

Содержание

От издательства	12
Об авторе	13
Колофон	14
Предисловие	15
Глава 1. Введение	17
Что такое линейная алгебра и зачем ее изучать?	17
Об этой книге.....	18
Предварительные требования	19
Математика.....	19
Отношение.....	19
Программирование	20
Математические доказательства в противовес интуитивному пониманию на основе программирования	20
Рабочий код в книге и предназначенный для скачивания онлайн	22
Упражнения по программированию	22
Как пользоваться этой книгой (для учителей и самообучающихся)	23
Глава 2. Векторы. Часть 1	24
Создание и визуализация векторов в NumPy.....	24
Геометрия векторов.....	27
Операции на векторах.....	28
Сложение двух векторов	28
Вычитание двух векторов.....	29
Геометрия сложения и вычитания векторов	30
Умножение вектора на скаляр	31
Сложение скаляра с вектором.....	32
Геометрия умножения вектора на скаляр.....	32

Транспонирование	33
Транслирование векторов в Python	34
Модуль вектора и единичные векторы	35
Точечное произведение векторов	36
Точечное произведение является дистрибутивным	38
Геометрия точечного произведения	39
Другие умножения векторов	40
Адамарово умножение	40
Внешнее произведение	41
Перекрестное и тройное произведения	42
Ортогональное разложение векторов	42
Резюме	46
Упражнения по программированию	46
Глава 3. Векторы. Часть 2	49
Множества векторов	49
Линейно-взвешенная комбинация	50
Линейная независимость	51
Математика линейной независимости	53
Независимость и вектор нулей	54
Подпространство и охват	54
Базис	57
Определение базиса	60
Резюме	61
Упражнения по программированию	62
Глава 4. Применения векторов	64
Корреляция и косинусное сходство	64
Фильтрация временных рядов и обнаружение признаков	67
Кластеризация методом k -средних	68
Упражнения по программированию	71
Упражнения по корреляции	71
Упражнения по фильтрации и обнаружению признаков	73
Упражнения по алгоритму k -средних	75
Глава 5. Матрицы. Часть 1	76
Создание и визуализация матриц в NumPy	76
Визуализация, индексация и нарезка матриц	76
Специальные матрицы	78
Матричная математика: сложение, умножение на скаляр, адямарово умножение	80
Сложение и вычитание	80
«Сдвиг» матрицы	81

Умножение на скаляр и адамарово умножение	82
Стандартное умножение матриц	82
Правила допустимости умножения матриц	83
Умножение матриц	84
Умножение матрицы на вектор	85
Линейно-взвешенные комбинации	86
Результаты геометрических преобразований	86
Матричные операции: транспонирование	88
Обозначение точечного и внешнего произведений	88
Матричные операции: LIVE EVIL (порядок следования операций)	89
Симметричные матрицы	89
Создание симметричных матриц из несимметричных	90
Резюме	91
Упражнения по программированию	92

Глава 6. Матрицы. Часть 2

Нормы матриц	97
След матрицы и норма Фробениуса	99
Пространства матрицы (столбцовое, строчное, нуль-пространство)	100
Столбцовое пространство	100
Строчное пространство	104
Нуль-пространства	104
Ранг	108
Ранги специальных матриц	110
Ранг сложенных и умноженных матриц	112
Ранг сдвинутых матриц	113
Теория и практика	113
Применения ранга	114
В столбцовом пространстве	115
Линейная независимость множества векторов	116
Определитель	117
Вычисление определителя	117
Определитель с линейными зависимостями	119
Характеристический многочлен	119
Резюме	121
Упражнения по программированию	123

Глава 7. Применения матриц

Матрицы ковариаций многопеременных данных	128
Геометрические преобразования посредством умножения матриц на векторы	131
Обнаружение признаков изображения	135
Резюме	138
Упражнения по программированию	138

