)}$ оказываются отличными от нуля, а рассмотрим общий случай, когда первые p из этих чисел отличны от нуля:

$$a_{11} \neq 0, \quad a_{22}^{(1)} \neq 0, \quad ..., \quad a_{pp}^{(p-1)} \neq 0 \quad (p \leqslant n-1), \tag{7}$$

что дает возможность (на p-м этапе приведения) привести исходную систему уравнений к виду

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots\ldots\ldots\ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + \ldots\ldots\ldots\ldots\ldots\ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{pp}^{(p-1)}x_p + \ldots\ldots\ldots\ldots + a_{pn}^{(p-1)}x_n &= y_p^{(p-1)}, \\ a_{p+1,p+1}^{(p)}x_{p+1} + \ldots + a_{p+1,n}^{(p)}x_n &= y_{p+1}^{(p)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n,p+1}^{(p)}x_{p+1} + \ldots + a_{nn}^{(p)}x_n &= y_n^{(p)}. \end{aligned} \tag{8}$$