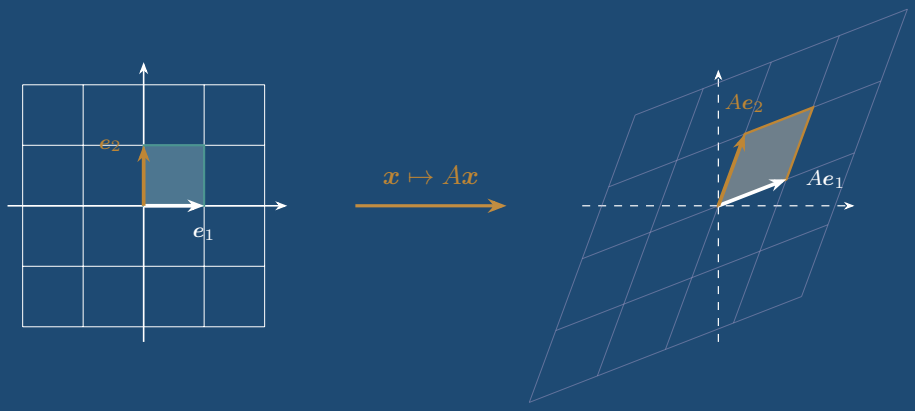
线性代数

从 $Ax = b$ 到奇异值分解：四个子空间的故事



原课程主讲：Gilbert Strang

独立中文整理本

MIT OPENCOURSEWARE · 18.06SC

线性代数逐讲讲义

从 $Ax = b$ 到奇异值分解：四个子空间的故事

原课程主讲：Gilbert Strang

原课程：MIT 18.06SC　2011 年秋季学期

数学基础自学讲义项目　·　2026 年 11 月修订

依据 MIT OpenCourseWare 材料独立整理与扩充

来源、版本与使用说明

原课程：MIT OpenCourseWare 18.06SC Linear Algebra，2011 年秋季学期
原课程主讲：Gilbert Strang

课程页面：https://ocw.mit.edu/courses/18-06sc-linear-algebra-fall-2011/

中文整理与扩充：数学基础自学讲义项目
整理本版本：2026 年 11 月修订版　**修订日期**：2026-11-04

本书依据原课程的课时摘要、习题、解答与考试材料独立整理，补充中文讲解、证明步骤、例题与练习。它是学习辅助材料，不代表 MIT 或原课程教师发布、审定或背书；各讲的具体依据见「对应材料」。

参考读物：

- 原课程配套读物：Gilbert Strang, *Introduction to Linear Algebra*, 4th/5th ed., Wellesley-Cambridge Press。章节对应以 OCW 课时页面所列版本为准。
- 延伸阅读：Sheldon Axler, *Linear Algebra Done Right*, 4th ed., Springer。

许可与使用：原课程材料的署名与许可按其来源页面标注。本整理本的改写与新增内容采用署名–非商业性使用–相同方式共享 4.0 国际许可协议（CC BY-NC-SA 4.0）。转载或改编时，请保留来源、署名与许可说明，并标明所作修改；第三方材料仍适用各自的许可。课程网站的说明见MIT OCW 使用条款。

排版：XeLaTeX 与 ctex；图形主要由 TikZ 绘制。

前言：从方程组到线性变换

线性方程组 $Ax = b$ 可以从计算和几何两个角度理解。消元给出解的具体表达；把 Ax 看成 A 的列向量的线性组合，则可以看出哪些 b 能由这些列构成，以及解为何可能不唯一。两种视角相互补充，是本书的起点。

本书的路线

第一部分从消元、矩阵乘法和逆矩阵出发，建立基、维数以及四个基本子空间。对实矩阵 $A \in \mathbb{R}^{m\times n}$，行空间与零空间在 $\mathbb{R}^n$ 中正交互补，列空间与左零空间在 $\mathbb{R}^m$ 中正交互补。这个结构把解的存在性、自由变量和秩联系在一起。

第二部分研究正交投影、最小二乘、行列式和特征值。当 b 不在列空间中时，可以先把它投影到列空间，再求最小二乘解；投影向量唯一，但系数解未必唯一。第三部分讨论对称矩阵、复矩阵、奇异值分解和基变换。对实矩阵，奇异值分解把线性变换表示为输入空间的正交变换、沿坐标方向的伸缩和输出空间的正交变换；正交变换可以包含反射。

课程材料与本整理本

本书依据 MIT OpenCourseWare 18.06SC 的课时材料扩充，共 31 讲和三次复习课。书中把例题的矩阵计算与几何解释放在一起，并补充练习、答案和复习材料。

本书是一份独立的中文学习辅助材料。讲义编号与原课程视频编号不一定相同；配合原课学习时，以每讲开头的「对应材料」为准。先修要求和各部分的学习路线列在后面的「如何使用本讲义」中。

如何使用本讲义

一讲的阅读顺序

每讲开头的「对应材料」说明来源；讲次与原课程课次的具体关系以该说明为准。

(1) **看目标与动机**：先明确本讲要解决的问题，再进入定义和定理。

(2) **读证明与例题**：先尝试说出思路，再逐步核对条件和计算；遇到「想一想」时，先自己作答。

(3) **做练习再看答案**：答案用于核对步骤。做错的题应回到相应定义、定理或例题查找原因。

(4) **用小结与复习课自测**：合上正文复述主要结论；完成一部分后，再做相应的 R1、R2、…… 复习课。

练习难度按本讲内容标记：★ 直接应用定义或定理；★★ 组合结论或完成较长计算；★★★ 需要额外构造或推广。星级用于安排学习顺序。

版面元素与编号

标记	读法
定义、公理	青色色条给出精确表述。检查对象、条件和量词；青色粗体标记首次引入的术语。
定理、命题、引理	蓝色色条给出结论。先读条件，再读结论；蓝色粗体提示需要记住的内容。
例　解	橙色方块标记例题，解答以青色 ♦ 结束；可以遮住解答先算一遍。
思路　证明	思路说明证明计划，证明给出论证；证明以蓝色 ■ 结束。
想一想	暂停阅读，检查自己对刚出现的概念或结论的理解，再核对「答」。
直观理解	图像或估算用于帮助理解；正式结论仍以定义、定理及证明为准。
易错提醒	核对常见错误及其原因，尝试说明错误方法在哪一步失效。
核心思想	提炼一节或一讲的方法；复习时结合本讲小结阅读。

「定理 3.2」表示第 3 讲的第 2 条定理类结论；「定理 R1.3」属于复习课 1。复习课放在每部分末尾，包含核心结论、题型例题、常见错误与自测。

本课程的准备与路线

主线先修：高中代数、向量与坐标。课程从 $Ax = b$ 出发，依次研究四个基本子空间、正交投影和矩阵分解，结合几何图像与手算。微分方程与 Fourier 应用需要一元微积分；Jacobian 与 Hessian 应用还用到偏导数和多元 Taylor 展开。相关位置注明了阅读准备和可暂缓的范围。

建议的读法：每讲配合 Strang 对应的课堂视频（第 k 讲对应的 OCW 课时写在「对应材料」里）和教材《Introduction to Linear Algebra》的对应章节。例题里的矩阵都手算一遍，再做该课时的 problem 并对照官方解答。复习课 R1、R2、R3 对应三次单元考试，R3 同时兼顾期末。

各部分的学习目标

部分	讲次	主要内容	学完应能做到
Unit I	1–12，R1	方程组的行图与列图；消元、逆矩阵、基、维数和四个基本子空间。	从消元结果写出全部解，判断存在性与唯一性，核对四个子空间的维数。
Unit II	13–23，R2	投影、最小二乘、QR、行列式、特征值与对角化；微分方程、Markov 矩阵与 Fourier 级数。	区分投影与系数解的唯一性；用矩阵分解处理迭代、拟合和线性微分方程。
Unit III	24–31，R3	对称与正定矩阵；复矩阵与 FFT；Jordan 形、SVD、线性变换、基变换与伪逆。	检查分解条件和矩阵尺寸；解释奇异值与低秩近似，计算不同基下的矩阵表示。

常用符号与记号约定

除复矩阵章节另作说明外，本书默认实数矩阵与列向量。分解公式的存在性、尺寸及具体条件见对应讲次。

记号	含义与约定	查阅位置
$A \in \mathbb{R}^{m\times n},\ x \in \mathbb{R}^n$	m 行 n 列矩阵与列向量；坐标元组可简写为 $(x_1,\dots,x_n)$。	第 1–3 讲
$0,\ e_i,\ I_n$	零向量或零矩阵、标准基向量、n 阶单位矩阵；尺寸由上下文确定。	第 1–5 讲
$A^{\mathsf{T}},\ A^{\mathrm{H}}$	转置与共轭转置；$A^{\mathrm{H}} = \overline{A}^{\mathsf{T}}$，实矩阵时二者相同。	第 5、25 讲
$A^{-1},\ A^{+}$	逆矩阵仅对可逆方阵定义；A^{+} 为 Moore–Penrose 伪逆。	第 3、31 讲
$\operatorname{span}\{v_1,\dots,v_k\}$	所列向量的全部线性组合构成的子空间。	第 6、9 讲
$\operatorname{rank} A,\ \dim V$	矩阵的秩与向量空间的维数。	第 9–10 讲
$C(A),\ N(A)$	列空间与零空间，分别位于 $\mathbb{R}^m$、$\mathbb{R}^n$。	第 6 讲
$C(A^{\mathsf{T}}),\ N(A^{\mathsf{T}})$	行空间与左零空间；分别是 $N(A)$、$C(A)$ 在相应欧氏空间中的正交补。	第 10、13 讲
$x^{\mathsf{T}}y,\ \|x\|$	实点积与欧氏范数，$\|x\|^2 = \sum_i \lvert x_i\rvert^2$；复数情形使用共轭转置。	第 13、25 讲
$V^{\perp}$	在指定内积空间中与 V 的每个向量均正交的向量全体。	第 13 讲
$P,\ p = Pb$	正交投影矩阵与投影向量；实数情形 $P^{\mathsf{T}} = P$、$P^2 = P$。公式 $A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$ 要求 A 满列秩。	第 14–15 讲
$A = QR$	正交分解；薄 QR 中，满列秩 A 对应列标准正交的 Q 和可逆上三角 R。	第 16 讲
$A = LU,\ PA = LU$	消元分解及含行交换的形式；L 单位下三角、U 上三角。无换行分解的条件见正文。	第 4 讲
$\det A,\ \operatorname{tr} A$	方阵的行列式与迹；实矩阵的体积倍率是 $\lvert\det A\rvert$。	第 17–20 讲
$Ax = \lambda x$	特征值与特征向量关系，要求 $x \neq 0$ 且 A 为方阵。	第 20 讲
$A = S\Lambda S^{-1}$	仅在 A 可对角化时使用；S 的列组成特征向量基。	第 21 讲
$A = Q\Lambda Q^{\mathsf{T}}$	实对称矩阵的谱分解；Q 为正交矩阵。	第 24 讲
$x^{\mathrm{H}}y,\ \|x\|$	复内积与范数：$x^{\mathrm{H}}y = \sum_i \overline{x_i}y_i$，$\|x\|^2 = x^{\mathrm{H}}x$。	第 25 讲
$A = U\Sigma V^{\mathsf{T}},\ \sigma_i$	任意实矩阵的奇异值分解；U, V 正交，Σ 为矩形对角矩阵，非零对角元为正奇异值。	第 28 讲

目录

前言：从方程组到线性变换 ii

如何使用本讲义 iii

常用符号与记号约定 v

第一部分　$Ax = b$ 与四个基本子空间 1

第 1 讲　线性方程组的几何 2

1.1　一个方程组，三种看法 . . . 2
1.2　向量与线性组合 . . . 4
1.3　列图像：列向量的组合 . . . 6
1.4　矩阵乘向量：一条公式，两种算法 . . . 8
1.5　哪些右端能够达到 . . . 10
练习 . . . 13

第 2 讲　用矩阵做消元 16

2.1　先问清楚：什么操作不会丢掉解 . . . 16
2.2　从增广矩阵到上三角系统 . . . 18
2.3　零出现在主元位置时 . . . 20
2.4　用一个矩阵记录一步消元 . . . 22
2.5　交换行与交换列：矩阵乘在哪一边 . . . 25
2.6　原习题综合：Pascal 矩阵的逐层消元 . . . 26
练习 . . . 28

第 3 讲　矩阵乘法与逆矩阵 31

3.1　矩阵乘法：行乘列 . . . 31
3.2　同一个乘积的列、行与外积 . . . 33
3.3　运算律与分块：可以改什么次序 . . . 34
3.4　逆矩阵：把每个输入准确恢复 . . . 36
3.5　Gauss–Jordan：同时求多个右端 . . . 38
3.6　方阵可逆的等价检验 . . . 40

练习 . . . 42

第 4 讲 LU 分解 45

4.1 乘积的逆与转置 . . . 46
4.2 从 $EA = U$ 到 $A = LU$ . . . 48
4.3 为什么乘子直接进入 L . . . 50
4.4 用 LU 解方程组，LDU 与唯一性 . . . 54
4.5 消元要花多少代价 . . . 57
4.6 换行与置换矩阵 . . . 59
练习 . . . 61

第 5 讲 转置、置换与实向量子空间 64

5.1 转置与对称结构 . . . 64
5.2 置换矩阵与行的重排 . . . 67
5.3 实向量空间与子空间 . . . 68
5.4 张成、列空间与线性方程 . . . 70
5.5 矩阵组成的向量空间 . . . 71
5.6 把判据用于新的集合 . . . 73
练习 . . . 75

第 6 讲 列空间与零空间 77

6.1 子空间回顾 . . . 77
6.2 两个子空间的交、并与和 . . . 79
6.3 列空间 $C(A)$ . . . 82
6.4 什么时候 $Ax = b$ 有解 . . . 84
6.5 零空间 $N(A)$ . . . 87
6.6 $Ax = b$ 的全部解 . . . 89
6.7 构造子空间的两种方式 . . . 92
练习 . . . 94

第 7 讲 齐次方程：主元变量、自由变量与特殊解 96

7.1 第二列没有主元，还能消元吗？ . . . 96
7.2 自由变量如何产生全部解？ . . . 98
7.3 特殊解为什么足以描述零空间？ . . . 99
7.4 简化行阶梯形怎样直接给出零空间？ . . . 101
7.5 零矩阵、满列主元与宽矩阵 . . . 104
7.6 原课习题：消元结果与列关系互相核对 . . . 105

练习 . . . 107

第 8 讲 非齐次方程：相容条件、特解与全部解 109

8.1 零行怎样限制右端？ . . . 109
8.2 一个特解怎样给出全部解？ . . . 111
8.3 右端必须跟着系数一起变换 . . . 113
8.4 有解与唯一：两件事分别由什么控制？ . . . 115
8.5 四种秩情形各有多少解？ . . . 116
8.6 特解、特殊解与同一解集 . . . 118
8.7 同一个右端下，矩阵是否相同？ . . . 118
练习 . . . 119

第 9 讲 线性无关、基与维数 121

9.1 哪些方向可以被其他方向替代？ . . . 121
9.2 张成回答的是哪些向量能够得到 . . . 123
9.3 基让每个向量有且只有一套坐标 . . . 124
9.4 维数是每一组基都有的向量数 . . . 126
9.5 列空间的基：主元列要回到原矩阵 . . . 127
9.6 零空间的基：特殊解记录自由方向 . . . 128
9.7 两道综合题：从关系到几何 . . . 129
练习 . . . 131

第 10 讲 四个基本子空间 133

10.1 四个子空间分别在哪里？ . . . 133
10.2 行最简形怎样给出两种基？ . . . 135
10.3 两个零空间怎样给出基？ . . . 136
10.4 行空间与零空间为何正交？ . . . 138
10.5 秩如何分配四个空间的维数？ . . . 139
10.6 $Ax = b$ 何时有解？ . . . 140
10.7 矩阵集合也有基和维数 . . . 142
练习 . . . 143

第 11 讲 矩阵空间、秩 1 矩阵与小世界图 145

11.1 矩阵空间 . . . 145
11.2 子空间的交、并、和与维数公式 . . . 149
11.3 函数空间：微分方程的解 . . . 151
11.4 秩 1 矩阵 . . . 153

11.5 图距离与小世界 . . . 156
练习 . . . 158

第 12 讲 图、网络与关联矩阵 161
12.1 图、网络与边的方向 . . . 161
12.2 关联矩阵记录端点差 . . . 162
12.3 连通性决定秩与循环数 . . . 164
12.4 Kirchhoff 电流定律与循环流 . . . 167
12.5 Ohm 定律与电路方程 . . . 168
练习 . . . 171

复习课 1 线性方程、消元与子空间 173
R1.1 知识地图 . . . 173
R1.2 核心结论回顾 . . . 174
R1.3 题型与例题 . . . 176
R1.4 常见错误 . . . 179
练习 . . . 180

第二部分 正交、行列式与特征值 182

第 13 讲 正交向量与子空间 183
13.1 点积记录了什么? . . . 183
13.2 垂直意味着什么? . . . 185
13.3 一个子空间的正交补是什么? . . . 187
13.4 四个基本子空间怎样互相垂直? . . . 189
13.5 为什么 $A^{\mathsf{T}}A$ 记住了零空间? . . . 192
练习 . . . 194

第 14 讲 投影到子空间 196
14.1 投影点为什么是最近点? . . . 196
14.2 怎样计算投影到一条直线? . . . 198
14.3 怎样投影到列空间? . . . 200
14.4 补空间如何表示? . . . 203
14.5 方程无解时为什么仍要投影? . . . 205
练习 . . . 206

第 15 讲　投影矩阵与最小二乘 **208**

15.1 投影矩阵有哪些代数特征? . . . 208
15.2 无解时怎样找最近的列组合? . . . 211
15.3 $A^{\mathsf{T}}A$ 什么时候可逆? . . . 213
15.4 怎样拟合三个测量点? . . . 214
15.5 拟合值与误差分别属于哪里? . . . 216
练习 . . . 218

第 16 讲　正交矩阵与 Gram–Schmidt 正交化 **220**

16.1 正交与正交归一：垂直之外还要单位长度 . . . 220
16.2 正交矩阵：长度、内积与距离保持 . . . 221
16.3 两个向量的 Gram–Schmidt 正交化 . . . 225
16.4 多个向量：逐个减去已有方向 . . . 226
16.5 QR 分解：把列写成正交方向的组合 . . . 228
练习 . . . 230