Упражнения по матрицам ковариаций и корреляций	138
Упражнения по геометрическим преобразованиям	140
Упражнения по обнаружению признаков изображения	142
Глава 8. Обратные матрицы	144
Обратная матрица	144
Типы обратных матриц и условия обратимости	145
Вычисление обратной матрицы	146
Обратная матрица матрицы 2×2	146
Обратная матрица диагональной матрицы	148
Инвертирование любой квадратной полноранговой матрицы	149
Односторонние обратные матрицы	151
Уникальность обратной матрицы	153
Псевдообратная матрица Мура–Пенроуза	154
Численная стабильность обратной матрицы	155
Геометрическая интерпретация обратной матрицы	156
Резюме	158
Упражнения по программированию	158
Глава 9. Ортогональные матрицы и QR-разложение	162
Ортогональные матрицы	162
Процедура Грама–Шмидта	164
QR-разложение	165
Размеры матриц Q и R	166
Почему матрица R является верхнетреугольной	168
QR и обратные матрицы	169
Резюме	169
Упражнения по программированию	170
Глава 10. Приведение строк и LU-разложение	174
Системы уравнений	174
Конвертирование уравнений в матрицы	175
Работа с матричными уравнениями	176
Приведение строк	178
Метод устранения по Гауссу	180
Метод устранения по Гауссу–Жордану	181
Обратная матрица посредством метода устранения по Гауссу–Жордану	182
LU-разложение	183
Взаимообмен строками посредством матриц перестановок	185
Резюме	186
Упражнения по программированию	186

Глава 11. Общие линейные модели и наименьшие квадраты	189
Общие линейные модели.....	190
Терминология.....	190
Настройка общей линейной модели.....	190
Решение общих линейных моделей.....	192
Является ли решение точным?	193
Геометрическая перспектива наименьших квадратов.....	194
В чем причина работы метода наименьших квадратов?	195
Общая линейная модель на простом примере.....	197
Наименьшие квадраты посредством QR-разложения	201
Резюме	202
Упражнения по программированию	203
Глава 12. Применения метода наименьших квадратов	207
Предсказывание количеств велопрокатов на основе погоды.....	207
Регрессионная таблица с использованием библиотеки statsmodels	212
Мультиколлинеарность.....	213
Регуляризация	213
Полиномиальная регрессия.....	215
Поиск в параметрической решетке для отыскания модельных параметров.....	218
Резюме	220
Упражнения по программированию	221
Упражнения по аренде велосипедов	221
Упражнения по мультиколлинеарности	222
Упражнения по регуляризации	223
Упражнение по полиномиальной регрессии.....	224
Упражнения по поиску в параметрической решетке.....	225
Глава 13. Собственное разложение	227
Интерпретации собственных чисел и собственных векторов	228
Геометрия.....	228
Статистика (анализ главных компонент)	229
Подавление шума	230
Уменьшение размерности (сжатие данных).....	231
Отыскание собственных чисел	231
Отыскание собственных векторов	234
Неопределенность собственных векторов по знаку и шкале	235
Диагонализация квадратной матрицы	236
Особая удивительность симметричных матриц.....	238
Ортогональные собственные векторы.....	238
Действительно-значные собственные числа	240

Собственное разложение сингулярных матриц.....	241
Квадратичная форма, определенность и собственные числа	243
Квадратичная форма матрицы.....	243
Определенность	245
$A^T A$ является положительной (полу)определенной.....	245
Обобщенное собственное разложение	246
Резюме	248
Упражнения по программированию	249

Глава 14. Сингулярное разложение..... 254

Общая картина сингулярного разложения	254
Сингулярные числа и ранг матрицы.....	256
Сингулярное разложение на Python	256
Сингулярное разложение и одноранговые «слои» матрицы.....	257
Сингулярное разложение из собственного разложения	259
Сингулярное разложение матрицы $A^T A$	260
Конвертация сингулярных чисел в дисперсию: объяснение	260
Кондиционное число.....	261
Сингулярное разложение и псевдообратная матрица Мура–Пенроуза.....	262
Резюме	263
Упражнения по программированию	264

Глава 15. Применения собственного и сингулярного разложений 268

Анализ главных компонент с использованием собственного и сингулярного разложений.....	268
Математика анализа главных компонент	269
Шаги выполнения PCA.....	271
PCA посредством сингулярного разложения.....	272
Линейный дискриминантный анализ	273
Низкоранговая аппроксимация посредством сингулярного разложения	275
Сингулярное разложение для шумоподавления.....	276
Резюме	276
Упражнения	277
Анализ главных компонент (PCA).....	277
Линейный дискриминантный анализ (LDA)	281
Сингулярное разложение для низкоранговых аппроксимаций.....	285
Сингулярное разложение для шумоподавления в изображениях	287

Глава 16. Краткое руководство по языку Python..... 291

Почему Python и какие есть альтернативы?.....	291
Интерактивные среды разработки.....	292
Использование Python локально и онлайн	292

Работа с файлами исходного кода в Google Colab.....	293
Переменные.....	294
Типы данных	296
Индексация.....	297
Функции	297
Методы в качестве функций	299
Написание своих собственных функций	299
Библиотеки	301
NumPy.....	301
Индексация и нарезка в NumPy.....	302
Визуализация.....	303
Переложение формул в исходный код.....	305
Форматирование печати и F-строки.....	308
Поток управления	309
Компараторы	309
Инструкции if	310
Инструкции elif и else	310
Несколько условий	311
Циклы for.....	312
Вложенные инструкции управления	312
Измерение времени вычислений.....	313
Получение помощи и приобретение новых знаний	313
Что делать, когда дела идут наперекосяк.....	314
Резюме	314
Дополнение А. Теорема о ранге и нульности.....	315
Тематический указатель	317

От издательства

Отзывы и пожелания

Мы всегда рады отзывам наших читателей. Расскажите нам, что вы думаете об этой книге – что понравилось или, может быть, не понравилось. Отзывы важны для нас, чтобы выпускать книги, которые будут для вас максимально полезны.

Вы можете написать отзыв на нашем сайте www.dmkpress.com, зайдя на страницу книги и оставив комментарий в разделе «Отзывы и рецензии». Также можно послать письмо главному редактору по адресу dmkpress@gmail.com; при этом укажите название книги в теме письма.

Если вы являетесь экспертом в какой-либо области и заинтересованы в написании новой книги, заполните форму на нашем сайте по адресу http://dmkpress.com/authors/publish_book/ или напишите в издательство по адресу dmkpress@gmail.com.

Список опечаток

Хотя мы приняли все возможные меры для того, чтобы обеспечить высокое качество наших текстов, ошибки все равно случаются. Если вы найдете ошибку в одной из наших книг, мы будем очень благодарны, если вы сообщите о ней главному редактору по адресу dmkpress@gmail.com. Сделав это, вы избавите других читателей от недопонимания и поможете нам улучшить последующие издания этой книги.

Нарушение авторских прав

Пиратство в интернете по-прежнему остается насущной проблемой. Издательство «ДМК Пресс» очень серьезно относится к вопросам защиты авторских прав и лицензирования. Если вы столкнетесь в интернете с незаконной публикацией какой-либо из наших книг, пожалуйста, пришлите нам ссылку на интернет-ресурс, чтобы мы могли применить санкции.

Ссылку на подозрительные материалы можно прислать по адресу электронной почты dmkpress@gmail.com.

Мы высоко ценим любую помощь по защите наших авторов, благодаря которой мы можем предоставлять вам качественные материалы.

Об авторе

Майк Икс Коэн – адъюнкт-профессор неврологии¹ в Институте Дондерса (Медицинский центр Университета Радбуда) в Нидерландах. Имеет более чем 20-летний опыт преподавания научного программирования, анализа данных, статистики и смежных тем, а также является автором нескольких онлайн-курсов² и учебников. У него подозрительно сухое чувство юмора, и ему нравится все фиолетовое.

¹ См. <https://oreil.ly/Ee23F>.

² См. <https://oreil.ly/BurUH>.

Колофон

Животное на обложке книги «Прикладная линейная алгебра для исследователей данных» – это антилопа ньяла, также именуемая низменной ньялой либо просто ньялой (*Tragelaphus angasii*). Самки и молодые ньялы обычно имеют светло-красновато-коричневый окрас шерсти, в то время как взрослые самцы имеют темно-коричневую или даже сероватую шерсть. Как у самцов, так и у самок есть белые полосы вдоль тела и белые пятна на боках. У самцов спиралевидные рога, которые вырастают до 33 дюймов в длину, а их шерсть намного более лохматая, с длинной бахромой, свисающей от горла до задних конечностей, и гривой густых черных волос вдоль позвоночника. Самки весят около 130 фунтов, тогда как самцы могут весить до 275 фунтов.