第 17 讲　行列式的性质 **232**

17.1 行列式要回答的问题 . . . 232
17.2 三条基本性质 . . . 233
17.3 三角矩阵与消元计算 . . . 236
17.4 乘积公式：把第二个矩阵固定 . . . 238
17.5 转置、列规则与置换奇偶性 . . . 239
17.6 两道原题：行和与 Vandermonde 矩阵 . . . 241
练习 . . . 243

第 18 讲　行列式公式与代数余子式 **246**

18.1 从三条性质出发：2×2 公式 . . . 246
18.2 3×3：27 块里只剩 6 块 . . . 248
18.3 排列公式与定义的存在性 . . . 249
18.4 代数余子式：把排列项按位置分组 . . . 251
18.5 三对角矩阵：展开产生低阶递推 . . . 252
18.6 改变一个元素：代数余子式给出变化率 . . . 254
练习 . . . 256

第 19 讲　Cramer 法则、逆矩阵与体积 **258**

19.1 逆矩阵的代数余子式公式 . . . 258
19.2 从公式读出结构 . . . 261

19.3 Cramer 法则：每个坐标是两个行列式之比 . . . 262
19.4 面积：先分清符号与大小 . . . 263
19.5 体积与线性变换 . . . 265
19.6 Jacobi 行列式：局部的小盒子 . . . 266
练习 . . . 268

第 20 讲 特征值与特征向量 271
20.1 方向不变的向量 . . . 271
20.2 从几何直接读出特征值：投影与反射 . . . 274
20.3 特征方程：先求倍数，再求方向 . . . 275
20.4 迹与行列式：两个快速检查 . . . 276
20.5 旋转为什么需要复数 . . . 277
20.6 重根与三角矩阵 . . . 278
20.7 从已知特征值推导新矩阵 . . . 279
练习 . . . 280

第 21 讲 对角化与矩阵幂 283
21.1 把特征方向选作坐标轴 . . . 283
21.2 什么时候有足够多的特征方向 . . . 284
21.3 对角化把矩阵幂变成标量幂 . . . 285
21.4 差分方程：初值决定各方向的系数 . . . 287
21.5 把二阶递推改写成一阶系统 . . . 288
21.6 同一方法求解其他递推 . . . 290
21.7 检查对角化并处理特殊情形 . . . 291
练习 . . . 292

第 22 讲 微分方程与 e^{At} 295
22.1 从 $du/dt = au$ 到 $du/dt = Au$ . . . 295
22.2 用特征向量求解：几个完整的例子 . . . 298
22.3 解耦与矩阵指数 e^{At} . . . 300
22.4 稳定性 . . . 305
22.5 二阶方程与长度守恒 . . . 308
练习 . . . 310

第 23 讲 Markov 矩阵与 Fourier 级数 313
23.1 列和为 1 意味着什么？ . . . 313
23.2 稳态何时吸引所有初值？ . . . 315

23.3 从正交坐标到函数坐标 . . . 319
23.4 三角函数的正交性与 Fourier 系数 . . . 319
练习 . . . 323

复习课 2 正交、行列式与线性系统 325
R2.1 知识地图 . . . 325
R2.2 核心结论回顾 . . . 326
R2.3 题型与例题 . . . 329
R2.4 常见错误 . . . 333
练习 . . . 334

第三部分 正定矩阵、奇异值分解与线性变换 336

第 24 讲 对称矩阵与正定性 337
24.1 对称矩阵有什么特别 . . . 337
24.2 实特征值与正交特征向量 . . . 339
24.3 谱定理：用正交矩阵对角化 . . . 341
24.4 谱分解：对称矩阵是投影的组合 . . . 343
24.5 主元的符号与特征值的符号 . . . 345
24.6 正定矩阵 . . . 347
练习 . . . 350

第 25 讲 复矩阵与快速 Fourier 变换 352
25.1 复数回顾与单位根 . . . 352
25.2 复向量的长度与内积 . . . 354
25.3 Hermite 矩阵 . . . 356
25.4 酉矩阵 . . . 358
25.5 Fourier 矩阵与离散 Fourier 变换 . . . 360
25.6 快速 Fourier 变换：偶数项与奇数项 . . . 362
练习 . . . 365

第 26 讲 正定矩阵与极小值 367
26.1 二次型的正负和几何 . . . 367
26.2 2×2：四种判据是一件事 . . . 368
26.3 更高维：消元把能量写成平方和 . . . 370
26.4 Hessian 与局部极小值 . . . 372
26.5 非对称矩阵的两种情形 . . . 375

练习 . . . 377

第 27 讲 相似矩阵与 Jordan 标准形 379

27.1 先补一块：$A^{\mathsf{T}}A$ 是正定的 . . . 379

27.2 相似矩阵就是换一组基 . . . 381

27.3 Jordan 块记录特征向量不足 . . . 384

27.4 重根为什么会分成不同的相似类 . . . 388

练习 . . . 390

第 28 讲 奇异值分解 392

28.1 哪些方向会被送到正交方向？ . . . 392

28.2 每个矩阵都有奇异值分解 . . . 393

28.3 计算：对角化 $A^{\mathsf{T}}A$ . . . 396

28.4 几何意义：正交变换、伸缩、正交变换 . . . 397

28.5 零空间、正交列与两个原题 . . . 399

练习 . . . 402

第 29 讲 线性变换及其矩阵 404

29.1 不用坐标看变换 . . . 404

29.2 线性变换的定义 . . . 405

29.3 矩阵给出的线性变换 . . . 408

29.4 基决定一切 . . . 409

29.5 线性变换的矩阵 . . . 411

29.6 同一变换，换基后矩阵怎样变 . . . 413

29.7 核与像：变换压缩了什么 . . . 414

练习 . . . 416

第 30 讲 基变换与图像压缩 418

30.1 坐标换了，向量变了吗？ . . . 418

30.2 同一个变换为何换了矩阵？ . . . 419

30.3 图像为什么能换一种坐标记录？ . . . 421

30.4 保留重要方向：低秩截断与误差 . . . 424

30.5 怎样挑一组合适的基？ . . . 426

练习 . . . 428

第 31 讲 左逆、右逆与伪逆 430

31.1 三种逆：方阵、左逆与右逆 . . . 430

31.2 左逆、最小二乘与列空间投影 . . . 432

31.3 右逆与行空间投影 433
31.4 伪逆：把行空间送回去 435
31.5 一般最小二乘与最小范数解 438
练习 439

复习课 3 第三单元与全课程复习 442

R3.1 知识地图 442
R3.2 核心结论回顾 443
R3.3 题型与例题 446
R3.4 常见错误 449
R3.5 全课程回顾 450
练习 451

第一部分

$Ax = b$ 与四个基本子空间

第 1 讲

线性方程组的几何

The Geometry of Linear Equations

对应材料：MIT 18.06SC Session 1.1 「The geometry of linear equations」课时摘要（Strang 第 1 讲视频）；Session 1.1 习题 1.1–1.3 及解答；Strang《Introduction to Linear Algebra》§1.1、§1.3、§2.1。

本讲目标

(1) 对同一个线性方程组，能从三个角度看它：行图像（直线、平面的交点）、列图像（列向量的线性组合）、矩阵形式 $Ax = b$。

(2) 理解线性组合 $cv + dw$ 的含义，知道一个、两个、三个向量的全部组合在几何上分别是什么。

(3) 会用两种方法计算矩阵乘向量 Ax：按列（列的组合）和按行（行与 x 的点积），并掌握维数规则。

(4) 能在二维和三维中画出（或想象出）两种图像，判断方程组有唯一解、无解还是有无穷多解。

(5) 理解「是否对每个 b 都能解 $Ax = b$」等价于「A 的列的组合能否填满整个空间」，初步认识奇异矩阵与线性相关。

1.1 一个方程组，三种看法

线性代数的基本问题是解线性方程组：n 个方程，n 个未知数。我们从最小的非平凡情形 $n = 2$ 开始，本讲会反复使用下面这个方程组：

$$\begin{aligned} 2x - \ \ y &= 0, \\ -x + 2y &= 3. \end{aligned} \tag{1.1}$$

这个方程组你在中学就会解。本讲的重点不在于求出答案，而在于用三种不同的方式看待它。这三种看法会贯穿整门课：

- 行图像：一次看一个方程（矩阵的一行）。每个方程是平面上的一条直线，解是直线的交点。
- 列图像：一次看一个未知数（矩阵的一列）。每个未知数乘着一个向量，解是让这些向量「拼出」右端的系数。
- 矩阵形式：把所有系数排成矩阵 A，方程组写成一个式子 $Ax = b$。

行图像是大家熟悉的，列图像大多数人第一次见。到了三维、以及更高的维数，行图像会变

得难以想象，列图像却始终清楚。这是 Strang 在第一讲就强调列图像的原因。

1.1.1 行图像：直线的交点

先看熟悉的行图像。第一个方程 $2x-y=0$ 即 $y=2x$，是过原点、斜率为 2 的直线，经过 $(0,0)$ 和 $(1,2)$。第二个方程 $-x+2y=3$ 即 $y=\frac{1}{2}x+\frac{3}{2}$，与 x 轴交于 $(-3,0)$，与 y 轴交于 $(0,\frac{3}{2})$。两条直线画在图 1.1 中。

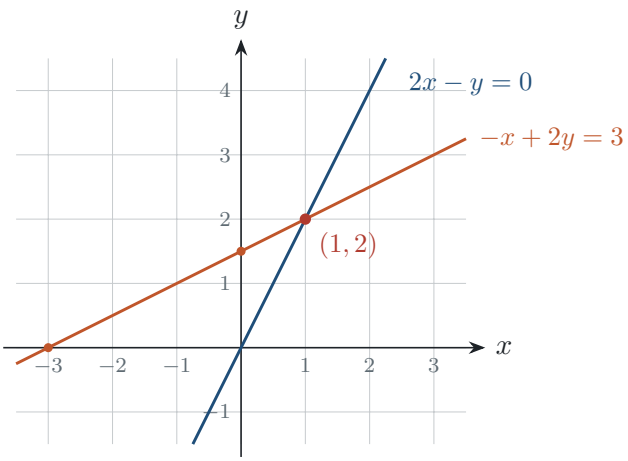


图 1.1　方程组 (1.1) 的行图像。两条直线交于唯一的一点 $(1,2)$，它就是方程组的解。

■ **例 1.1　行图像求解**　用代入法解方程组 (1.1)，并验证所得的点确实在两条直线上。

解　由第一个方程得 $y=2x$。代入第二个方程：

$$-x+2(2x)=3 \implies 3x=3 \implies x=1, \qquad y=2x=2.$$

验证：第一个方程 $2\cdot 1-2=0$ 成立；第二个方程 $-1+2\cdot 2=3$ 成立。所以点 $(1,2)$ 同时在两条直线上，正是图 1.1 中的交点。 ♦

平面上两条直线的位置关系只有三种：相交于一点、平行不重合、完全重合。当两行方程的系数都不全为零时，可以直接用这三种位置关系判断解数。含系数全零行的退化情形将在 §1.4 用统一的代数证明处理。

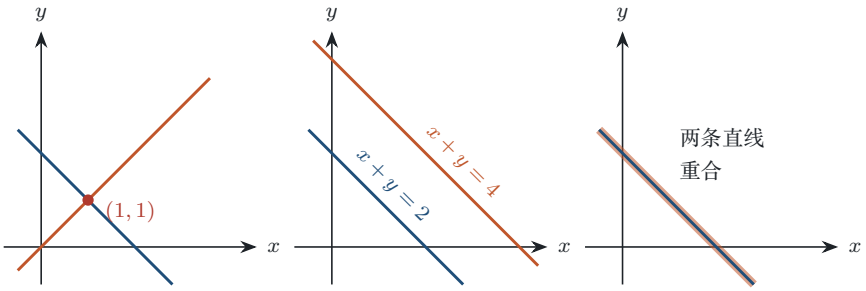
■ **例 1.2　三种结局**　画出下列三个方程组的行图像，并判断各有几个解：

$$\text{(a)}\ \begin{aligned} x+y&=2\\ x-y&=0 \end{aligned} \qquad \text{(b)}\ \begin{aligned} x+y&=2\\ x+y&=4 \end{aligned} \qquad \text{(c)}\ \begin{aligned} x+y&=2\\ 2x+2y&=4 \end{aligned}$$

解　(a) 两式相加得 $2x=2$，所以 $x=1$，再由 $x-y=0$ 得 $y=1$。验证：$1+1=2$，$1-1=0$。两条直线斜率分别为 -1 和 1，不平行，交于唯一一点 $(1,1)$。

(b) 第二式减第一式得 $0=2$，矛盾，所以无解。几何上，两条直线 $y=2-x$ 和 $y=4-x$ 斜率都是 -1，截距不同，是两条平行线，永不相交。

(c) 第二个方程恰好是第一个的 2 倍，两个方程描述的是同一条直线 $y = 2 - x$。直线上每一点都是解：对任意实数 t，$(x, y) = (t,\, 2 - t)$ 满足两个方程（$t + (2 - t) = 2$，$2t + 2(2 - t) = 4$），所以有无穷多个解。图 1.2 并列展示这三种情况。 ♦



(a) 唯一解　(b) 无解　(c) 无穷多解

图 1.2　两条直线的三种位置关系，对应例 1.2 的三个方程组。

想一想　两个方程两个未知数的方程组，能不能恰好有两个解？

答　不能。两条不同的直线至多有一个公共点，因为过两个不同的点只有一条直线。所以只要有两个公共点，两条直线就重合，解就有无穷多个。§1.4 会对任意大小的线性方程组证明同样的结论：解的个数只能是 0、1 或 ∞。

1.2　向量与线性组合

要讲列图像，先要有向量和线性组合这两个概念。它们是整门课的基本词汇。

1.2.1　向量及其两种运算

定义 1.1（向量）　由 n 个实数排成的一列

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

称为一个 n 维列向量（column vector），v_i 是它的第 i 个分量（component）。所有 n 维实向量组成的集合记作 $\mathbb{R}^n$。两个向量相等，当且仅当对应分量全部相等。

$\mathbb{R}^2$ 中的向量 $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ 可以画成从原点指向点 (v_1, v_2) 的箭头，$\mathbb{R}^3$ 同理。为了节省版面，行文中也把列向量写成 (v_1, v_2) 或 (v_1, v_2, v_3)，它们仍然表示列向量。所有分量都是 0 的向量叫零向量（zero vector），记作 0。

对向量只做两种运算：

- **向量加法**：对应分量相加，$v+w=(v_1+w_1,\ \dots,\ v_n+w_n)$。几何上把 w 的起点接到 v 的终点，从原点到 w 终点的箭头就是 $v+w$；也可以看成以 v, w 为邻边的平行四边形的对角线。
- **数乘**：每个分量乘同一个数，$cv=(cv_1,\ \dots,\ cv_n)$。几何上把箭头伸缩 $|c|$ 倍；$c<0$ 时方向反过来。

1.2.2 线性组合

把两种运算合在一起，就得到线性代数里最重要的构造。

> **定义 1.2（线性组合）**　设 $v, w\in\mathbb{R}^n$，c, d 是实数。向量 $cv+dw$ 称为 v 与 w 的一个**线性组合**（linear combination），c, d 称为组合的系数。一般地，$c_1v_1+c_2v_2+\cdots+c_kv_k$ 是 $v_1,\dots,v_k$ 的线性组合。

几个特殊的组合：取 $c=1, d=0$ 得到 v 本身；取 $c=d=0$ 得到零向量，所以**零向量是任何一组向量的线性组合**；取 $c=1, d=-1$ 得到差 $v-w$。

线性组合只允许「数乘，再相加」这两步，没有向量与向量相乘。一个非零向量 v 的全部倍数 cv 组成一条直线，而且这条直线**一定过原点**（取 $c=0$）。所以不过原点的直线，例如 $y=x+1$，不是任何一个向量全部倍数的集合。「一组向量的全部线性组合必定包含零向量」这一点以后会反复用到。

■ **例 1.3　计算线性组合，再反过来求系数**　设 $v=(1,1)$，$w=(2,-1)$。(1) 计算 $3v+2w$ 和 $v-w$。(2) 求 c, d 使 $cv+dw=(4,1)$。(3) 对任意 $b=(b_1,b_2)$，求 c, d 使 $cv+dw=b$。

解　(1) 按分量计算：

$$3\begin{bmatrix}1\\1\end{bmatrix}+2\begin{bmatrix}2\\-1\end{bmatrix}=\begin{bmatrix}3+4\\3-2\end{bmatrix}=\begin{bmatrix}7\\1\end{bmatrix},\qquad \begin{bmatrix}1\\1\end{bmatrix}-\begin{bmatrix}2\\-1\end{bmatrix}=\begin{bmatrix}-1\\2\end{bmatrix}.$$

(2) 方程 $cv+dw=(4,1)$ 按分量写开，是关于 c, d 的两个方程：

$$c+2d=4,\qquad c-d=1.$$

第一式减第二式：$3d=3$，$d=1$；代回 $c-d=1$ 得 $c=2$。验证：$2(1,1)+1\cdot(2,-1)=(4,1)$。

(3) 同样写成 $c+2d=b_1$，$c-d=b_2$。相减得 $3d=b_1-b_2$；第一式加第二式的 2 倍得 $3c=b_1+2b_2$。所以

$$c=\frac{b_1+2b_2}{3},\qquad d=\frac{b_1-b_2}{3}.$$

对 $b=(4,1)$ 代入得 $c=2, d=1$，与 (2) 一致。对 $b=(1,0)$ 得 $c=d=\frac{1}{3}:\frac{1}{3}(1,1)+\frac{1}{3}(2,-1)=(1,0)$。因为对**每一个** b 都解出了 c, d，所以 v 和 w 的线性组合**填满了整个平面** $\mathbb{R}^2$。 ♦

注意 (2) 中求系数的过程：要问「b 是不是 v, w 的组合」，就得解一个方程组。图 1.3 画出几个具体组合；下一节将用这种组合解释列图像。

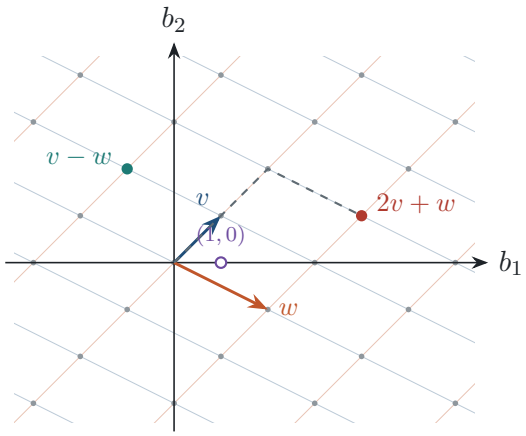


图 1.3　$v=(1,1)$ 与 $w=(2,-1)$ 的线性组合。灰点是整数系数的组合，构成一张斜格网；允许实数系数后，格子之间的点也能取到（如紫色空心点 $(1,0)=\frac{1}{3}v+\frac{1}{3}w$），组合填满整个平面。