Ньялы обитают в лесах Юго-Восточной Африки, ареал которых включает Малави, Мозамбик, Южную Африку, Эсватини, Замбию и Зимбабве. Они пугливы и предпочитают пастись ранним утром, ближе к вечеру либо ночью, проводя большую часть жаркой части дня, отдыхая в укрытии. Ньялы образуют свободные стада численностью до десяти особей, хотя самцы постарше ведут одиночный образ жизни. Они не являются территориальными животными, хотя самцы будут бороться за доминирование во время спаривания.

Ньялы считаются видом, вызывающим наименьшее беспокойство, хотя выпас скота, сельское хозяйство и потеря среды обитания представляют для них угрозу. Многие животные на обложках издательства O'Reilly находятся под угрозой исчезновения, и все они важны для нашего мира.

Иллюстрация на обложке выполнена Карен Монтгомери по мотивам старинной линейной гравюры из *Histoire Naturelle*.

Предисловие

Условные обозначения в книге

В книге используются следующие типографические условные обозначения:

курсивный шрифт

обозначает новые термины, URL-адреса, адреса электронной почты, имена файлов и расширения файлов.

моноширинный шрифт

используется для листингов программ, а также внутри абзацев для ссылки на элементы программ, такие как переменные или имена функций, базы данных, типы данных, переменные среды, инструкции и ключевые слова.



Данный элемент обозначает общее замечание.



Данный элемент обозначает предупреждение или предостережение.

Использование примеров исходного кода

Дополнительные материалы (примеры исходного кода, упражнения и т. д.) доступны для скачивания по адресу <https://github.com/mikexcohen/LinAlg4DataScience>.

Если у вас есть технический вопрос или проблема с использованием примеров исходного кода, то, пожалуйста, отправьте электронное письмо по адресу bookquestions@oreilly.com.

Благодарности

Должен признаться, я действительно не люблю писать разделы с признаниями. И это не потому, что мне не хватает благодарности или я считаю, что мне некого благодарить, – совсем наоборот: у меня слишком много людей, которых нужно поблагодарить, и я не знаю, с чего начать, кого перечислить по имени, а кого пропустить. Должен ли я поблагодарить своих родителей за их роль в формировании меня таким человеком, который написал эту книгу? Возможно, их родителей за то, что они сформировали моих родителей? Помню, как моя учительница в четвертом классе говорила мне, что я, должно быть, стану писателем, когда вырасту. (Я не помню ее имени и не уверен, когда я вырасту, но, возможно, она оказала некоторое влияние на эту книгу.) Я написал большую часть этой книги во время поездок на Канарские острова с работой на удалении; возможно, мне следует поблагодарить пилотов, кото-

рые доставили меня туда? Или электриков, которые устанавливали проводку в коворкингах? Вероятно, я должен быть благодарен Оздемиру Паше за его роль в популяризации кофе, который одновременно облегчал и отвлекал меня от писательства. И давайте не будем забывать фермеров, которые выращивали вкусную еду, которая меня поддерживала и делала счастливым.

Видите, к чему это ведет: мои пальцы печатали, но потребовалась вся история человеческой цивилизации, чтобы создать меня и среду, которая позволила мне написать эту книгу – и которая позволила вам прочитать эту книгу. Таким образом, спасибо человечеству!

Ну да ладно, я также могу посвятить один абзац более традиционному разделу благодарностей. Самое главное, я благодарен всем моим студентам на моих курсах в университете и летней школе с живым преподаванием, а также на моих онлайн-курсах Udemu за то, что они доверили мне преподавание у них и мотивировали меня продолжать совершенствовать свои объяснения прикладной математики и других технических тем. Я также благодарен Джесс Хаберману, редактору отдела закупок в O'Reilly, которая установила «первый контакт», чтобы спросить, не буду ли я заинтересован в написании этой книги. Шира Эванс (редактор разработки), Джонатан Оуэн (производственный редактор), Элизабет Оливер (выпускающий редактор), Кристен Браун (менеджер по контентным услугам) и два эксперта – технических рецензента сыграли непосредственную роль в трансформации моих нажатий клавиш в книгу, которую вы сейчас читаете. Уверен, что этот список неполон, потому что другие люди, которые помогли опубликовать эту книгу, мне неизвестны либо потому, что я забыл их из-за потери памяти в моем преклонном возрасте¹. Благодарности всем читающим этот текст, кто чувствует, что внес хотя бы ничтожный вклад в эту книгу.