直观理解　「全部线性组合」在几何上是什么，取决于向量的个数和它们的相对位置：

- 一个非零向量 v：全部 cv 是一条**过原点的直线**。
- $\mathbb{R}^2$ 中两个不共线的向量：全部 $cv+dw$ 是**整个平面**（例 1.3）。
- $\mathbb{R}^3$ 中两个不共线的向量：全部 $cv+dw$ 是一个**过原点的平面**。
- $\mathbb{R}^3$ 中三个向量：若前两个不共线，且第三个不在它们张成的平面里，全部组合是**整个空间**；退化时可得到平面、直线，或只有原点。

本讲先用这些几何图像理解组合，末节给出线性相关的定义；第 9 讲再系统讨论相关性与生成空间。

想一想　设 $v=(1,2)$，$w=(-2,-4)$。v 与 w 的全部线性组合是什么？$(1,0)$ 是它们的组合吗？

答　$w=-2v$，所以 $cv+dw=(c-2d)v$ 只是 v 的倍数，全部组合是过原点、方向为 $(1,2)$ 的直线 $b_2=2b_1$。$(1,0)$ 不满足 $b_2=2b_1$（$0\neq 2$），所以不是它们的组合。两个向量共线时，组合填不满平面。

1.3　列图像：列向量的组合

1.3.1　把方程组写成一个向量方程

回到方程组 (1.1)。这次不按行读，而是**按列读**：把 x 的系数排成一列，y 的系数排成一列，右端也排成一列：

$$x\begin{bmatrix}2\\-1\end{bmatrix}+y\begin{bmatrix}-1\\2\end{bmatrix}=\begin{bmatrix}0\\3\end{bmatrix}. \tag{1.2}$$

这和原方程组是一回事：左边按分量算出来是 $\begin{bmatrix}2x-y\\-x+2y\end{bmatrix}$，两个向量相等当且仅当对应分量相等，这恰好就是两个原方程。

于是问题换了一种问法：

> 怎样组合两个列向量 $\begin{bmatrix}2\\-1\end{bmatrix}$ 和 $\begin{bmatrix}-1\\2\end{bmatrix}$，才能得到右端 $b=\begin{bmatrix}0\\3\end{bmatrix}$？

我们已经知道答案 $x=1, y=2$。用列图像验证：

$$1\begin{bmatrix}2\\-1\end{bmatrix}+2\begin{bmatrix}-1\\2\end{bmatrix}=\begin{bmatrix}2-2\\-1+4\end{bmatrix}=\begin{bmatrix}0\\3\end{bmatrix}.$$

图 1.4 把它画了出来：先走一份第一列 $(2,-1)$，再走两份第二列，即 $(-2,4)$，正好到达 $b=(0,3)$。

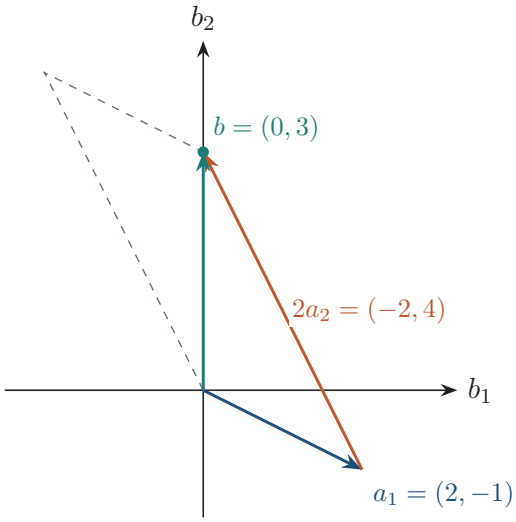


图 1.4　按列组合时，系数 $1,2$ 分别控制两列的倍数；将 a_1 与 $2a_2$ 首尾相接，终点为右端向量 b。

图中的橙色箭头从第一列的终点出发，箭头表示的向量仍为 $(-2,4)$。平移箭头只用于展示加法，不改变它的分量。青色箭头从原点直达最终结果，是两个列向量组合后的输出。未知数现在是两个组合系数，所求的点 $(x,y)=(1,2)$ 和被组合得到的向量 $b=(0,3)$ 属于两个不同的角色，不能将它们混为一谈。

例 1.4　从任意右端反求组合　对任意 $b=(b_1,b_2)$，求出 x,y 使 $x(2,-1)+y(-1,2)=b$。由此判断每个右端是否都能被这两列组合出来。

解　按分量写成 $2x-y=b_1$、$-x+2y=b_2$。第一式乘 2 后加第二式，得 $3x=2b_1+b_2$；第一式加第二式的 2 倍，得 $3y=b_1+2b_2$。故候选为

$$x=\frac{2b_1+b_2}{3},\qquad y=\frac{b_1+2b_2}{3}.$$

代回第一式得 $[2(2b_1+b_2)-(b_1+2b_2)]/3=b_1$，代回第二式得 $[-(2b_1+b_2)+2(b_1+2b_2)]/3=b_2$，所以候选确为解。任何解都必须满足刚才导出的两个等式，故没有别的解。对任意 b 都给出了唯一系数，因而两列的全部组合填满平面。取 $b=(0,3)$ 即回到 $(x,y)=(1,2)$。 ♦

这次结论的量词比只解一个方程组强：先任取 b，再给出系数公式，证明了所有右端均可达到。只找到某个特定 b 的表示，无法得出两列填满平面的结论。例如两个共线向量也能组合成零向量，但它们达不到整张平面。

固定矩阵与变化的右端　「$Ax=b$ 有解」只谈给定的一个 b；「对每个 b 都有解」要求所有目标均可达到。矩阵 $\begin{bmatrix}1&2\\2&4\end{bmatrix}$ 能达到 $b=(3,6)$，却不能达到 $b=(3,7)$，因为输出第二分量总是第一分量的 2 倍。

1.4　矩阵乘向量：一条公式，两种算法

如果每次都把各列分别写成长向量，三个未知数时就占满一行。把这些列依次排进矩阵，组合系数排进向量，便得到 Ax 的记号。矩阵写法节省的是记录工作，保留的仍然是逐行方程与逐列组合的同一个等式。

定义 1.3（矩阵与矩阵乘向量）　有 m 行、n 列的实数表 $A=(a_{ij})$ 称为 $m\times n$ 矩阵。若 $x\in\mathbb{R}^n$，定义 $Ax\in\mathbb{R}^m$ 的第 i 个分量为

$$(Ax)_i=\sum_{j=1}^{n}a_{ij}x_j,\qquad 1\le i\le m.$$

输入的分量个数必须等于矩阵列数，输出分量个数等于行数。例如 2×3 矩阵乘三维向量得到二维向量，不能乘二维向量；方阵只是行数与列数相等的特殊情形。零矩阵把任何输入送到零向量，单位矩阵保持输入不变，这两种边界情形也符合上述求和定义。

命题 1.4（行算法等于列算法）　设 A 的列为 $a_1,\dots,a_n\in\mathbb{R}^m$，$x=(x_1,\dots,x_n)$。则

$$Ax=x_1a_1+\cdots+x_na_n.$$

因此 $Ax=b$ 与 m 个逐行方程 $\sum_j a_{ij}x_j=b_i$ 完全等价。

证明　右边组合的第 i 个分量为 $\sum_j x_ja_{ij}$，实数乘法交换律将它写成 $\sum_j a_{ij}x_j$，恰为 $(Ax)_i$。两个向量的所有对应分量相等，故向量相等。再按向量相等的定义，$Ax=b$ 等价于每个分量均等于 b_i，也就是那 m 个方程。 ■

■ **例 1.5　原课程的两种乘法计算**　分别按列和按行计算

$$\begin{bmatrix}2&5\\1&3\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}.$$

解　列算法先缩放再相加：$1(2,1)+2(5,3)=(12,7)$。行算法分别计算两个点积：第一行给 $2\cdot1+5\cdot2=12$，第二行给 $1\cdot1+3\cdot2=7$。所以输出为 $(12,7)$，两种算法一致。前者直接展示输出怎样由列生成，后者便于逐个算分量。 ♦

■ **例 1.6　原习题 1.2 的三维乘法**　计算 $A(3,-2,1)$，其中

$$A=\begin{bmatrix}1&2&0\\2&0&3\\4&1&1\end{bmatrix}.$$

解　按行核对三个分量：$3-4+0=-1$，$6+0+3=9$，$12-2+1=11$，故输出为 $(-1,9,11)$。按列再次核对：

$$3(1,2,4)-2(2,0,1)+(0,3,1)=(-1,9,11).$$

两次计算先后顺序不同，却必须给出同一结果。若有一个分量不同，应逐行回查该分量所用的列分量与系数，不能因为两个数「很接近」就略过符号错误。 ♦

1.4.1　线性性与解的个数

对固定矩阵 A，两个输入相加后再乘 A，与各自先乘 A 再相加，结果相同。这个事实来自分配律，也决定了线性方程组的解不能恰好只有两个。

命题 1.5（矩阵乘向量的线性性）　设 A 为 $m\times n$ 实矩阵，$u,v\in\mathbb{R}^n$，$c,d\in\mathbb{R}$，则

$$A(cu+dv)=cAu+dAv.$$

证明　逐个比较分量。对 $1\le i\le m$，求和的分配律给出

$$\begin{aligned}\big(A(cu+dv)\big)_i&=\sum_{j=1}^{n}a_{ij}(cu_j+dv_j)\\&=c\sum_{j=1}^{n}a_{ij}u_j+d\sum_{j=1}^{n}a_{ij}v_j=\big(cAu+dAv\big)_i.\end{aligned}$$

所有对应分量都相等，便得到向量等式。取 $c=d=1$ 得加法相容，取 $d=0$ 得数乘相容；取 $c=d=0$ 还得到 $A0=0$。 ■

推论 1.6（解的个数只有三种可能）　实线性方程组 $Ax=b$ 的解集或者为空，或者只有一个元素，或者有无穷多个元素。若 $u\neq v$ 是两个解，则每个实数 t 都给出解 $u+t(v-u)$。

证明　若没有解或只有一个解，已是前两种情况。若有不同的解 u,v，由线性性得

$$A\big(u+t(v-u)\big)=Au+t(Av-Au)=b+t(b-b)=b.$$

还需确认不同参数产生不同的解。因 $v-u\neq0$，它有某个分量非零；若 $u+t(v-u)=u+s(v-u)$，比较该分量即得 $t=s$。实数参数有无穷多个，故解也有无穷多个。 ■

这证明只使用矩阵乘向量的定义，没有要求方阵，也没有要求各行方程都不退化。比如单个方程 $0x=0$ 的每个实数都是解；$0x=1$ 则没有解。把每一行都说成直线时必须排除系数全零的情况，而上面的代数论证同时包含这些边界情形。

■ **例 1.7　两解之间及之外的解**　对方程组 $x+y=2$、$2x+2y=4$，取两个解 $u=(0,2)$、$v=(2,0)$。写出推论中的全部点，并判断它们是否覆盖所有解。

解　代入参数公式得 $u+t(v-u)=(2t,2-2t)$。每一对分量之和都为 2，故满足两式。反之，任意解 (x,y) 都有 $y=2-x$；取 $t=x/2$，即可写成上述形式。所以本例中参数式覆盖所有解。$t\in[0,1]$ 给两点之间的线段，允许所有实数 t 才得到整条直线，例如 $t=-1$ 给出 $(-2,4)$。 ♦

1.4.2　原习题中的矩阵维数

原习题 1.3 问：3×2 矩阵乘 2×3 矩阵，是否得到 3×3 矩阵？只需把右边的矩阵按列看，就能用已经会的矩阵乘向量来回答，无需先背一个新的运算口诀。

若 A 为 $m\times n$ 矩阵，B 为 $n\times p$ 矩阵，把 B 的列写成 $b_1,\dots,b_p$，定义 AB 的列依次为 $Ab_1,\dots,Ab_p$。每个 b_j 有 n 个分量，可以输入 A；每个 Ab_j 有 m 个分量。因此结果有 m 行、p 列。尺寸规则为

$$(m\times n)(n\times p)=m\times p.$$

本讲只用这一按列定义确认尺寸，矩阵乘法的进一步规则留到后面的讲次。

■ **例 1.8　尺寸可乘，不表示交换后相同**　设 A 是 3×2 矩阵，B 是 2×3 矩阵。判断 AB、BA 是否有定义，结果各是什么尺寸。

解　AB 中，B 的每列为二维向量，恰好可以输入 A，输出三维向量；B 有三列，故 AB 为 3×3 矩阵。原习题的断言为真。对 BA，A 的两列是三维向量，可以输入有三列的 B，每列输出二维向量，故 BA 为 2×2。两种顺序都可乘，但尺寸不同，不能写 $AB=BA$。若将 B 改成 4×3，则 AB 连定义所需的尺寸都不满足。 ♦

1.5　哪些右端能够达到

例 1.4 对任意右端都给出了唯一组合系数。现在改看矩阵

$$B=\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}.$$

它的第二列是第一列的两倍，因此 $B(x,y)=(x+2y,2x+4y)$ 的第二个分量始终是第一个的两倍。这个限制与系数取得多大无关：所有输出都在直线 $b_2=2b_1$ 上。

> **定义 1.7（方阵的奇异性与列相关）**　若一个 $n\times n$ 实矩阵的列的全部线性组合不能填满 $\mathbb{R}^n$，称它为**奇异矩阵**（singular matrix）；能够填满时称为非奇异。向量组 $a_1,\dots,a_n$ 称为**线性相关**（linearly dependent），如果存在不全为零的系数 $c_1,\dots,c_n$，使 $\sum_j c_j a_j=0$；否则称为线性无关。

第二个定义的否定形式应当仔细读：线性无关要求得到零向量的组合只有全部系数为零的一种。任意向量组都可以用全零系数组合成零向量，故这个平凡组合不能证明相关。对 B，取系数

$(-2,1)$ 得 $-2(1,2)+(2,4)=0$，系数不全为零，的确证明两列相关。对主矩阵 A，通式给出 $Ax=0$ 只有 $x=0$，所以两列无关。

若一个向量组包含零向量，令这一列的系数为 1，其余列系数为 0，便得到非平凡的零组合，所以它相关；单个非零向量则无关，因为 $cv=0$ 时比较一个非零分量即可推出 $c=0$。不能单凭列数多或少作判断，还要检查这些向量之间的关系。

例 1.9　同一奇异矩阵的两个右端　分别求 $B(x,y)=(3,6)$ 和 $B(x,y)=(3,7)$ 的全部解。

解　第一问的两式为 $x+2y=3$ 和 $2x+4y=6$，后一式是前一式的两倍。令 $y=t$，则全部解为 $(x,y)=(3-2t,t)$，$t\in\mathbb{R}$。回代得第一分量 $3-2t+2t=3$，第二分量为 6；任意解都能以自己的第二分量作参数，故没有遗漏。

第二问要求同一个数 $x+2y$ 同时等于 3 与 $7/2$，矛盾。因此无解。在行图像中，前者的两条直线重合，后者平行不重合；在列图像中，前者的目标在可达直线上，后者的目标不在。这两幅图表达同一个代数事实。 ♦

想一想　若 $Au=0$，其中 $u\neq 0$，而 $Ax_0=b$ 已有一个解，能立刻得到什么？能据此断言每个右端都有解吗？

答　对所有 $t\in\mathbb{R}$，$A(x_0+tu)=Ax_0+tAu=b$，且不同 t 给出不同解，所以这个右端有无穷多解。但存在非零的零输入方向，并不能保证任意右端都可达。刚才的 B 有 $u=(-2,1)$，却无法达到 $(3,7)$。需要先确认一个特定右端有解，才能沿着这样的方向产生更多解。

在方阵情形，「列能填满整个空间」「每个右端恰有一个解」「存在逆矩阵」会在后续课程中证明等价。这里对主矩阵 A 已直接给出系数公式，对 B 已直接求出所有可达右端，无需提前引用一般等价定理。奇异性描述固定矩阵的整体性质，某个方程组是否相容还要看它的右端。

1.5.1　原习题中的三个共面向量

例 1.10　原习题 1.1：不全为零的组合　设 $w_1=(1,2,3)$，$w_2=(4,5,6)$，$w_3=(7,8,9)$。找一个非平凡零组合，写出包含三列的平面，并判断三列的全部组合是否恰好填满该平面。

解　先比较相邻向量：$w_2-w_1=(3,3,3)$，$w_3-w_2=(3,3,3)$。相减得

$$w_1-2w_2+w_3=(0,0,0).$$

系数 $(1,-2,1)$ 不全为零，所以三列线性相关。逐列核对 $1-2\cdot 2+3=0$、$4-2\cdot 5+6=0$、$7-2\cdot 8+9=0$，可见每列均在过原点的平面

$$P=\{b\in\mathbb{R}^3 : b_1-2b_2+b_3=0\}$$

上。线性组合仍满足这个等式，因为相应三项的加权和仍为 0，所以可达输出都在 P 中。

要证明恰好填满，必须补反向包含。任取 $b\in P$，令

$$s=\frac{4b_2-5b_1}{3},\qquad t=\frac{2b_1-b_2}{3}.$$

计算得 $s+4t=b_1$、$2s+5t=b_2$，第三分量为 $3s+6t=-b_1+2b_2=b_3$。故 $sw_1+tw_2=b$，第三列的系数取 0 就够了。这样证明可达集合恰为平面 P。例如 $(1,0,0)$ 不在 P 中，所以三列填不满 $\mathbb{R}^3$，以它们为列的方阵 W 是奇异的。 ♦

这里先由零组合发现冗余，再用平面方程限制输出，最后用系数公式确认每个平面内目标均能达到。只证明「所有输出在某平面内」，尚不足以说「输出恰为整个平面」：三个零向量的输出只有原点，它们也都在这个平面里。写出任意目标的组合，是确认没有遗漏可达点的方法。

零组合必须检查系数　「这组向量能组合成零，所以相关」缺少**系数不全为零**这个条件。向量 $(1,0)$、$(0,1)$ 用系数 $(0,0)$ 也能得到零，但若 $c(1,0)+d(0,1)=0$，逐分量只能有 $c=d=0$，所以它们无关。非平凡零组合指的是系数不全为零，不是组合结果非零。

1.5.2　三维行图像与矩形矩阵

■ **例 1.11　三个平面可以交成直线**　求方程组 $x+y+z=3$、$x-y+z=1$、$2x+2z=4$ 的全部解，说明三行与三列各提供了什么信息。

解　前两式相减得 $2y=2$，故 $y=1$；代回得 $x+z=2$。第三式是前两式相加，没有增加条件。令 $x=t$，全部解为 $(t,1,2-t)$，$t\in\mathbb{R}$。回代三行分别得到 $3,1,4$，反向推导已表明每个解都必须有此形式。

行图像中，前两个平面相交成直线，而第三平面包含这条直线。列图像中，矩阵的第一列与第三列相等，都是 $(1,1,2)$；第二列为 $(1,-1,0)$。因此输出只取决于 $x+z$ 和 y，无法区分 x,z 分别取了多少。对右端 $(3,1,4)$，需要 $x+z=2$ 与 $y=1$，故第一列与第三列的总系数可任意分配。两种图像给出的解集一致。 ♦

若三行组成单位矩阵，则行方程为 $x=b_1, y=b_2, z=b_3$。每个平面固定一个坐标，与相应坐标平面平行或重合；三者的唯一公共点是 (b_1,b_2,b_3)。列图像中，三个坐标方向的系数分别直接取 b_1,b_2,b_3。例如右端 $(1,2,3)$ 给出唯一解 $(1,2,3)$。因此「三个方程」本身既不保证唯一，也不保证不唯一，关键在这些方程提供的限制是否彼此重复，以及右端是否相容。

■ **例 1.12　三个方程，两个未知数**　设 $C=\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}$，求 $Cx=(2,-1,1)$ 的全部解，并描述哪些三维右端可以被 C 达到。

解　把输入写为 (s,t)，输出为 $(s,t,s+t)$。前两个分量强制 $s=2, t=-1$，第三分量检查 $s+t=1$，确实成立，故这个右端有唯一解 $(2,-1)$。对一般右端 b，前两行仍强制 $s=b_1, t=b_2$，所以第三行可满足当且仅当 $b_3=b_1+b_2$。当条件成立时组合系数唯一；不成立时无解。可达集合为过原点的平面 $b_3=b_1+b_2$，没有右端有无穷多解。 ♦

这里 C 不是方阵，我们只说它的列不能填满 $\mathbb{R}^3$，不用本讲仅对方阵约定的「奇异」一词。两

列在三维中无关，却也填不满三维空间。这说明「列无关」与「所有右端可达」对一般矩形矩阵是不同的要求，后面将分别用主元列与主元行来判断。

面对一个方程组的三个检查　先写清矩阵的行数、列数和右端的维数，再问目标是否在列的全部组合中；若有解，继续问组合系数是否唯一。行图像帮助识别相容性，列图像帮助说明可达范围，回代原方程检查每一项可防止计算遗漏。

本讲小结

- **同一个等式的三种读法**：逐行看 $\sum_j a_{ij}x_j = b_i$，逐列看 $\sum_j x_j a_j = b$，整体记为 $Ax = b$。
- **线性组合**：系数可以是任意实数。全部组合必含零向量；一个非零向量的倍数组成过原点的直线。
- **尺寸规则**：A 为 $m \times n$ 时，输入 $x \in \mathbb{R}^n$，输出 $Ax \in \mathbb{R}^m$；B 为 $n \times p$ 时，AB 为 $m \times p$。
- **线性性**：$A(cu + dv) = cAu + dAv$。若 $Au = Av = b$ 且 $u \neq v$，则 $u + t(v - u)$ 为无穷多个不同的解。
- **可达性**：$Ax = b$ 有解当且仅当 b 是各列的组合；每个右端有解当且仅当全部列组合填满输出空间。
- **相关性**：存在不全为零的系数使 $\sum_j c_j a_j = 0$ 才称列相关。若已有一个解，加上这样的零组合方向可得到更多解。
- **方阵奇异**：列组合填不满整个输出空间。奇异方阵仍可能对某些右端有解，不能把「奇异」直接读成「这个右端无解」。

下一讲把逐行相加减的操作整理成消元法：在不改变解集的前提下，按固定顺序消去未知数。每一步仍可用本讲的行与列两种图像核对。

练习

1.1 ★　解 $2x - y = 4$、$-x + 2y = 1$，先按行消去未知数，再写成两个列向量的组合并回代。

1.2 ★　令 $v = (1, -2, 0)$、$w = (0, 1, 3)$，求 $2v - w$ 与 $v + 2w$。判断 $(2, -5, -3)$ 和 $(1, 0, 0)$ 是否为 v, w 的线性组合；有则给系数，无则说明理由。

1.3 ★　按列与按行各算一次 $\begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -2 \end{bmatrix}(2, -1, 3)$。写出矩阵、输入及输出各自的尺寸。

1.4 ★　判断真假并说明理由：(a) 三个方程三个未知数总有唯一解；(b) 两列各自非零就线性无关；(c) 任意向量组都能用某组系数组合成零；(d) 若一个实线性方程组有两个不同的解，它必有无穷多个解。

1.5 ★★　对 $B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$，求方程 $Bx = (a, b)$ 有解的充要条件，并在条件满足时写出全部解。分别应用于 $(a, b) = (0, 0)$ 与 $(1, 1)$。

1.6 ★★ 对例 1.10 的矩阵 W，写出 $Wx=0$ 的全部解。求 $Wx=(4,5,6)$ 的全部解，解释两组答案之间的关系。

1.7 ★★ 证明：若 $Ax_0=b$，则 $Ax=b$ 的全部解恰为 $x=x_0+u$，其中 u 遍历 $Au=0$ 的全部解。两个包含方向都要证明。

1.8 ★ A 为 3×2，B 为 2×4，D 为 4×1。判断 AB、BA、BD、AD 是否有定义；有定义的写出尺寸。每次都说明内部尺寸与外部尺寸。

1.9 ★★ 对矩形矩阵 $C=\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}$，分别判断右端 $(1,2,3)$ 与 $(1,2,4)$ 是否可达。证明两列线性无关，并解释这为什么没有推出所有三维右端可达。

1.10 ★★★ 若已知 $Ax=b$、$Ay=c$，给出 $Az=2b-c$ 的一个解。它一定是唯一解吗？用 $A=\begin{bmatrix}1&2\\2&4\end{bmatrix}$、$x=(1,0)$、$y=(0,1)$ 具体检验，再给出同一右端的另一个解。

练习答案与提示

1.1 第一式乘 2 再加第二式，得 $3x=9$，故 $x=3$，$y=2x-4=2$。回代得 $6-2=4$、$-3+4=1$。列形式为 $3(2,-1)+2(-1,2)=(4,1)$。逐行算出的交点 $(3,2)$ 是输入，组合得到的 $(4,1)$ 是右端，二者角色不同。

1.2 按分量计算，$2v-w=(2,-5,-3)$，$v+2w=(1,0,6)$，所以第一个目标的系数为 $(2,-1)$。若 $cv+dw=(1,0,0)$，第一分量强制 $c=1$，第三分量强制 $3d=0$，故 $d=0$；但第二分量便为 $-2\neq 0$，矛盾。因此第二个目标不是线性组合。取两个分量解出候选后，必须检查剩下的分量。

1.3 按列算 $2(1,0)-(-1,3)+3(2,-2)=(9,-9)$；这里第二个系数是 -1，所以第二列贡献为 $(1,-3)$。按行算得第一分量 $2+1+6=9$，第二分量 $0-3-6=-9$。矩阵为 2×3，输入为 3×1，输出为 2×1，尺寸满足乘法要求。

1.4 (a) 假。例如三行都为 $x+y+z=0$，有无穷多解；若第二行改为 $x+y+z=1$，又变为无解。(b) 假，$(1,2)$ 与 $(2,4)$ 均非零，但 $-2(1,2)+(2,4)=0$ 是非平凡零组合。(c) 真，所有系数取零即可；它并不等于相关的定义，因为相关还要求系数不全为零。(d) 真，由推论 1.6，连接两个解的整条参数直线都由解组成，且不同参数给出不同点。

1.5 必须有 $b=2a$，因为输出的第二分量始终为第一分量的两倍。反之，若 $b=2a$，取 $x_2=t$、$x_1=a-2t$，则 $Bx=(a,2a)=(a,b)$，故条件充分。这也是所有解，因为任何解都必须满足第一行 $x_1+2x_2=a$。对 $(0,0)$，全部解为 $(-2t,t)$；对 $(1,1)$，$1\neq 2\cdot 1$，所以无解。右端为零时总有零解，但本例还有无穷多个非零解。

1.6 写出三行 $x_1+4x_2+7x_3=0$、$2x_1+5x_2+8x_3=0$、$3x_1+6x_2+9x_3=0$。第二行减第一行的 2 倍，得 $-3x_2-6x_3=0$，所以 $x_2=-2x_3$；再由第一行得 $x_1=x_3$。第三行回代为 $3x_3-12x_3+9x_3=0$，无新限制。因此零右端的全部解为 $t(1,-2,1)$。

对 $(4,5,6)=w_2$，有特解 $(0,1,0)$。若 $Wx=w_2$，逐行减去这个特解，得到 $W(x-(0,1,0))=0$，所以全部解为 $(t,1-2t,t)$。回代时三列组合为 $w_2+t(w_1-2w_2+w_3)=w_2$。第一组是所有不改变输出的方向，第二组是在一个特解上加这些方向。

1.7 先证构造出的点都是解：若 $Au=0$，线性性给出 $A(x_0+u)=Ax_0+Au=b$。再证每个解都在其中：若 $Ax=b$，令 $u=x-x_0$，则 $Au=Ax-Ax_0=b-b=0$，且 $x=x_0+u$。两个方向都完成，

故所给集合恰为全部解。条件 $Ax_0 = b$ 很重要；没有特解时，零方程的解集仍存在，却不能凭它制造非零右端的解。

1.8 AB 有定义，内部尺寸为 2 与 2，结果外部尺寸为 3×4。BA 的内部尺寸为 4 与 3，不相等，故无定义。BD 的内部尺寸为 4 与 4，结果为 2×1；AD 的内部尺寸为 2 与 4，故无定义。只看最终想要几行几列，不能绕过内部尺寸必须相等的要求。

1.9 $C(s,t) = (s, t, s+t)$。第一个右端由 $(s,t) = (1,2)$ 达到，且前两分量使这个输入唯一。第二个右端也强制 $s = 1, t = 2$，但第三分量变为 $3 \neq 4$，所以不可达。若 $c(1,0,1) + d(0,1,1) = 0$，第一分量给 $c = 0$、第二分量给 $d = 0$，故两列无关。它们仍只组合出平面 $b_3 = b_1 + b_2$，所以无关只保证可达目标的系数唯一，不保证全部三维目标都可达。

1.10 由线性性，取 $z = 2x - y$ 就有 $Az = 2Ax - Ay = 2b - c$。这个构造只保证存在一个解，唯一性还要排除非零的零组合方向。给定矩阵下，$b = (1,2)$、$c = (2,4)$，故 $2b - c = 0$；所造的 $z = (2,-1)$ 确有 $Az = 0$。另一个解为 $(0,0)$，不同于 z。事实上所有解为 $(-2t, t)$，z 对应 $t = -1$。因此原断言不能加强成「总是唯一解」。

第 2 讲

用矩阵做消元

Elimination with Matrices

对应材料：MIT 18.06SC Session 1.2 「Elimination with matrices」课时摘要、习题 2.1–2.2 及完整解答；补充采用 Strang《ZoomNotes for Linear Algebra》§2.2–2.4 中的回代、三元消元和换行例题。一般置换性质留到第 5 讲，原摘要中的行列式预告留到第 17 讲，分解 $A = LU$ 留到第 4 讲。

本讲目标

(1) 能说明行交换、非零倍数与行替换为什么保持方程组的全部解，并同步处理右端。

(2) 会在增广矩阵中找主元、算乘数、向下消元，再对非零对角元的上三角系统从下往上回代。

(3) 能区分零主元位置需要换行、当前列确实没有候选，以及无解与有无穷多解。

(4) 会写一步消元的矩阵，按正确次序合并多步操作，并逐步撤销这些操作。

(5) 能解释置换矩阵左乘换行、右乘换列的区别，完成原课程的 Pascal 矩阵综合题。

2.1 先问清楚：什么操作不会丢掉解

第 1 讲把同一个方程组写成行方程、列组合和 $Ax = b$。列图像回答目标是否能够达到，实际求组合系数时，我们仍需要一套按顺序完成的计算。面对三个方程，逐个猜交点已经不方便；消元法的目的，是把原方程变成容易依次解出的方程，同时保留原来的全部解。先看原课程习题 2.1 的系统：

$$2x + 3y = 5, \qquad 6x + 15y = 12. \tag{2.1}$$

第二式减第一式的 3 倍，x 的系数变为 $6 - 3 \cdot 2 = 0$，于是第二式变成 $6y = -3$。注意右端也做了 $12 - 3 \cdot 5 = -3$；若只在左边消去 x，却仍保留右端 12，所得系统就已经改变。这个看似简单的细节，是手算和编程中都必须遵守的规则。

定义 2.1（等价方程组与行操作） 两个未知数相同的方程组，若它们的全部解组成同一个集合，称为**等价方程组**（equivalent systems）。对方程组的整行，包括右端，可作三种**初等行操作**（elementary row operations）：交换两行；将一行乘以非零实数；把另一行的某个实数倍加到一行上。

把第 i 行记为 R_i，第三种操作写成 $R_i \leftarrow R_i + cR_j$，其中 $i \neq j$。箭头右边的两行都是操作之前的行；操作完成后只替换第 i 行，第 j 行保持原样。系数 c 可以为零，此时什么也没有改变。

第二种操作不允许乘零，因为乘零无法恢复原来那一条限制。

命题 2.2（可撤销的行操作保持全部解）　对一个实线性方程组实施上述任意一次行操作，得到的新方程组与原方程组等价。有限多次这样的操作仍保持全部解。

证明　交换两行只改变方程的排列，每个同时满足各式的输入仍满足各式，再交换一次即可恢复。将一行乘非零实数 c，原等式成立时乘后的等式成立；反过来，用 $1/c$ 乘新等式便恢复原式，故两个方向成立。对行替换，设原来的第 i, j 式是 $a_i x = b_i$、$a_j x = b_j$，这里 a_i, a_j 是行向量。原系统的解满足 $(a_i + c a_j)x = b_i + c b_j$，所以必为新系统的解。反过来，新系统仍含未改变的第 j 式；用新的第 i 式减去它的 c 倍，得 $a_i x = b_i$，其余各行均未改动，故新系统的解也为原系统的解。每一步的解集相等，连续有限步后，最初与最后的解集便相等。这同时证明了「不丢解」与「不增加额外的解」，不只是在一个已知解上作了检验。 ■

■ **例 2.1　原习题 2.1：第一次完整消元**　对 (2.1) 写出矩阵形式，求消元乘数、两个主元和全部解，并回代原系统。

解　矩阵形式为 $\begin{bmatrix} 2 & 3 \\ 6 & 15 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 12 \end{bmatrix}$。第一行的首个非零系数为 2；要消去下方的 6，乘数是 $6/2 = 3$。作 $R_2 \leftarrow R_2 - 3R_1$ 后，系统变成 $2x + 3y = 5$、$6y = -3$，两个主元是 $2, 6$。从第二行得 $y = -1/2$；代入第一行，$2x = 5 + 3/2 = 13/2$，所以 $x = 13/4$。两个求解步骤都由非零系数唯一决定，故没有其他解。原第一式为 $2 \cdot 13/4 + 3(-1/2) = 5$，原第二式为 $6 \cdot 13/4 + 15(-1/2) = 78/4 - 30/4 = 12$，两式都成立。 ♦

只知道新旧第二条直线经过同一个交点，并不足以证明等价；真正保留的还有第一条方程。图 2.1 将三条直线放在一张图里：蓝线不变，橙线被青线替换，蓝线与后两条线的公共点相同。

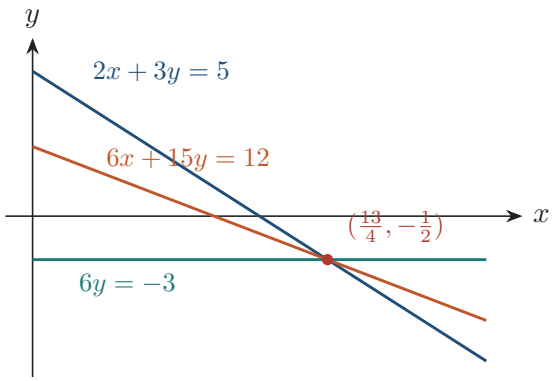


图 2.1　保留第一行，把第二行减去第一行的三倍；橙线与青线本身不同，所得两个系统的公共解相同。

想一想　既然第一行能加到第二行，能不能先把第一行加到第二行，再把第二行加到第一行，最后称「两行同时相加」？

答　不能把依次操作与同时读取旧行混为一谈。依次操作的第二步会用到已经更新的第二行；若把两行都同时替换成旧两行之和，两条新方程完全相同，原来不同的限制便可能丢失。例如 $x=0, y=0$ 的唯一解是原点，同时替换后只有 $x+y=0$，得到整条直线。行操作必须写清楚哪一行改变、哪一行保持，以及每一步读取的是哪个版本。

把一行乘零会丢失信息　方程 $x=2$ 的唯一解为 2，乘零后得到 $0=0$，任意实数都满足。新系统含有原解，却增加了额外解，不能称为等价。证明行操作可用时，必须说明操作可以撤销；不能只说「原来的解代进去还成立」。

2.2　从增广矩阵到上三角系统

两元系统只需消去一个系数。三元系统需要分阶段：第一阶段消去第一列主元下方的项，第二阶段消去第二列主元下方的项，最后一行只剩一个未知数。每个阶段都要让已经得到的零保持为零，否则前面的工作会被破坏。

定义 2.3（增广矩阵、主元与消元乘数）　把右端 b 接在 A 右边，得到的 $[A \mid b]$ 称为**增广矩阵**（augmented matrix）。消元时用来消去同列下方系数的非零数称为**主元**（pivot）。若当前主元为 p，目标系数为 a，则**消元乘数**（elimination multiplier）为 $\ell = a/p$，操作为「目标行减去主元行的 ℓ 倍」。

「当前」十分重要。主元通常已被前面的行操作改变，不能总到原矩阵里去取；乘数的分子也必须取当前目标行的系数。增广矩阵中的竖线只用来区分未知数系数与右端，整行运算跨过竖线进行，但右端列不当作未知数列去选主元。