¹ ЛОЛ, когда я написал эту книгу, мне было 42 года.

Глава 1

Введение

ЧТО ТАКОЕ ЛИНЕЙНАЯ АЛГЕБРА И ЗАЧЕМ ЕЕ ИЗУЧАТЬ?

Линейная алгебра имеет интересную историю в математике, восходящую к XVII веку на Западе и гораздо раньше в Китае. Матрицы – развернутые таблицы чисел, лежащие в основе линейной алгебры, – использовались для обеспечения компактной системы счисления с целью хранения наборов чисел, таких как геометрические координаты (это было первоначальное применение матриц Декартом), и систем уравнений (впервые введенных Гауссом). В XX веке матрицы и векторы использовались для многопеременной математики, включая математическое исчисление, дифференциальные уравнения, физику и экономику.

Но до недавнего времени большинству людей не нужно было заботиться о матрицах. Однако, как оказалось, компьютеры чрезвычайно эффективны в работе с матрицами. И поэтому современные вычисления породили современную линейную алгебру. Современная линейная алгебра является вычислительной, в то время как традиционная линейная алгебра является абстрактной. Современную линейную алгебру лучше всего изучать на исходном коде и приложениях в области графики, статистики, науки о данных, искусственного интеллекта и численного моделирования, в то время как традиционная линейная алгебра изучается на доказательствах и размышлениях о бесконечномерных пространствах векторов. Современная линейная алгебра предоставляет структурные элементы, которые поддерживают почти каждый алгоритм, реализованный на компьютерах, в то время как традиционная линейная алгебра зачастую является интеллектуальной пищей для продвинутых студентов математических университетов.

Добро пожаловать в современную линейную алгебру.

Стоит ли вам изучать линейную алгебру? Это зависит от того, хотите ли вы понимать алгоритмы и процедуры или же просто применять методы, разработанные другими. Я не хочу принижать последнее – в сущности, нет ничего

плохого в использовании инструментов, которые вы не понимаете (я пишу это на ноутбуке, который я могу использовать, но не смог создать его с нуля). Но, учитывая, что вы читаете книгу с таким названием в коллекции книг O'Reilly, я предполагаю, что вы (1) хотите узнать принцип работы алгоритмов либо (2) хотите разрабатывать или адаптировать вычислительные методы. Так что да, вы должны изучать линейную алгебру, и вы должны изучить ее современную версию.

ОБ ЭТОЙ КНИГЕ

Предназначение этой книги – научить вас современной линейной алгебре. Но речь идет не о том, чтобы запомнить несколько ключевых уравнений и пройти по абстрактным доказательствам; цель в том, чтобы научить вас думать о матрицах, векторах и операциях на них. Вы разовьете геометрическое понимание на уровне интуиции относительно того, почему линейная алгебра такова, какая она есть. И вы поймете, как реализовывать концепции линейной алгебры в рабочем коде Python, уделяя особое внимание приложениям в области машинного обучения и науки о данных.

Во многих традиционных учебниках по линейной алгебре избегаются числовые примеры в интересах обобщений, ожидается, что вы самостоятельно получите сложные доказательства, и преподается множество концепций, которые имеют мало отношения либо вообще не имеют отношения к применению или реализации на компьютерах. Я пишу эти строки не как критику – абстрактная линейная алгебра прекрасна и элегантна. Но если ваша цель – использовать линейную алгебру (и математику в целом) в качестве инструмента для понимания данных, статистики, глубокого обучения, обработки изображений и т. д., то традиционные учебники по линейной алгебре, возможно, покажутся досадной тратой времени, которая поставит вас в замешательство и обеспокоит, взяв под сомнение ваши потенциальные возможности в технической области.

Эта книга написана с учетом того, что учащиеся занимаются самообразованием. Возможно, у вас есть степень в области математики, инженерии или физики, но вам нужно научиться реализовывать линейные алгоритмы в рабочем коде. Или же, вполне вероятно, вы не изучали математику в университете и теперь осознаете важность линейной алгебры для вашей учебы или работы. В любом случае, эта книга является самостоятельным ресурсом; она не представляет собой исключительно дополнение к лекционному курсу (хотя ее можно было бы использовать для этой цели).

Если, читая последние три абзаца, вы кивнули головой в знак согласия, то эта книга определена для вас.

Если вы хотите погрузиться в линейную алгебру поглубже, с большим количеством доказательств и разведывательной работы, то существует несколько отличных учебников, о прочтении которых вы можете подумать,

включая мой учебник «Линейная алгебра: теория, интуиция, исходный код (Sincxpress BV)¹.

ПРЕДВАРИТЕЛЬНЫЕ ТРЕБОВАНИЯ

Я пытался написать эту книгу для увлеченных учеников с минимальной формальной подготовкой. И тем не менее ничего никогда не изучается по-настоящему с нуля.

Математика

Вы должны чувствовать себя комфортно в математике уровня средней школы. Просто основы алгебры и геометрии, и ничего особенного.

Для этой книги требуется абсолютно нулевой уровень в исчислении (хотя дифференциальное исчисление имеет большую важность для приложений, в которых линейная алгебра используется часто, таких как глубокое обучение и оптимизация).

Но самое главное – вам нужно чувствовать себя комфортно, думая о математике, рассматривая уравнения и графики и не боясь противостоять интеллектуальным трудностям, которые возникают при изучении математики.

Отношение

Линейная алгебра – это раздел математики, и, следовательно, данная книга посвящена математике. Изучение математики, в особенности во взрослом возрасте, требует определенного терпения, целеустремленности и напористости. Выпейте чашку кофе, сделайте глубокий вдох, уберите свой телефон в другую комнату и погрузитесь в работу.

В глубине вашей головы будет звучать голос, говорящий о том, что вы слишком стары или слишком глупы, чтобы изучать продвинутую математику. Иногда этот голос будет звучать громче, а иногда мягче, но он всегда будет присутствовать. И это не только у вас – он есть у всех. Этот голос невозможно подавить или уничтожить; даже не пытайтесь. Просто примите, что некоторая неуверенность в себе – это часть человеческого бытия. Всякий раз, когда этот голос заговаривает, вам приходится доказывать, что он неправ.

¹ Приношу извинения за бесстыдную саморекламу; обещаю, что в данной книге это единственный случай, когда я позволяю себе подобное послабление.

Программирование

Данная книга посвящена линейно-алгебраическим приложениям в рабочем коде. Я написал эту книгу, ориентируясь на Python, потому что Python в настоящее время является наиболее широко используемым языком в науке о данных, машинном обучении и смежных областях. Если вы предпочитаете другие языки, такие как MATLAB, R, C или Julia, то я надеюсь, что вам будет легко переложить рабочий код Python на эти языки.

Я постарался сделать код Python как можно проще, оставляя его при этом релевантным для приложений. Глава 16 содержит базовое введение в программирование на Python. Стоит ли вам просматривать эту главу? Все зависит от вашего уровня владения языком Python:

Средний/продвинутый уровень (опыт программирования > 1 года)

Полностью пропустите главу 16 либо, по возможности, пролистайте ее, чтобы получить представление о типе рабочего кода, который появится в остальной части книги.

Некоторые знания (опыт < 1 года)

Рекомендую проработать эту главу на случай, если в ней есть новый материал или вам нужно его освежить. Но вы должны быть в состоянии пройти ее довольно быстро.

Полный новичок

Подробно ознакомьтесь с этой главой. Поймите, что данная книга не является полным руководством по Python, поэтому если вы обнаружите, что вам трудно разобраться с рабочим кодом в главах книги, то, возможно, вы захотите отложить эту книгу, поработать со специальным курсом или книгой по Python, а затем вернуться к этой книге.

МАТЕМАТИЧЕСКИЕ ДОКАЗАТЕЛЬСТВА В ПРОТИВОВЕС ИНТУИТИВНОМУ ПОНИМАНИЮ НА ОСНОВЕ ПРОГРАММИРОВАНИЯ

Цель изучения математики в общем и целом состоит в том, чтобы понять математику. Как понять математику? Давайте посчитаем способы:

Строгие доказательства

Доказательство в математике – это последовательность утверждений, демонстрирующая то, что набор допущений приводит к логическому заключению. Доказательства, несомненно, имеют большую важность в чистой математике.