■ **例 2.2　原课程的三元主例**　对 $A=\begin{bmatrix}1&2&1\\3&8&1\\0&4&1\end{bmatrix}$、$b=(2,12,2)$，用前向消元得到上三角系统，再求全部解。

解　第一主元为 1。为了消去第二行第一列的 3，取 $\ell_{21}=3/1=3$，作 $R_2 \leftarrow R_2-3R_1$。第三行第一列已经为零，相应乘数 $\ell_{31}=0$，无需改变第三行。得到

$$\left[\begin{array}{ccc|c}1&2&1&2\\3&8&1&12\\0&4&1&2\end{array}\right] \longrightarrow \left[\begin{array}{ccc|c}1&2&1&2\\0&2&-2&6\\0&4&1&2\end{array}\right].$$

第二行的三个系数依次是 $3-3=0$、$8-3\cdot 2=2$、$1-3\cdot 1=-2$，右端为 $12-3\cdot 2=6$。尤其第二主

元是现在的 2，不是原矩阵第二行第二列的 8。下一步的乘数为 $\ell_{32}=4/2=2$，作 $R_3\leftarrow R_3-2R_2$：

$$\left[\begin{array}{ccc|c}1&2&1&2\\0&2&-2&6\\0&4&1&2\end{array}\right]\longrightarrow\left[\begin{array}{ccc|c}1&2&1&2\\0&2&-2&6\\0&0&5&-10\end{array}\right]=[U\mid c].$$

第三行的末系数是 $1-2(-2)=5$，右端是 $2-2\cdot 6=-10$。三个主元依次为 $1,2,5$，都非零。最后三条方程为 $x+2y+z=2$、$2y-2z=6$、$5z=-10$。从最后一行开始，$z=-2$；第二行给 $2y+4=6$，故 $y=1$；第一行给 $x+2-2=2$，故 $x=2$。原系统回代为 $2+2-2=2$、$6+8-2=12$、$4-2=2$。每步行变换等价，且回代的三个未知数各被唯一决定，故 $(2,1,-2)$ 是全部解。 ♦

第一阶段之后，第一列的主元下方全为零；第二阶段使用的第二行在第一列也是零，所以它的倍数不会重新制造第一列的非零数。以此类推，在第 k 阶段，只用从第 k 列起才可能非零的主元行处理更下面的行，前 $k-1$ 列中已经得到的零便能保留下来。

定义 2.4（上三角矩阵与回代）　方阵 $U=(u_{ij})$ 若在主对角线下方满足 $u_{ij}=0$（$i>j$），称为**上三角矩阵**（upper triangular matrix）。在系统 $Ux=c$ 中，从最后一行开始，逐行用已知的后面分量求前面分量，称为**回代**（back substitution）。

上三角只要求对角线下方为零，不要求对角线或上方非零。单位矩阵是上三角矩阵，零矩阵也是；但零矩阵的方程组并不能总用除法唯一求解。真正允许回代中每一步除法的条件，是全部对角元非零。刚才得到的三个对角元正是三个主元。

命题 2.5（非零对角元的回代给出唯一解）　设 $n\ge 1$，U 为 $n\times n$ 上三角实矩阵，且 $u_{ii}\neq 0$ 对每个 i 成立。对任意 $c\in\mathbb{R}^n$，$Ux=c$ 有唯一解，它由

$$x_i=\frac{c_i-\sum_{j=i+1}^{n}u_{ij}x_j}{u_{ii}},\qquad i=n,n-1,\dots,1$$

依次确定。当 $i=n$ 时，分子中的和为空，按 0 计算。

直观理解　最后一条方程没有前面的未知数，先固定 x_n；倒数第二条只需再决定 x_{n-1}。每向上一行，右边用到的分量都已经求出。这个次序把一个多变量问题化为依次解单个非零系数的一元方程。

证明　第 n 行是 $u_{nn}x_n=c_n$，非零系数使它恰有一个解 $x_n=c_n/u_{nn}$。假设 $x_{i+1},\dots,x_n$ 已被唯一确定。上三角性质使第 i 行只有 $u_{ii}x_i+\sum_{j=i+1}^{n}u_{ij}x_j=c_i$，其中求和部分已经是已知数，$u_{ii}\neq 0$ 使 x_i 恰由公式确定。从 $i=n$ 倒推到 1，确实构造出全部分量；构造时每一行都被满足，所以存在解。另一方面，任意解都必须先取同一个 x_n，再取同一个 x_{n-1}，依次直到 x_1，所以解唯一。存在性与唯一性都由相同的递推次序得到。 ■

■ 例 2.3 **ZoomNotes 的第二组数据**　求 $\begin{bmatrix}2&3&4\\4&11&14\\2&8&17\end{bmatrix}x=(19,55,50)$。记录每一阶段实际使用的主元和乘数。

解　第一主元为 2，乘数为 $\ell_{21}=4/2=2$、$\ell_{31}=2/2=1$。两步都读取未改变的第一行，得到

$$\left[\begin{array}{ccc|c}2&3&4&19\\0&5&6&17\\0&5&13&31\end{array}\right].$$

第二主元为 5，乘数 $\ell_{32}=5/5=1$；第三行减第二行后，末系数为 $13-6=7$，右端为 $31-17=14$。上三角系统为

$$\begin{bmatrix}2&3&4\\0&5&6\\0&0&7\end{bmatrix}x=(19,17,14).$$

回代得 $x_3=2$，$x_2=(17-6\cdot2)/5=1$，$x_1=(19-3\cdot1-4\cdot2)/2=4$。原方程三行分别给 $8+3+8=19$、$16+11+28=55$、$8+8+34=50$。三个主元 $2,5,7$ 都非零，故唯一解为 $(4,1,2)$。♦

在同一个系数矩阵下换一个右端时，所有消元乘数保持相同，因为它们只从系数部分产生。改变的是消元后右端和最终解。这一点意味着可以先记录对 A 做过的操作，再把同样操作施加到许多不同的 b；本讲后半段将把这份操作记录写成矩阵。

想一想　上三角系统能不能从第一行开始解？第一行也是一个方程，为什么要绕到最后一行？
答　第一行一般同时含有多个尚未知道的分量，例如主例为 $x+2y+z=2$，单凭它无法唯一求 x。最后一行只有一个未知分量，先求它才能让上一行变成一元方程。如果 U 是对角矩阵，各行互不依赖，顺序确实无关；回代的规定次序适用于一般上三角情形。

2.3　零出现在主元位置时

上节每一次除法都有非零分母。现在设主元位置恰好是零，不能写出「用下方系数除以它」的乘数。我们需要先查明：是行的排列不合适，还是这一列在尚未处理的部分中已经找不到非零数？这两个情形要求不同的处理。

■ 例 2.4 **第一步就要换行**　求 $\begin{bmatrix}0&2\\3&4\end{bmatrix}(x,y)=(4,11)$。左上角为零，能否据此说系统没有唯一解？

解　第一条方程是 $2y=4$，第二条是 $3x+4y=11$，它们的限制没有重复。把两行连同右端交换，得到 $\begin{bmatrix}3&4\\0&2\end{bmatrix}(x,y)=(11,4)$。新主元依次为 $3,2$，回代得 $y=2$、$x=(11-8)/3=1$。原两行给 $2\cdot2=4$、$3\cdot1+4\cdot2=11$，验证成立，而且回代保证唯一。原来只有主元的位置为零；下方同

列的 3 可以被换上来。矩阵中出现某个零，乃至对角线上出现零，都不能单独判定系统没有唯一解。必须看经过允许换行后，能否得到一整列的消元主元。 ♦

■ **例 2.5　第二阶段才出现零主元**　对 $J=\begin{bmatrix}1&1&1\\2&2&3\\1&3&4\end{bmatrix}$、$b=(6,15,19)$ 消元。第一阶段后若第二主元位置为零，应该怎样继续？

解　第一主元为 1，作 $R_2\leftarrow R_2-2R_1$、$R_3\leftarrow R_3-R_1$，得到

$$\left[\begin{array}{ccc|c}1&1&1&6\\0&0&1&3\\0&2&3&13\end{array}\right].$$

第二主元位置为零，下方同列的 2 非零。交换现在的第二、三行，得到

$$\left[\begin{array}{ccc|c}1&1&1&6\\0&2&3&13\\0&0&1&3\end{array}\right].$$

已经是上三角形，主元为 $1,2,1$。回代给 $z=3$、$2y=13-9=4$，故 $y=2$、$x=6-2-3=1$。原三行分别为 $1+2+3=6$、$2+4+9=15$、$1+6+12=19$。交换发生在尚未处理的两行之间，它们的第一列都为零，所以不会破坏第一阶段的结果。不能回头与已完成的第一行任意交换后仍沿用原有主元记录；若改变已完成的顺序，就需要相应重新追踪整个过程。 ♦

2.3.1　一列真的找不到非零候选

如果当前第 k 个主元位置以及它下面的同列项全部为零，换行无法制造一个非零数。此时「每列都有一个非零对角主元」的方阵回代路线停止，但并不意味着所有右端都无解。第 1 讲已经说明，列的组合可能只填满一条线；线上的右端可达，线外的右端不可达。

■ **例 2.6　同一个失败位置，两种解集**　系数矩阵为 $B=\begin{bmatrix}1&2\\2&4\end{bmatrix}$。分别把右端取为 $(3,6)$、$(3,7)$，消元后各说明什么？

解　第一主元为 1，作 $R_2\leftarrow R_2-2R_1$。两个增广矩阵分别变为

$$\left[\begin{array}{cc|c}1&2&3\\0&0&0\end{array}\right],\qquad\left[\begin{array}{cc|c}1&2&3\\0&0&1\end{array}\right].$$

第一种的第二行是恒等式 $0=0$，不给新的限制。只剩 $x+2y=3$，令 $y=t$，全部解为 $(3-2t,t)$。任意此形输入都使输出为 $(3,6)$，反过来任何解都由第一行具有此形，因此没有遗漏。第二种的第二行要求 $0=1$，任意输入都无法满足，故无解。系数部分在两种情形中完全相同：同一个零主元能配上恒等式，也能配上矛盾式，必须保留右端才能作出判断。 ♦

零主元不是无解的同义词　左上角为零的例 2.4 有唯一解；消元得到全零系数行的例 2.6，则依右端不同而有无穷多解或无解。检查顺序是先看当前同列下方能否换入非零项，再看零系数行的右端。任何一步都不允许除以零。

还可能在第一列就没有候选。例如 $\begin{bmatrix}0 & 1\\ 0 & 2\end{bmatrix}(x,y)=(1,2)$ 的第一列全零，但方程依然要求 $y=1$，x 可任取，全部解为 $(t,1)$。继续研究这种系统时，要跳过全零的候选列，再寻找后面的主元；第 7、8 讲会用主元变量与自由变量作系统整理。本讲先把方阵是否能沿对角线得到全部主元的问题讲清楚。

命题 2.6（方阵缺少对角主元时不能有唯一解）　对 $n\times n$ 系统按列消元，前 $k-1$ 个非零主元已经得到。若当前第 k 列从第 k 行起全部为零，则任意右端的原系统都不可能恰有一个解。

证明　把已经得到的系数矩阵记为 T。前 $k-1$ 列的对角元非零，且各自下方为零。构造一个输入 u：取 $u_k=1$，$u_{k+1}=\cdots=u_n=0$。从第 k 行往下，前 $k-1$ 列已为零，第 k 列由假设也为零，后面的分量又被取零，因此这些行的 Tu 都等于零。剩下前 $k-1$ 行，按 $i=k-1,\dots,1$ 的次序选取 u_i。第 i 行对 u_i 的系数是已经找到的非零主元，后面各分量已知，所以能恰好取一个 u_i 使该行等于零。这便得到 $Tu=0$ 且 $u\neq 0$。若 $k=1$，这段回代为空，直接取 $u=(1,0,\dots,0)$ 即可。行操作保持齐次系统的全部解，所以原矩阵也满足 $Au=0$。若给定右端没有解，当然谈不上唯一；若已有一个解 x_0，线性性给出 $A(x_0+tu)=b$ 对所有实数 t 成立，且 $u_k=1$ 保证不同 t 得到不同输入。于是有解时必有无穷多个解。 ■

这个证明不使用行列式，也不使用尚未介绍的秩。它直接指出缺少主元留下了一个不改变输出的非零方向。与之相对，若通过换行与消元得到全部非零主元，命题 2.5 则对每个右端给出唯一解，且等价行操作把这一结论带回原系统。

2.4　用一个矩阵记录一步消元

主例的第一步是「第二行减去第一行的 3 倍」。能不能不再把这句话逐次写在矩阵旁边，而用一个固定矩阵把这步操作记录下来？第 1 讲已定义矩阵乘积的各列：AB 的第 j 列是 A 乘 B 的第 j 列。由同一个定义，也能从行的角度理解左乘。

命题 2.7（左乘矩阵是在组合右边矩阵的行）　设 F 为 $m\times m$ 矩阵，A 为 $m\times n$ 矩阵。乘积 FA 的第 i 行为 A 的各行的线性组合，组合系数依次是 F 的第 i 行：

$$\operatorname{row}_i(FA)=\sum_{k=1}^{m} f_{ik}\operatorname{row}_k(A).$$

证明　按列定义，乘积的对应元素为

$$(FA)_{ij}=\sum_{k=1}^{m} f_{ik}a_{kj}.$$

把 $j=1,\dots,n$ 的分量排成一行，就是各行按 f_{ik} 加权求和。这里出现的行组合与上讲的列组合同属一个乘法定义，无需再约定另一种运算。 ■

定义 2.8（消元矩阵）　对 $i\neq j$，把单位矩阵 I_m 的第 (i,j) 元从 0 改为 $-\ell$，所得矩阵记为 $E_{ij}(\ell)$，称为这一步的**消元矩阵**（elimination matrix）。左乘它实现 $R_i\leftarrow R_i-\ell R_j$，其余各行不变。下标记录被改变的行与被使用的行，括号中的 ℓ 是被减去的乘数。

例如主例的第一步对应

$$E_{21}=\begin{bmatrix}1&0&0\\-3&1&0\\0&0&1\end{bmatrix},\qquad E_{21}A=\begin{bmatrix}1&2&1\\0&2&-2\\0&4&1\end{bmatrix},\qquad E_{21}b=(2,6,2).$$

第二行的组合系数为 $(-3,1,0)$，所以它取的是 $-3R_1+R_2$。容易写反的符号现在可以用行组合检查：消去正的 3 时，需要矩阵中的 -3。若当前目标本来为零，$\ell=0$，所得矩阵就是 I；主例不改变第三行第一列的那一步可记作 $E_{31}=I$。

■ **例 2.7　从操作写矩阵，再把它撤销**　写出「第三行减去第一行的 4 倍」的 3×3 矩阵，并作用到 $X=\begin{bmatrix}1&2&3\\0&1&4\\4&5&6\end{bmatrix}$。如何恢复 X？

解　只有第 $(3,1)$ 元需要改为 -4，故

$$F=\begin{bmatrix}1&0&0\\0&1&0\\-4&0&1\end{bmatrix},\qquad FX=\begin{bmatrix}1&2&3\\0&1&4\\0&-3&-6\end{bmatrix}.$$

末行依次为 $4-4\cdot 1=0$、$5-4\cdot 2=-3$、$6-4\cdot 3=-6$。要恢复，给末行加回第一行的 4 倍，即左乘 $G=\begin{bmatrix}1&0&0\\0&1&0\\4&0&1\end{bmatrix}$。由于第一行一直不变，$G(FX)$ 恢复 X 的每一行；直接相乘也有 $GF=FG=I_3$。 ♦

命题 2.9（一步消元的反操作）　对 $i\neq j$ 和任意实数 ℓ，矩阵 $E_{ij}(\ell)$ 的双侧逆为 $E_{ij}(-\ell)$。因而

$$Ax=b\quad\Longleftrightarrow\quad E_{ij}(\ell)Ax=E_{ij}(\ell)b.$$

证明　先减去第 j 行的 ℓ 倍，再加回它的 ℓ 倍，两次操作都不改变第 j 行，故每一行恢复原状。将这两个操作施加到 I_m，得到两个顺序的乘积都等于 I_m，即 $E_{ij}(-\ell)E_{ij}(\ell)=E_{ij}(\ell)E_{ij}(-\ell)=I_m$。前向蕴含由左乘得到；反向用反操作矩阵左乘便恢复 $Ax=b$。这与命题 2.2 的逐行论证一致。 ■

这里已使用「逆」的具体含义：一个矩阵在左右两种相乘顺序下都把另一个矩阵恢复为单位矩阵。本讲只验证消元与交换这些特定矩阵的逆；一般逆矩阵的定义、唯一性和求法留到第 3 讲。不要据此推断任意方阵都有逆，例 2.6 的 B 就保留了非零的零输入方向。

2.4.1 先做的操作为什么写在右边

主例先做 E_{21}，再做 $E_{32}=\begin{bmatrix}1&0&0\\0&1&0\\0&-2&1\end{bmatrix}$，因此上三角结果为 $E_{32}(E_{21}A)$。括号告诉我们：先把 A 变为 $E_{21}A$，再对这个中间矩阵用 E_{32}。新操作总乘在现有结果的左边，不能把时间先后顺序照抄成由左到右的因子排列。

> **命题 2.10（复合操作与括号）**　设 F,G 为 $m\times m$ 矩阵，A 为 $m\times n$ 矩阵，则 $F(GA)=(FG)A$。特别地，对输入 $v\in\mathbb{R}^m$，$F(Gv)=(FG)v$。

证明　固定输出行 i 和列 j，左式的对应元素为

$$\begin{aligned}\big(F(GA)\big)_{ij}&=\sum_{k=1}^{m}f_{ik}\Big(\sum_{r=1}^{m}g_{kr}a_{rj}\Big)\\&=\sum_{r=1}^{m}\Big(\sum_{k=1}^{m}f_{ik}g_{kr}\Big)a_{rj}=\big((FG)A\big)_{ij}.\end{aligned}$$

第二步只重排有限个实数乘积，使用实数的结合律与分配律。所有元素相等，故矩阵相等。令 A 只有一列便得到向量情形。证明只改变括号，没有交换 F,G 的位置。 ■

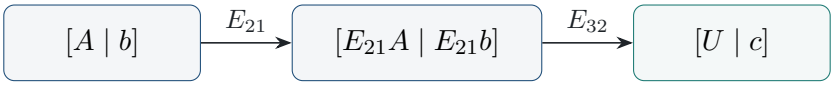
■ **例 2.8　两步合成时产生的新系数**　求主例的合成消元矩阵 $E=E_{32}E_{21}$，核对 $EA=U$、$Eb=c$，并比较 $E_{21}E_{32}$。