Визуализации и примеры

Четко написанные объяснения, диаграммы и численные примеры помогут приобретать понимание концепций и операций в линейной алгебре на уровне интуиции. Для удобства визуализации большинство примеров выполнено в 2D либо 3D, но принципы применимы и к более высоким измерениям.

Разница между ними заключается в том, что формальные математические доказательства обеспечивают строгость, но редко понимание на уровне интуиции, в то время как визуализации и примеры обеспечивают прочное понимание на уровне интуиции благодаря практическому опыту, но рискуют приводить к неточностям, будучи основанными на конкретных примерах, которые не обобщаются.

Доказательства важных утверждений включены, но я больше концентрируюсь на упрочении интуитивного понимания посредством объяснений, визуализаций и примеров исходного кода.

И это подводит меня к пониманию математики на уровне интуиции на основе программирования (тому, что я иногда называю «мягкими доказательствами»). Вот идея: вы исходите из того, что в Python (и таких библиотеках, как NumPy и SciPy) правильно реализована низкоуровневая обработка чисел, в то время как вы концентрируетесь на принципах путем обследования многочисленных числовых примеров в исходном коде.

Краткий пример: мы «мягко докажем» принцип коммутативности умножения, который гласит, что $a \times b = b \times a$:

```
a = np.random.randn()
b = np.random.randn()
a * b - b * a
```

Приведенный выше исходный код генерирует два случайных числа и проверяет гипотезу о том, что изменение порядка умножения не влияет на результат. Если принцип коммутативности истинен, то в третьей строке будет выведено значение 0.0. Если вы выполните этот исходный код несколько раз и всегда будете получать 0.0, то, увидев один и тот же результат в многочисленных примерах с разными числами, вы приобретете понимание коммутативности на уровне интуиции.

Для ясности: понимание на уровне интуиции на основе исходного кода не заменяет строгого математического доказательства. Все дело в том, что «мягкие доказательства» позволяют понимать математические концепции, не беспокоясь о деталях абстрактного математического синтаксиса и аргументов. Это особенно выгодно для программистов, которым не хватает продвинутого математического образования.

В сухом остатке все сводится к тому, что *много математических концепций могут усваиваться при помощи небольшой порции исходного кода.*

РАБОЧИЙ КОД В КНИГЕ И ПРЕДНАЗНАЧЕННЫЙ ДЛЯ СКАЧИВАНИЯ ОНЛАЙН

Данную книгу можно читать, не глядя на исходный код и не решая упражнения по программированию. Все в порядке; вы обязательно чему-нибудь научитесь. Но не расстраивайтесь, если ваши знания будут поверхностными и мимолетными. Если вы хотите действительно *понять* линейную алгебру, то вам нужно решать задачи. Вот почему эта книга снабжена демонстрациями исходного кода и упражнениями по каждой математической концепции.

Важный исходный код напечатан непосредственно в книге. Я хочу, чтобы вы читали текст и уравнения, смотрели на графики и одновременно *видели исходный код*. Это позволит вам связывать концепции и уравнения с исходным кодом.

Но листинг исходного кода в книге может занимать много места, а ручное его копирование на вашем компьютере будет утомительным. Поэтому на страницах книги печатаются только ключевые строки кода; онлайн-код содержит дополнительный исходный код, комментарии, графические украшения и т. д. Онлайн-код также содержит решения упражнений по программированию (всех упражнений, а не только выборочных задач!). Вам обязательно следует скачать исходный код и ознакомиться с ним во время работы с книгой.

Весь исходный код можно получить из репозитория на сайте GitHub <https://github.com/mikexcohen/LinAlg4DataScience>. Этот репозиторий можно клонировать либо просто скачать целиком в виде ZIP-файла (вам не потребуется регистрироваться, входить в систему либо платить за скачивание кода).

Я написал исходный код в среде Google Colab, используя блокнот Jupyter. Я решил использовать Jupyter, потому что это дружественная и простая в применении среда. Тем не менее призываю вас использовать любую интегрированную среду разработки на Python, которую вы предпочитаете. Для удобства онлайн-код также предоставляется в виде простых файлов .ру.