解　用 E_{32} 的末行 $(0,-2,1)$ 组合 E_{21} 的行，得到末行 $(6,-2,1)$，所以

$$E=\begin{bmatrix}1&0&0\\-3&1&0\\6&-2&1\end{bmatrix},\qquad EA=\begin{bmatrix}1&2&1\\0&2&-2\\0&0&5\end{bmatrix},\qquad Eb=(2,6,-10).$$

新出现的 6 表示：第三行减去的是*更新后*的第二行的两倍，而更新后的第二行为 R_2-3R_1，故最后的第三行是 $R_3-2(R_2-3R_1)=R_3-2R_2+6R_1$。这不是额外做了一步，是两步操作展开后的同一结果。若顺序反过来，$E_{21}E_{32}=\begin{bmatrix}1&0&0\\-3&1&0\\0&-2&1\end{bmatrix}$，它作用于 A 后的第三行为 $R_3-2R_2=(-6,-12,-1)$，第一列并不为零。因此交换因子会改变结果，矩阵乘法一般不满足交换律。 ♦

图 2.2 按时间从左向右展示消元；箭头上最后施加的矩阵，在合成结果中写到最左边。阅读一串乘积时，从靠近输入的右端因子开始，就能还原实际的操作次序。



$$[U \mid c] = E_{32}E_{21}[A \mid b]$$

图 2.2　对系数和右端施加相同操作；先施加的 E_{21} 在合成矩阵的右边。

想一想　如果知道 E_{21} 和 E_{32} 各自的反操作，把反操作按原来的顺序放在一起，能恢复 A 吗？
答　要先撤销最后做的 E_{32}，再撤销最先做的 E_{21}，所以从 U 恢复 A 的合成矩阵是 $E_{21}^{-1}E_{32}^{-1}$。括号核对为 $E_{21}^{-1}\big(E_{32}^{-1}(E_{32}(E_{21}A))\big) = A$。恢复的时间次序与原操作相反，合成矩阵中的因子次序也相应反转；这依靠相邻反操作抵消，不能随意把不相邻的两个因子跨过去抵消。

具体在主例中，E_{21}^{-1} 的第 $(2,1)$ 元为 3，E_{32}^{-1} 的第 $(3,2)$ 元为 2，乘积为

$$G = E_{21}^{-1}E_{32}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

左乘 U 时，先对第三行加回第二行的两倍，再对第二行加回第一行的三倍，确有 $GU = A$；右端也被恢复：$Gc = (2, 12, 2) = b$。验证 $GE = EG = I$ 说明这些具体操作确实可以完全撤销。第 4 讲会把这种恢复记录整理成一种重要分解，本讲暂不需要那种记号。

2.5　交换行与交换列：矩阵乘在哪一边

例 2.4 已经用到交换方程。它也能记录成左乘矩阵：把单位矩阵的相应两行交换，就得到一个执行相同行交换的矩阵。与行替换一样，这个记录可以同时作用到系数和右端，保证方程的排列一致。

定义 2.11（置换矩阵）　把单位矩阵的行按某个次序重新排列，所得矩阵称为**置换矩阵**（permutation matrix）。每行、每列各恰有一个 1，其他元素为 0。仅交换第 i, j 两行的置换矩阵记为 P_{ij}。

对三维输入，交换第一、二分量的矩阵为 $P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$。它的第一行选取输入的第二分量，第二行选取第一分量，第三行保留第三分量。单位矩阵本身也是置换矩阵，表示维持原次序；「置换」不要求一定改变某个分量。

命题 2.12（左乘换行，右乘换列）　设 A 为 $m \times n$ 矩阵。P_{ij} 为 m 阶交换矩阵时，$P_{ij}A$ 交换 A 的第 i, j 行；Q_{rs} 为 n 阶交换矩阵时，AQ_{rs} 交换 A 的第 r, s 列。每个交换矩阵连续作用两次为单位矩阵，故它本身就是自己的逆。

证明　左乘的第 i 行由交换矩阵的第 i 行决定，这一行只在第 j 个位置为 1，故输出是 A 的第 j 行；同理第 j 行来自第 i 行，其他行保持。右乘时按列读，Q_{rs} 的第 r 列是第 s 个坐标向量，A 乘它就选出 A 的第 s 列；第 s 列反过来选出第 r 列，其他列不变。对同一对行或同一对列交换两次，位置全部恢复。尤其将两次交换作用于单位矩阵，得到 $P_{ij}^2 = I_m$ 与 $Q_{rs}^2 = I_n$，这给出了双侧反操作。 ■

■ **例 2.9　同一个交换矩阵，左右两种结果**　取 $H = \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix}$、$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$，分别求 PH, HP。对方程组 $Hx = b$，哪一种操作只是在交换方程？

解　左乘给 $PH = \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix}$，右乘给 $HP = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$，两者不同。前者把整行重排，连同右端作 Pb 就保持原未知数的解；后者重排的是未知数前面的系数列。若右乘之后仍把未知数记成原来的 x，一般就改变了输入的含义。要保持对应关系，写 $x = Py$，原方程才变成 $(HP)y = b$；其中 y 是把原两个分量交换后的新输入。例 2.4 原解为 $x = (1, 2)$，换列后的新解是 $y = (2, 1)$，两者通过 $x = Py$ 对应，而非保持同一个坐标元组。 ♦

对于矩形矩阵，左右的交换矩阵甚至有不同尺寸：左边要有与行数相同的阶数，右边要有与列数相同的阶数。不能拿一个 m 阶交换矩阵随意右乘 $m \times n$ 矩阵，除非尺寸确实允许。把「交换行、交换列」写成矩阵式，也就把这种尺寸检查放进了公式本身。

每一步同时检查三件事　先确定操作改变的是方程行还是未知数列，再确认相乘尺寸，最后写出反操作。行操作连同右端一起做，保持同一输入的解；换列则需要记录未知数的重排。多步合成时，最后实施的操作写在最左边。

2.6　原习题综合：Pascal 矩阵的逐层消元

前面的消元矩阵一次只改变一行。原课程习题 2.2 给出一个更有结构的例子：一次把若干行分别减去它们上面的相邻行。矩阵乘法始终从旧矩阵读取各行，所以必须先弄清「同时作差」与「按某个次序依次作差」的关系。

■ **例 2.10　原习题 2.2：先缩小一层**　求下三角矩阵 D，使

$$D \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix}.$$

再求一个合成矩阵 M，把左边的 Pascal 矩阵 C 完全化为 I_4。

解　先直接比较目标各行。第一行保持；第二行是原第二行减原第一行；第三行是原第三行减原

第二行；第四行是原第四行减原第三行。于是

$$D=\begin{bmatrix}1&0&0&0\\-1&1&0&0\\0&-1&1&0\\0&0&-1&1\end{bmatrix}.$$

例如目标第三行为 $(1,2,1,0)-(1,1,0,0)=(0,1,1,0)$，第四行为 $(1,3,3,1)-(1,2,1,0)=(0,1,2,1)$。第一列从第二行起全为零，右下角剩一个三阶 Pascal 形状。下一次只在右下三阶部分作相邻差，得到

$$D_2=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&-1&1&0\\0&0&-1&1\end{bmatrix},\qquad D_2DC=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&1&1\end{bmatrix}.$$

第三次只需第四行减第三行，取

$$D_3=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&-1&1\end{bmatrix},\qquad D_3D_2DC=I_4.$$

所以合成矩阵为

$$M=D_3D_2D=\begin{bmatrix}1&0&0&0\\-1&1&0&0\\1&-2&1&0\\-1&3&-3&1\end{bmatrix}. \tag{2.2}$$

逐行相乘可检查 $MC=I_4$；例如末行为 $-R_1+3R_2-3R_3+R_4=(0,0,0,1)$。同样直接核对 $CM=I_4$，所以这里确实得到 C 的双侧逆。这个具体题的逆来自逐步消元，不需要先掌握一般求逆公式。 ♦

2.6.1 为什么相邻差要从下往上实施

写成 DC 时，每条新行都用 C 的旧行。若手上只保留一个会被覆盖的数组，想用三次行替换实现同一个结果，应先 $R_4\leftarrow R_4-R_3$，再 $R_3\leftarrow R_3-R_2$，最后 $R_2\leftarrow R_2-R_1$。每一步需要读取的上方行还没有被改变，所以与同时作差相符。反过来，若从上往下，先把第二行改为 $(0,1,0,0)$，第三行再减去这个*新*第二行，就会变成 $(1,1,1,0)$，与目标 $(0,1,1,0)$ 不同；再改第四行则得到 $(0,2,2,1)$，也已经不是原来的相邻差。问题不在减法算错，而在读取的数据版本不同。遇到一批看似可以同时实施的操作时，应先写成每个新行对旧行的组合，再确定安全的逐步次序。

想一想　Pascal 题中的 D 次对角线上有三个 -1，能不能把它的逆直接写成「把这三个 -1 都改为 1」？

答　不能。一步消元的反操作只改变一个非对角位置，而这里三个行替换相互依赖。D 把输入 (x_1, x_2, x_3, x_4) 变成 $(x_1, x_2 - x_1, x_3 - x_2, x_4 - x_3)$；要恢复 x_4，必须把前面所有差逐项累加。逆矩阵因此是下三角的累加矩阵，第 i 行前 i 个位置都是 1，其余为 0。只改次对角线的符号会漏掉这些累计贡献。

Pascal 的名字不承担证明工作。题中的每一步都可以用本讲已经证明的行组合与复合次序逐行核验。它展示了两种规模：单次操作只涉及两行，合成矩阵却可能含有更远位置的系数。主例合成矩阵出现的 6 与本题出现的交替符号，都是连续读取已更新行后留下的明确记录。

本讲小结

- **等价操作**：交换行、非零倍乘行、$R_i \leftarrow R_i + cR_j$ 都能撤销，因而保持全部解；操作包括右端。
- **乘数与主元**：用当前非零主元 p 消去同列系数 a，乘数为 $\ell = a/p$，不能除以零。
- **前向消元与回代**：先得 $[U \mid c]$；若 $u_{ii} \neq 0$，按 $i = n, \ldots, 1$ 唯一求出 $x_i = (c_i - \sum_{j>i} u_{ij} x_j)/u_{ii}$。
- **零主元处理**：下方同列有非零项则换行；确无候选时，方阵不能对任何右端恰有一个解，仍需区分矛盾行与冗余行。
- **消元矩阵**：$E_{ij}(\ell)$ 在单位阵第 (i, j) 元放 $-\ell$；它的逆为 $E_{ij}(-\ell)$。
- **复合顺序**：先做 E_1、再做 E_2，合成为 $E_2 E_1$；撤销时为 $E_1^{-1} E_2^{-1}$，一般不能交换因子。
- **置换的方向**：PA 交换行且右端也变为 Pb；AQ 交换列，保持原问题时需令原输入 $x = Qy$。

下一讲系统整理矩阵乘法的几种算法与运算律，并把这里见过的反操作推广为逆矩阵。消元过程已经给出了具体问题：怎样一并处理多个右端，以及怎样把一个可逆操作完整记下来。

练习

2.1 ★　不看例 2.1，对 $2x + 3y = 5$、$6x + 15y = 12$ 写出增广矩阵、乘数、主元、上三角系统和全部解。说明若只改系数、不改右端，会在哪一步出现错误。

2.2 ★★　保留例 2.2 的 A，将右端改为 $(1, 0, 1)$。沿用同样的两个乘数求新右端 c，再回代并检验原三行。

2.3 ★　对 $H = \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix}$、$b = (4, 11)$，写交换矩阵 P、PH、Pb。求原系统的解，并说明换列后 $(HP)y = b$ 的解与原输入有什么关系。

2.4 ★★　对例 2.5 的系统，逐次写出两次行替换和一次换行。为何交换第二、三行没有破坏第一列的零？将最终三个主元与原矩阵对角元比较。

2.5 ★★ 对 $B=\begin{bmatrix}1&2\\2&4\end{bmatrix}$、右端 (a,b)，从增广消元得出有解条件和全部解。说明为什么单看系数矩阵无法区分无解与无穷多解。

2.6 ★ 写出「第三行减去第一行的 4 倍」的矩阵 F 与反操作矩阵 G，算出例 2.7 中的 FX，验证 $GF=FG=I_3$。解释为什么需要 $i\neq j$。

2.7 ★★ 设 E_{21},E_{32} 为例 2.8 的两个消元矩阵。分别求 $E_{32}E_{21}$ 与 $E_{21}E_{32}$，解释前一个乘积第 $(3,1)$ 元为什么为 6。写恢复原矩阵的因子次序。

2.8 ★★ 设 P_{12} 交换三维输入的前两个分量，P_{23} 交换后两个分量。分别将 $P_{23}P_{12}$、$P_{12}P_{23}$ 作用于 $(1,2,3)$，判断两者是否相等。两个单次交换各自的逆是什么？

2.9 ★★ 对本讲的 Pascal 矩阵 C，用 (2.2) 中的 M 求 $Cx=(1,3,8,20)$。再把所得向量代入 C，逐行检查四个分量。

2.10 ★★★ 讨论参数系统 $x+y=2$、$2x+ay=5$ 的全部解。再将第二式右端改为 4，重新讨论。求解过程中说明在哪个参数值必须暂停除法，以及此时如何读增广矩阵。

练习答案与提示

2.1 增广矩阵为 $\left[\begin{array}{cc|c}2&3&5\\6&15&12\end{array}\right]$。乘数为 $6/2=3$；第二行减第一行的三倍后是 $[\,0\ \ 6\mid -3\,]$，故主元为 $2,6$，上三角系统为 $2x+3y=5$、$6y=-3$。回代给 $y=-1/2$、$x=13/4$，主元非零使输入唯一。原两行分别得到 $5,12$。若保留旧右端 12，便错写成 $6y=12$；这条式子不是原第二式减第一式三倍所得的结果，等价性已在那一步丢失。

2.2 第一阶段对第二个右端分量作 $0-3\cdot 1=-3$，第三分量仍为 1；第二阶段作 $1-2(-3)=7$，故 $c=(1,-3,7)$。回代先得 $z=7/5$，再由 $2y-14/5=-3$ 得 $y=-1/10$；最后 $x=1-2(-1/10)-7/5=-1/5$。代入原三行，依次为 $-1/5-1/5+7/5=1$、$-3/5-4/5+7/5=0$、$-2/5+7/5=1$。主元仍为 $1,2,5$，故为唯一解。

2.3 $P=\begin{bmatrix}0&1\\1&0\end{bmatrix}$，$PH=\begin{bmatrix}3&4\\0&2\end{bmatrix}$、$Pb=(11,4)$。回代得原输入 $x=(1,2)$。右乘时 $HP=\begin{bmatrix}2&0\\4&3\end{bmatrix}$，解为 $y=(2,1)$，因为 $2y_1=4$、$4y_1+3y_2=11$。对应关系为 $x=Py$，不能把 y 的分量仍解释成原来同序的未知数。

2.4 两次行替换为 $R_2\leftarrow R_2-2R_1$、$R_3\leftarrow R_3-R_1$，得到 $\left[\begin{array}{ccc|c}1&1&1&6\\0&0&1&3\\0&2&3&13\end{array}\right]$；随后交换第二、三行。两条被交换的行第一列都为零，所以交换后第一列主元下方仍全为零。最终主元为 $1,2,1$，而原对角元为 $1,2,4$，后二者不能照抄为消元主元。回代得 $(1,2,3)$，代入原矩阵得 $(6,15,19)$。

2.5 作 $R_2\leftarrow R_2-2R_1$ 后，末行为 $[\,0\ \ 0\mid b-2a\,]$。有解必须且只需 $b=2a$；必要性来自末行不能要求 0 等于非零数，充分性由全部候选 $(x_1,x_2)=(a-2t,t)$ 给出。每个候选代入得到 $(a,2a)$，每个解也由第一式具有此形。条件满足时有无穷多解，不满足时无解；系数矩阵固定，但末行右端随 a,b 改变。

2.6 $F=\begin{bmatrix}1&0&0\\0&1&0\\-4&0&1\end{bmatrix}$，$G=\begin{bmatrix}1&0&0\\0&1&0\\4&0&1\end{bmatrix}$。$FX$ 的前两行不变，第三行为 $(4,5,6)-4(1,2,3)=(0,-3,-6)$。相乘 GF 时末行是 $4(1,0,0)+(-4,0,1)=(0,0,1)$，前两行是单位阵前两行；$FG$ 的末行则为

$-4(1,0,0)+(4,0,1)=(0,0,1)$，故两个乘积都为 I_3。条件 $i\neq j$ 保证作为参照的行不变；若把一行减去自身，整行成为零，无法用「加回来」恢复旧行。

2.7 两个乘积依次为 $\begin{bmatrix}1&0&0\\-3&1&0\\6&-2&1\end{bmatrix}$ 与 $\begin{bmatrix}1&0&0\\-3&1&0\\0&-2&1\end{bmatrix}$。前一个先更新第二行，再从第三行减它的两倍，因此第三行为 $R_3-2(R_2-3R_1)$，第一行的系数是 $(-2)(-3)=6$。后一个先更新第三行，此时读取的仍是旧第二行，没有这项贡献。恢复的时间次序为先撤销 E_{32}、再撤销 E_{21}，合成写成 $E_{21}^{-1}E_{32}^{-1}$。

2.8 对第一个乘积，从右边开始，$(1,2,3)\xrightarrow{P_{12}}(2,1,3)\xrightarrow{P_{23}}(2,3,1)$。对第二个，$(1,2,3)\xrightarrow{P_{23}}(1,3,2)\xrightarrow{P_{12}}(3,1,2)$。同一个输入的两个输出不同，说明矩阵不相等。每个单次交换的逆都是它自身，因为交换同一对分量两次会恢复原顺序；但多次交换的反操作仍必须按相反的时间次序执行。

2.9 乘 M 给 $x_1=1$、$x_2=-1+3=2$、$x_3=1-2\cdot3+8=3$、$x_4=-1+3\cdot3-3\cdot8+20=4$，所以 $x=(1,2,3,4)$。原 C 的四行输出依次为 1、$1+2=3$、$1+2\cdot2+3=8$、$1+3\cdot2+3\cdot3+4=20$。因为每一步消元可撤销，且最后为单位阵，这确为全部解；也可利用 C 非零对角元由上往下代入，依次唯一求出相同的四个分量。

2.10 第二行减第一行的两倍，得到 $(a-2)y=1$。若 $a\neq2$，可以除以 $a-2$，唯一解为 $y=1/(a-2)$、$x=2-1/(a-2)$；回代第二式为 $2(2-y)+ay=4+(a-2)y=5$。若 $a=2$，末行是 $0=1$，无解，不能把分母为零的表达式当作答案。将原第二式右端改为 4 后，消元变为 $(a-2)y=0$：$a\neq2$ 时唯一解为 $(2,0)$；$a=2$ 时末行为 $0=0$，全部解为 $(2-t,t)$，任意实数 t 都满足两行，且第一行保证没有别的形式。

第 3 讲

矩阵乘法与逆矩阵

Multiplication and Inverse Matrices

对应材料：MIT 18.06SC Session 1.3（Strang 第 3 讲视频）课时摘要；习题 3.1、3.2 及解答（3.2 即 Strang《Introduction to Linear Algebra》2.5 节第 24 题）；参考该书 2.4、2.5 节。

本讲目标

(1) 掌握矩阵乘法 AB 的形状规则与「行乘列」公式 $c_{ij}=\sum_k a_{ik}b_{kj}$，并知道它为什么这样定义。

(2) 会用另外三种方式看同一个乘积：按列、按行、列乘行之和；能根据问题挑选最方便的一种。

(3) 熟悉乘法的运算律：结合律、分配律成立，交换律不成立；会做分块乘法。

(4) 理解逆矩阵的定义与唯一性；知道「存在 $x\neq 0$ 使 $Ax=0$」意味着 A 不可逆。

(5) 会用 Gauss–Jordan 消元 $[\,A\mid I\,]\to[\,I\mid A^{-1}\,]$ 求逆，并能在消元中途识别奇异矩阵。

(6) 掌握方阵可逆的几个等价条件，以及「方阵的单侧逆就是逆」这一结论。

3.1 矩阵乘法：行乘列

第 2 讲里我们已经在用矩阵乘法了：消元矩阵 E_{21} 左乘 A 完成一步消元，$E_{32}(E_{21}A)=U$ 完成整个前向消元，而把括号挪成 $(E_{32}E_{21})A$ 用到了结合律。这一讲回头把乘法本身讲清楚，回答两个问题：两个矩阵怎样相乘？为什么要这样定义？然后转到除法的替代品——逆矩阵。

3.1.1 从「代换的复合」看乘法的定义

设 B 是 $n\times p$ 矩阵，A 是 $m\times n$ 矩阵。B 把 $x\in\mathbb{R}^p$ 变成 $y=Bx\in\mathbb{R}^n$，A 再把 y 变成 $z=Ay\in\mathbb{R}^m$。我们想找一个矩阵 C，一步把 x 直接变成 z。逐个分量写出来：

$$y_k=\sum_{j=1}^{p} b_{kj}x_j,$$

$$z_i=\sum_{k=1}^{n} a_{ik}y_k=\sum_{j=1}^{p}\Big(\sum_{k=1}^{n} a_{ik}b_{kj}\Big)x_j.$$

所以 $z=Cx$，其中 C 的 (i,j) 元必须是括号里的那个和。我们就把这个 C 定义为乘积 AB。乘法规则来自「先做 B、再做 A」这种复合，它不是凭空约定的。

定义 3.1（矩阵乘法）　设 A 是 $m\times n$ 矩阵，B 是 $n\times p$ 矩阵。它们的**乘积**（product）$C=AB$ 是 $m\times p$ 矩阵，其元素为

$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj},$$

即 A 的第 i 行与 B 的第 j 列的点积。只有当 A 的列数等于 B 的行数 时 AB 才有定义。

形状规则可以记成一句话：

$$\boxed{(m\times n)\,(n\times p)=(m\times p).}$$

中间两个 n 必须相同并且「消掉」，两头的 m,p 给出乘积的形状。图 3.1 画出了相乘所需的尺寸与复合顺序。

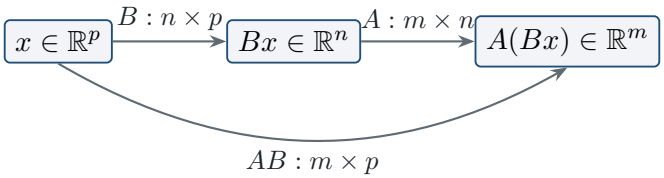


图 3.1　矩阵乘积记录先用 B、再用 A 的复合；计算每个输出分量时，内层指标逐项求和。

图中间的空间有 n 个坐标，所以内层相加恰有 n 项。AB 的尺寸只记录最终输入与输出，不能从最终尺寸猜测内层是否兼容。图里先作用的因子写在右端，和第 2 讲合成消元的顺序完全一致。

■ **例 3.1　把四个元素逐个算出来**　设 $A=\begin{bmatrix}1&2\\3&4\end{bmatrix}$，$B=\begin{bmatrix}5&7\\6&8\end{bmatrix}$，求 AB。说明左上角的 17 和右下角的 53 分别来自哪一行与哪一列。

解　按行乘列分别求和，第一行给 $1\cdot5+2\cdot6=17$、$1\cdot7+2\cdot8=23$，第二行给 $3\cdot5+4\cdot6=39$、$3\cdot7+4\cdot8=53$。因此

$$AB=\begin{bmatrix}17&23\\39&53\end{bmatrix}.$$

左上角取 A 第一行与 B 第一列，右下角取 A 第二行与 B 第二列。两个原矩阵都为 2×2，每个元素由两个乘积相加得到；不是把相同位置的两个元素直接相乘。若按位置相乘，左上角只得 5，它没有包含经过中间第二个分量的贡献。　♦

单位矩阵的作用也能逐元素核对：$I_mA=A$、$AI_n=A$，两侧的单位矩阵尺寸未必相同。例如一个 2×3 矩阵左边要用 I_2，右边要用 I_3。零矩阵在相容尺寸下与任意矩阵相乘，结果为对应尺寸的零矩阵，因为每个求和项都是零。

想一想　知道 AB 有定义，是否就知道 BA 也有定义？即使两者都有定义，能否保证相等？
答　若 A 为 2×3、B 为 3×4，AB 为 2×4，而 BA 的内部尺寸 $4,2$ 不同，没有定义。若两者都为同阶方阵，两种乘积都有定义，仍可能不相等；后面会给出明确反例。必须分别检查「有没有定义」与「是否相等」，它们是两个问题。

3.2　同一个乘积的列、行与外积

例 3.1 的四个数可以逐个求，也可以一次读出一列或一行。不同算法没有改变乘法本身，而是把同一个有限求和按不同方式组织。计算目的不同，最省力的组织方式也不同：多个右端适合按列看，行操作适合按行看，结构分解则适合一列乘一行。

命题 3.2（按列与按行读乘积）　设 A 为 $m\times n$，B 为 $n\times p$。若 b_j 为 B 第 j 列，则 AB 的第 j 列为 Ab_j；若 r_i 为 A 第 i 行，则 AB 的第 i 行为 r_iB。因此结果的列是 A 的列的组合，结果的行是 B 的行的组合。

证明　固定 j 后，$(AB)_{ij}=\sum_k a_{ik}b_{kj}$ 正是 A 乘向量 b_j 的第 i 分量，所以整列为 Ab_j。固定 i 后，同一公式对 $j=1,\dots,p$ 排成一行，为 $\sum_k a_{ik}\,\mathrm{row}_k(B)$，即 r_iB。前者的系数来自 B 的一列，后者的系数来自 A 的一行，不能把所组合的矩阵弄反。 ■

■ **例 3.2　重算同一个乘积**　不用逐个算四个元素，按列与按行重新计算例 3.1 的 AB。

解　按列，第一列为 $5(1,3)+6(2,4)=(17,39)$，第二列为 $7(1,3)+8(2,4)=(23,53)$。把它们依次排成两列，得到相同的矩阵。按行，第一行为 $1(5,7)+2(6,8)=(17,23)$，第二行为 $3(5,7)+4(6,8)=(39,53)$，把它们排成两行，仍得到相同的四个数。按列同时计算同一个右端的所有输出分量，按行同时计算同一个输出分量对所有右端的结果。两种组织方式都保留了原指标的含义；若按列算出 $(17,39)$ 却写成第一行，所得矩阵就被错误地重新排列了。 ♦

定义 3.3（一列乘一行与外积）　设 $u\in\mathbb{R}^m$ 是列向量，$r=(r_1,\dots,r_p)$ 是行向量，尺寸为 $1\times p$。矩阵 ur 为 $m\times p$，第 (i,j) 元是 u_ir_j，称为一个**外积**（outer product）。这里的乘法与行乘列的点积有不同的尺寸。

例如一条 1×3 行乘一条 3×1 列，结果是 1×1 的数；反过来一列乘一行，结果为 3×3 矩阵。两种写法不能只因都含三项就混用。若列或行全零，外积也全零；若两者都非零，外积的各列都是同一列的倍数，各行也是同一行的倍数。

■ **例 3.3　原摘要的一列乘一行**　求 $\begin{bmatrix}2\\3\\4\end{bmatrix}\begin{bmatrix}1&6\end{bmatrix}$，说明结果各行与各列的关系。

解　尺寸为 $(3\times 1)(1\times 2)=3\times 2$，每个元素只有一项乘积，得到 $\begin{bmatrix}2&12\\3&18\\4&24\end{bmatrix}$。第二列是第一列的 6 倍；三行依次为 $2,3,4$ 倍的行 $(1,6)$。作为二维点，三行 $(2,12),(3,18),(4,24)$ 都在过原点的直线 $y=6x$ 上。图 3.2 标出这些行所对应的点。列向量则生活在三维空间，不能把这两种图像画在同一个二维坐标系里。第 9、10 讲会用空间和维数把这种「所有行或所有列各沿一个方向」的现象系统表述。 ♦

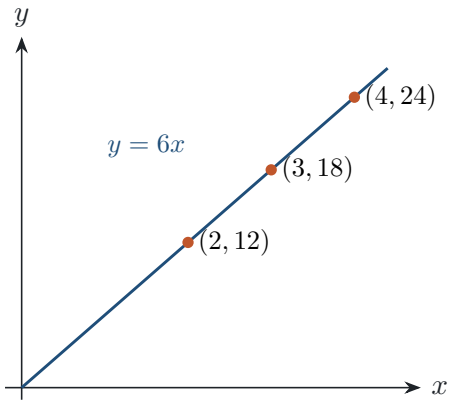


图 3.2　外积矩阵的三行作为二维点共线；整条过原点直线是行向量 $(1,6)$ 的所有实倍数。

命题 3.4（乘积是若干外积之和）　若 $A=[a_1\ \cdots\ a_n]$，B 的行依次为 $r_1,\dots,r_n$，则 $AB=\sum_{k=1}^{n}a_k r_k$。

证明　第 k 个外积的第 (i,j) 元为 $a_{ik}b_{kj}$，把 $k=1,\dots,n$ 的这些矩阵相加，对应元素之和为 $\sum_k a_{ik}b_{kj}$，恰为 AB 的定义。每一项尺寸都是 $m\times p$，所以矩阵加法也有定义。 ■

对例 3.1，两项外积分别为 $\begin{bmatrix}5&7\\15&21\end{bmatrix}$、$\begin{bmatrix}12&16\\24&32\end{bmatrix}$，相加正是 $\begin{bmatrix}17&23\\39&53\end{bmatrix}$。单项外积的行列均只有一个方向，多个外积相加却可以得到更丰富的矩阵；不能把单项的几何结论直接套给它们的总和。

3.3　运算律与分块：可以改什么次序

第 2 讲的消元矩阵已经说明因子位置会影响结果。我们需要区分三件事：改变括号、交换因子、把乘法分配到加法的各项。实数乘法中这些操作常被一并使用；矩阵乘法只保留其中部分规则，而且每个式子都必须先满足尺寸要求。

命题 3.5（结合律与分配律）　设 A 为 $m\times n$、B 为 $n\times p$、C 为 $p\times q$，则 $(AB)C=A(BC)$。当加法与乘法尺寸相容时，还成立 $A(B+D)=AB+AD$、$(A+F)B=AB+FB$；单位矩阵满足 $I_mA=A=AI_n$。

证明　结合律两端的第 (i,j) 元都等于有限双重和 $\sum_{k=1}^{n}\sum_{r=1}^{p} a_{ik}b_{kr}c_{rj}$。先对 k 求和得到 $(AB)C$，先对 r 求和得到 $A(BC)$；有限个实数的加法可重排，所以两端逐元素相同。分配律则把 $\sum_k a_{ik}(b_{kj}+d_{kj})$ 拆为两个和，或把 $\sum_k (a_{ik}+f_{ik})b_{kj}$ 拆开即可。单位矩阵在相应求和中只有一个非零系数 1，故选回原元素。 ■

结合律允许省略一串乘积中的括号，但不能改因子的先后。对于一个向量 x，$A(Bx)$ 与 $(AB)x$ 输出相同，计算量却可能不同：若只需要一个输出，先算 Bx 往往比先形成整个 AB 更省工作。输出相同不等于每种计算安排都同样合适。

■ **例 3.4　原习题 3.1：验证分配律**　取 $A=\begin{bmatrix}1&2\\3&4\end{bmatrix}$、$B=\begin{bmatrix}1&0\\0&0\end{bmatrix}$、$C=\begin{bmatrix}0&0\\5&6\end{bmatrix}$。分别算 $AB+AC$ 与 $A(B+C)$。

解　先分别乘，$AB=\begin{bmatrix}1&0\\3&0\end{bmatrix}$，$AC=\begin{bmatrix}10&12\\20&24\end{bmatrix}$，相加得 $\begin{bmatrix}11&12\\23&24\end{bmatrix}$。另一条路线先加 $B+C=\begin{bmatrix}1&0\\5&6\end{bmatrix}$，再乘 A，第一行给 $(1+10,0+12)$，第二行给 $(3+20,0+24)$，得到同一个结果。这道题检验的是已证明的逐元素分配规则。一次数值相等不能代替对所有矩阵的证明，但能帮助发现漏乘一项或混淆左右位置的错误。 ♦

■ **例 3.5　交换、消去因子与零乘积的反例**　取 $N=\begin{bmatrix}0&1\\0&0\end{bmatrix}$、$R=\begin{bmatrix}0&0\\1&0\end{bmatrix}$。比较 NR, RN, N^2；再取 $Z=\begin{bmatrix}1&0\\0&0\end{bmatrix}$，比较 ZI_2 与 $Z\begin{bmatrix}1&0\\0&2\end{bmatrix}$。

解　直接相乘得 $NR=\begin{bmatrix}1&0\\0&0\end{bmatrix}$、$RN=\begin{bmatrix}0&0\\0&1\end{bmatrix}$，所以因子不能任意交换。另一方面 $N\neq 0$，却有 $N^2=0$，所以乘积为零不能推出某个因子为零。末两项都等于 Z，但 $I_2\neq\begin{bmatrix}1&0\\0&2\end{bmatrix}$。因此由 $ZB=ZC$ 不能在没有额外条件时「约去 Z」。后面会证明：若左边的公共方阵确实可逆，左乘它的逆才有合法消去的依据。 ♦

标量展开不能原样照搬　对同阶方阵，$(A+B)^2=A^2+AB+BA+B^2$，不能总合并为 $A^2+2AB+B^2$，因为 AB 未必等于 BA。先按分配律保留因子的相对顺序，只有另行证明 $AB=BA$ 后才可合并。矩阵中没有一种随手写成 B/A 就能同时替代左乘逆与右乘逆的通用记号。

3.3.1　分块不是省掉尺寸检查

把矩阵按连续的若干行列切成小矩阵，称为**分块**（block partition）。分块后仍可按「行乘列」写乘积，但各块本身也要能相乘，同一求和中的结果块还要有相同尺寸。用线条把矩阵画成四格，并不会自动保证乘法正确。

命题 3.6（相容分块的乘法）　设 $A=\begin{bmatrix}A_1 & A_2\\ A_3 & A_4\end{bmatrix}$、$B=\begin{bmatrix}B_1 & B_2\\ B_3 & B_4\end{bmatrix}$，切分使 A 的两段列宽分别等于 B 的两段行高，则

$$AB=\begin{bmatrix}A_1B_1+A_2B_3 & A_1B_2+A_2B_4\\ A_3B_1+A_4B_3 & A_3B_2+A_4B_4\end{bmatrix}.$$

证明　每个结果元素的内层和按切分位置分成前后两段。处于左上结果块时，第一段是 A_1 相应行与 B_1 相应列的点积，第二段是 A_2 与 B_3 的点积，所以对应元素为 $A_1B_1+A_2B_3$。其余三块分别在另三组输出行列中作同样的两段拆分，给出公式。分块不改变任何元素，只改变有限求和的分组。 ■

■ **例 3.6　矩形矩阵的相容切分**　计算 $K=\begin{bmatrix}1 & 2 & 0\\ 3 & 4 & 1\end{bmatrix}$ 与 $L=\begin{bmatrix}1 & 0\\ 0 & 1\\ 2 & 3\end{bmatrix}$ 的乘积，把内部尺寸 3 切成 $2+1$。

解　将 K 按列写成 $[K_1\ K_2]$，其中 $K_1=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$ 为 2×2，$K_2=(0,1)^T$ 为 2×1 列。将 L 按行写成 $\begin{bmatrix}L_1\\ L_2\end{bmatrix}$，其中 $L_1=I_2$，$L_2=(2,3)$ 为 1×2 行。因此

$$KL=K_1L_1+K_2L_2=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}+\begin{bmatrix}0 & 0\\ 2 & 3\end{bmatrix}=\begin{bmatrix}1 & 2\\ 5 & 7\end{bmatrix}.$$

两项都是 2×2，可以相加。按普通行乘列核对末行为 $3\cdot 1+4\cdot 0+1\cdot 2=5$、$3\cdot 0+4\cdot 1+1\cdot 3=7$。块可以是矩形，单位块与外积块也可以出现在同一次分块计算中。 ♦

3.4　逆矩阵：把每个输入准确恢复

第 2 讲中，减去另一行的若干倍可以用加回来撤销，交换两行也能通过再交换恢复。任意方阵 A 是否也有一种统一的恢复操作？我们要求恢复操作对所有输入都有效，而不只把某一个特殊向量送回去。