УПРАЖНЕНИЯ ПО ПРОГРАММИРОВАНИЮ

Математика – это не зрительский вид спорта. В большинстве книг по математике есть бесчисленное множество письменных задач (и давайте будем честны: никто не решает их все). Но данная книга целиком посвящена *прикладной* линейной алгебре, и никто не применяет линейную алгебру на бумаге! Вместо этого линейная алгебра применяется в рабочем коде. Поэтому вместо ручных задач и утомительных доказательств, «оставленных читателю в качестве упражнения» (как любят писать авторы учебников по математике), в данной книге много упражнений по программированию.

Упражнения по программированию различаются по сложности. Если вы в Python и линейной алгебре – новичок, то некоторые упражнения, возмож-

но, покажутся вам действительно сложными. Если вы застряли, то вот вам совет: кратко взгляните на мое решение, чтобы вдохновиться, затем уберите его, чтобы не видеть мой исходный код, и продолжайте работать над своим исходным кодом.

Сравнивая свое решение с моим, имейте в виду, что существует много способов решения задач на Python. Важно найти правильный ответ; предпринимаяемые вами шаги, чтобы его получить, нередко зависят от личного стиля программирования.

КАК ПОЛЬЗОВАТЬСЯ ЭТОЙ КНИГОЙ (для учителей и самообучающихся)

Данная книга полезна в трех средах:

Самообучение

Я постарался сделать эту книгу доступной для читателей, которые хотят изучать линейную алгебру самостоятельно, вне формальной классной комнаты. Никаких дополнительных ресурсов или онлайн-лекций не требуется, хотя, конечно же, существует масса других книг, веб-сайтов, видеороликов YouTube и онлайн-курсов, которые студенты могут счесть полезными.

Первичный учебник на занятиях по науке о данных

Данная книга может использоваться в качестве первичного учебника в курсе математики, лежащей в основе науки о данных, машинного обучения, искусственного интеллекта и смежных тем. Содержимое состоит из 14 глав (исключая это введение и дополнительный материал по Python), и от студентов ожидается, что они будут работать над одной-двумя главами в неделю. Поскольку учащиеся имеют доступ к решениям всех упражнений, преподаватели, возможно, пожелают дополнить упражнения книги дополнительными наборами задач.

Вторичный учебник по математикоориентированному курсу линейной алгебры

Эта книга также может использоваться в качестве дополнения к курсу математики с сильным акцентом на доказательствах. В этом случае лекции были бы сосредоточены на теории и строгих доказательствах, в то время как на данную книгу можно было бы ссылаться с целью переложения концепций в исходный код с прицелом на приложения в науке о данных и машинном обучении. Как я написал выше, преподаватели, возможно, пожелают предоставить дополнительные упражнения, поскольку решения ко всем упражнениям из книги доступны онлайн.

Глава 2

Векторы. Часть 1

Векторы обеспечивают основу, на которой построена вся линейная алгебра (и, следовательно, остальная часть этой книги).

К концу этой главы вы будете знать о векторах все: что они собой представляют, что они делают, как их интерпретировать и как создавать и работать с ними на Python. Вы поймете наиболее важные операции на векторах, включая векторную алгебру и точечное произведение. Наконец, вы узнаете о разложениях векторов, являющихся одной из главных целей линейной алгебры.

СОЗДАНИЕ И ВИЗУАЛИЗАЦИЯ ВЕКТОРОВ В NUMPY

В линейной алгебре *вектор* – это упорядоченный список чисел. (В абстрактной линейной алгебре векторы могут содержать другие математические объекты, включая функции; однако поскольку данная книга ориентирована на приложения, мы будем рассматривать только те векторы, которые содержат числа.)

Векторы обладают несколькими важными характеристиками. Первые две, с которых мы начнем, – это¹:

размерность.

Число чисел в векторе;

ориентация.

Ориентирован ли вектор *вдоль столбца* (стоя в полный рост) либо *вдоль строки* (лежа ровно во всю ширь).

Размерность нередко указывается с помощью причудливо выглядящей записи \mathbb{R}^N , где \mathbb{R} обозначает действительно-значные числа (ср. с символом

¹ В книге проводится четкое различие между созвучными понятиями *dimensionality* (математическая размерность, протяженность, объемность) и *dimension* (геометрическое измерение, мерность как в слове двухмерный). Первый термин переводится как размерность, а второй в зависимости от контекста – как измерение либо мерность. – *Прим. перев.*

Конец ознакомительного фрагмента.

Приобрести книгу можно

в интернет-магазине

«Электронный универс»

e-Univers.ru