定义 3.7（逆矩阵与可逆）　设 A 为 $n\times n$ 实矩阵。若存在同阶矩阵 B 满足 $BA=I_n$ 且 $AB=I_n$，称 B 为 A 的**逆矩阵**（inverse matrix），记为 A^{-1}，并称 A **可逆**（invertible）。暂不预设每个方阵都可逆，两个乘法方向都属于定义。

例如 $I_n^{-1}=I_n$；交换同一对分量的矩阵以自身为逆；一步消元 $E_{ij}(\ell)$ 的逆为 $E_{ij}(-\ell)$。零方阵在 $n\geq 1$ 时不可逆，因为无论乘什么矩阵，结果都为零，不会是 I_n。对角矩阵若各对角元非零，把每个对角元取倒数便给出双侧逆；若某个对角元为零，对应坐标方向就被完全消去，无法恢复。

命题 3.8（逆的唯一性、解方程与消去因子）　若方阵 A 可逆，则逆矩阵唯一；对每个 b，$Ax=b$ 的唯一解为 $x=A^{-1}b$。同阶且尺寸相容时，$AB=AC$ 可推出 $B=C$，$BA=CA$ 也可推出 $B=C$。

证明　若 B,C 都为逆，利用结合律，$B=B(AC)=(BA)C=C$，所以唯一。对方程，任意解左乘 A^{-1} 得 $x=A^{-1}b$；反过来 $A(A^{-1}b)=(AA^{-1})b=b$，故这个候选确为解。前一方向排除其他解，后一方向保证存在。若 $AB=AC$，在两边*左*乘 A^{-1}，得 $B=C$。若 $BA=CA$，则在两边*右*乘 A^{-1}。选择相乘的一侧，是为了让 A 与它的逆相邻；不能在任意一侧乘后就跨过中间因子抵消。 ■

■ **例 3.7　原摘要的可逆矩阵**　核对 $V=\begin{bmatrix}1&3\\2&7\end{bmatrix}$ 的逆为 $W=\begin{bmatrix}7&-3\\-2&1\end{bmatrix}$，并解 $Vx=(5,12)$。

解　先核对两个方向。VW 的四个元素依次为 $7-6=1$、$-3+3=0$、$14-14=0$、$-6+7=1$。WV 的四个元素依次为 $7-6=1$、$21-21=0$、$-2+2=0$、$-6+7=1$。所以 $VW=WV=I_2$，确有 $W=V^{-1}$。乘右端，$x=W(5,12)=(35-36,-10+12)=(-1,2)$。回代 V 给 $-1+6=5$、$-2+14=12$。逆的存在保证这个解唯一。这里只验证了给定的候选，怎样从 V 出发把 W 算出来，要用下一节的增广消元。 ♦

命题 3.9（非零的零输入方向阻止可逆）　若 A 为方阵，存在 $u\neq 0$ 满足 $Au=0$，则 A 不可逆。

证明　假设有双侧逆，则 $u=Iu=(A^{-1}A)u=A^{-1}(Au)=A^{-1}0=0$，与 $u\neq 0$ 矛盾。逆若能恢复所有输入，就不能把一个非零输入与零输入都送到同一个零输出后还区分开来。 ■

原摘要用 $S=\begin{bmatrix}1&3\\2&6\end{bmatrix}$ 说明这个障碍：$S(3,-1)=(0,0)$，因为三倍第一列等于第二列。输入 $(3,-1)$ 与 $(0,0)$ 输出相同，所以任何后续矩阵都不可能把同一个零向量分别恢复为这两个输入。第 1 讲把「列组合填不满整个空间」的方阵称为奇异；稍后将借助消元证明，它与「不可逆」的用法一致。行列式判别留到第 17 讲。

命题 3.10（乘积的逆必须倒序）　若同阶方阵 A,B 都可逆，则 AB 可逆，且 $(AB)^{-1}=B^{-1}A^{-1}$。

证明　设候选为 $C=B^{-1}A^{-1}$。按结合律保留因子次序，有 $(AB)C=A(BB^{-1})A^{-1}=I$，也有 $C(AB)=B^{-1}(A^{-1}A)B=I$。两个方向都验证了逆的定义，唯一性使候选恰为 $(AB)^{-1}$。这个结论与先撤销最后一步、再撤销最初一步的操作顺序相符。 ■

■ **例 3.8　倒序与不倒序会给不同答案**　取 $A=\begin{bmatrix}1&2\\0&1\end{bmatrix}$、$B=\begin{bmatrix}1&0\\3&1\end{bmatrix}$，比较 $B^{-1}A^{-1}$ 与 $A^{-1}B^{-1}$。

解　分别撤销一个剪切操作，得 $A^{-1}=\begin{bmatrix}1&-2\\0&1\end{bmatrix}$、$B^{-1}=\begin{bmatrix}1&0\\-3&1\end{bmatrix}$。原乘积 $AB=\begin{bmatrix}7&2\\3&1\end{bmatrix}$。正确倒序给 $B^{-1}A^{-1}=\begin{bmatrix}1&-2\\-3&7\end{bmatrix}$，按刚才的双侧核验它就是 AB 的逆；不倒序则得 $A^{-1}B^{-1}=\begin{bmatrix}7&-2\\-3&1\end{bmatrix}$，与正确结果不同。例如把后一个错误候选乘在 AB 右边，左上角得到 $7\cdot7+2(-3)=43$，而单位阵该位置为 1，立刻排除。检查一项就能推翻一个错误候选；要确认一个候选正确，则仍需检查定义的完整条件。 ♦

想一想　既然可逆时能「除去 A」，为什么还坚持不用一个统一的 B/A 符号？

答　方程 $AX=B$ 的解是 $X=A^{-1}B$，而 $YA=B$ 的解是 $Y=BA^{-1}$，两式中的未知矩阵位置不同。这两个结果一般不同，也可能需要不同的尺寸。明确写出左乘还是右乘逆，才能保持原方程的排列。即使 A 可逆，也不会因此让所有矩阵都与它交换。

3.5　Gauss–Jordan：同时求多个右端

例 3.7 的逆矩阵有两列，它们分别将两个坐标右端解回输入。把 $V^{-1}=[x_1\ x_2]$ 写入 $VV^{-1}=I$，就是同时要求 $Vx_1=e_1$、$Vx_2=e_2$。同一个系数矩阵的消元乘数与右端无关，因此两列可以放在一起运算。

定义 3.11（Gauss–Jordan 消元）　对增广矩阵 $[A\mid I_n]$ 作等价行操作，不只消去主元下方的元素，还把主元归一为 1，消去它上方的其余元素，直到左半边变成 I_n，称为**Gauss–Jordan 消元**（Gauss–Jordan elimination）。若全过程成功，右半边给出逆矩阵。

这个定义描述成功时的目标，不保证任意 A 都能做到。所有行操作必须作用于整行共 $2n$ 个元素；右边的单位阵记录这些操作对各个坐标右端的影响。前向消元到上三角形后，可以从最后一个主元开始，依次消去上方；也可以在取得每个主元时同时清掉该列其他位置，只要步骤合法且追踪一致即可。

■ **例 3.9　原摘要的完整求逆过程**　从 $V=\begin{bmatrix}1&3\\2&7\end{bmatrix}$ 出发，不猜答案地求 V^{-1}。

解　写出两个右端，第一步 $R_2\leftarrow R_2-2R_1$：

$$\left[\begin{array}{cc|cc}1&3&1&0\\2&7&0&1\end{array}\right]\longrightarrow\left[\begin{array}{cc|cc}1&3&1&0\\0&1&-2&1\end{array}\right].$$

第二主元已是 1，不需要缩放。继续 $R_1\leftarrow R_1-3R_2$，得到

$$\left[\begin{array}{cc|cc}1&0&7&-3\\0&1&-2&1\end{array}\right].$$

右边两列分别是 $Vx=e_1$、$Vx=e_2$ 的解，故候选为 $\begin{bmatrix}7&-3\\-2&1\end{bmatrix}$。例 3.7 已独立核对两个方向的

乘积为 I_2，所以它确为逆。与只求一个 b 相比，左半边操作完全相同，右半边同时追踪两个独立的目标。 ♦

> **命题 3.12（为什么右半边确实是双侧逆）**　若一串等价行操作将 $[A \mid I_n]$ 变为 $[I_n \mid E]$，则 $E = A^{-1}$。

证明　每一步行操作都可表示为一个可逆矩阵左乘：行交换和行替换的反操作已在第 2 讲证明；非零倍乘一行则由同一位置乘倒数撤销。按实际次序将这些矩阵合成，得到可逆矩阵 F。右半边由 I_n 变为 $FI_n = F$，所以观测到的 E 就等于 F；左半边满足 $FA = I_n$。由于 F 可逆，左乘 F^{-1} 得 $A = F^{-1}$，故 $AF = F^{-1}F = I_n$。两侧乘积均为单位阵，$F = A^{-1}$。证明使用了整串操作本身可逆，不能只把任意矩阵的单侧等式直接当成双侧定义。 ■

若对 $S = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ 作相同的第一步，得到 $\left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right]$。左边第二行已经为零，没有第二主元；右边这行说明两个坐标右端中并非都可达到。继续对零系数行作缩放或与其他行组合，都不能得到左边的单位阵。算法遇到这个障碍应报告不可逆，而不能把任意一个残留右半块称为逆。

■ **例 3.10　原习题 3.2：参数上三角矩阵**　对任意实数 a, b, c，求 $U = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$ 的逆。每步说明使用的是已经更新的哪一行。

解　三个对角主元均为 1，无需向下消元。先 $R_2 \leftarrow R_2 - cR_3$，再从第一行消去第二列，使用*更新后的*第二行作 $R_1 \leftarrow R_1 - aR_2$，于是

$$[U \mid I_3] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & a & b & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & -c \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & b & 1 & -a & ac \\ 0 & 1 & 0 & 0 & 1 & -c \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right].$$

最后 $R_1 \leftarrow R_1 - bR_3$，便得到

$$U^{-1} = \begin{bmatrix} 1 & -a & ac - b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.1}$$

独立核对 UU^{-1} 的右上角为 $(ac - b) + a(-c) + b = 0$，$U^{-1}U$ 的右上角为 $b - ac + (ac - b) = 0$，其他非对角项为 $a - a$、$c - c$，对角项为 1。所以公式对所有实数参数成立，不需要排除某个参数值。 ♦

原解答的某条行操作把 $R_2 \leftarrow R_2 - cR_3$ 误写成减去 cR_2，末尾也把 U^{-1} 误记成 L^{-1}；这里按实际方程与乘积核对改正。若先消去第一行的第二列，也可得到中间项 $b - ac$，最后再消掉它，结果仍相同；必须随所选次序更新每一个系数，不能从两条路线各取一步拼在一起。

逆不是逐元素取倒数　取 $a=2,b=1,c=3$，公式给 $U^{-1}=\begin{bmatrix}1&-2&5\\0&1&-3\\0&0&1\end{bmatrix}$。右上角 $5=2\cdot3-1$ 包含间接经过第二分量的贡献；它既不是 $1/b$，也不是简单的 $-b$。逆要撤销整个矩阵作用，逐元素取倒数不符合矩阵乘法的定义。

3.6　方阵可逆的等价检验

现在把本讲用过的判断方法接起来：找到双侧逆当然可解每个右端；只知道齐次方程只有零解，是否已经够了？对方阵，消元把这个问题转成「每个主元能否出现」，不需要预先引用行列式或秩定理。

思路　逆矩阵立即给出每个方程的唯一解；唯一解条件又排除非零齐次解。再证明只有零解会迫使消元得到 n 个非零主元，最后继续作 Gauss–Jordan 行操作，把左半块化为单位阵并得到逆矩阵。命题 3.13 负责证明消元不能中断这一步。

命题 3.13（齐次方程阻止缺主元）　若方阵 A 经消元、允许交换未处理的行后，在第 k 列找不到非零主元，则存在 $u\neq0$ 使 $Au=0$。

证明　先设已得到的前 $k-1$ 个主元均非零。当前列从第 k 行起全为零，令当前变换后矩阵为 T。取 $u_k=1$，$u_{k+1}=\cdots=u_n=0$。从第 k 行往下，先前 $k-1$ 列下方均为零，第 k 列按假设也为零，所以这些行的 Tu 为零。从第 $k-1$ 行向上，依次选 u_i 使第 i 行的输出为零。第 i 行在第 $1,\dots,i-1$ 列的系数已经为零，而 $u_{i+1},\dots,u_n$ 都已选定；u_i 的系数是非零第 i 主元，故恰有一个可选值。得到 $Tu=0$ 且 $u\neq0$。变换 $T=EA$ 是有限个可逆行操作矩阵的乘积，$Tu=0$ 于是 $Au=E^{-1}Tu=0$。若 $k=1$，无需向上回代，直接取 $u=(1,0,\dots,0)$；结论仍成立。 ■

定理 3.14（方阵可逆的四种等价说法）　对 $n\times n$ 实矩阵 A，以下命题等价：

(1) A 可逆；

(2) 对每个 $b\in\mathbb{R}^n$，$Ax=b$ 有唯一解；

(3) $Ax=0$ 只有零解；

(4) 经允许的行交换与消元可得到 n 个非零主元。

定理 3.14 的证明　**(1)⇒(2)**。取 $x=A^{-1}b$，逆的定义保证 $Ax=b$；若 $Ax=Ay=b$，左乘 A^{-1} 得 $x=y$。

(2)⇒(3)。令 $b=0$。零向量是解，条件说它是唯一解。

(3)⇒(4)。若消元没有 n 个主元，命题 3.13 会给出非零齐次解，与 (3) 矛盾。因此每列都出现主元。

(4)⇒(1)。对 $[A\mid I_n]$ 同步消元，再从末行向上归一化主元并消去上方元素，得到 $[I_n\mid E]$。由命题 3.12，$E=A^{-1}$，故 A 可逆。

■

例 3.11　由齐次方程直接识别奇异　判定 $S=\begin{bmatrix}1 & 3\\ 2 & 6\end{bmatrix}$ 是否可逆，并说明对所有右端是否都有唯一解。

解　第二列等于第一列的 3 倍，所以 $S(3,-1)=(0,0)$，其中 $(3,-1)\neq 0$。由定理 3.14，S 不可逆；例如右端为零时至少有两个解 $(0,0)$ 与 $(3,-1)$，因此不可能对每个右端都有唯一解。实际上，第二输出分量永远是第一分量的两倍，所以右端 $(1,0)$ 根本不可达。 ♦

反例说明定理中「方阵」条件不能随意删去。矩形矩阵的左右逆可以只存在一个方向，且列无关与输出填满不是一回事。用 $C=\begin{bmatrix}1 & 0\end{bmatrix}$、$D=\begin{bmatrix}1\\ 0\end{bmatrix}$，有 $CD=I_1$，但 $DC=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\neq I_2$。尺寸允许的单侧恒等式不能直接当作双侧逆。

> **命题 3.15（方阵的单侧逆足以推出双侧逆）**　若 A,B 是同阶方阵且 $BA=I$，则 $AB=I$；若 $AB=I$，也必有 $BA=I$。

证明　若 $BA=I$，则 $Ax=0$ 蕴含 $x=BAx=B0=0$，所以齐次方程只有零解。由定理 3.14，A 可逆；在 $BA=I$ 两边右乘 A^{-1}，得 $B=A^{-1}$，于是 $AB=I$。若 $AB=I$，则 $Bx=0$ 蕴含 $x=ABx=A(Bx)=0$，所以 $Bx=0$ 只有零解。由定理 3.14，B 可逆；在 $AB=I$ 两边右乘 B^{-1}，得 $A=B^{-1}$，于是 $BA=I$。两个方向分别利用 A、B 的齐次方程只有零解，不需要提前使用转置的性质。 ■

由 $AB=I$ 也能看出 A 的每个右端至少有一个解：对任意 b，取 $x=Bb$ 就有 $Ax=ABb=b$。这只说明解存在，不能直接调用定理 3.14 的唯一解条件；上面对 B 的齐次方程的检查先推出 B 可逆，再由 $AB=I$ 得 $A=B^{-1}$，补足了另一侧乘积。矩形矩阵不满足这里使用的方阵判据，前面的 C,D 已给出单侧乘积不能自动补成双侧单位阵的例子。

本讲小结

- **形状与元素**：$(m\times n)(n\times p)=m\times p$，且 $(AB)_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}$；内维度必须相等。
- **乘积的四种视角**：行乘列给元素，AB 的列为 A 乘 B 的列，行按 A 的系数组合 B 的行，并且 $AB=\sum_k a_k r_k$。
- **运算律**：矩阵乘法满足结合律、左右分配律和单位矩阵律；交换律一般不成立，非零矩阵的乘积也可能为零。
- **分块**：分块乘法仍按行乘列规则组织；对应块必须可乘，参与相加的结果块必须同尺寸。
- **逆矩阵**：$A^{-1}A=AA^{-1}=I$，逆唯一；$Ax=b$ 的唯一解是 $A^{-1}b$，而 $(AB)^{-1}=B^{-1}A^{-1}$。
- **消元判据**：$[A\mid I]\to[I\mid A^{-1}]$ 可求逆；方阵可逆等价于每个右端唯一可解、齐次方程只有零解、消元得到 n 个主元。
- **单侧逆**：方阵中 $BA=I$ 或 $AB=I$ 任一式都推出另一式；矩形矩阵可只有一个方向的单侧逆。

下一讲将把消元操作记录成 $A=LU$。本讲的乘积次序和逆矩阵顺序会直接决定这个分解怎样写、怎样撤销。

练习

3.1 ★ 设 $A=\begin{bmatrix}1&0&2\\2&1&-1\end{bmatrix}$、$B=\begin{bmatrix}1&3\\2&0\\1&4\end{bmatrix}$。求 AB，说明尺寸，并写出第一行如何按列计算。

3.2 ★★ 对上题的 A,B，把 AB 写成三项外积之和，并检查结果的第二行。

3.3 ★ 取 $A=\begin{bmatrix}1&1\\0&2\end{bmatrix}$、$B=\begin{bmatrix}1&2\\3&0\end{bmatrix}$、$C=\begin{bmatrix}2\\1\end{bmatrix}$，分别计算 $(AB)C$ 与 $A(BC)$。

3.4 ★★ 对 $N=\begin{bmatrix}0&1\\0&0\end{bmatrix}$、$R=\begin{bmatrix}0&0\\1&0\end{bmatrix}$，计算 NR,RN,N^2。这三项分别说明什么？

3.5 ★★ 将 $K=\begin{bmatrix}1&2&0\\3&4&1\end{bmatrix}$ 按 $2+1$ 切开，$L=\begin{bmatrix}1&0\\0&1\\2&3\end{bmatrix}$ 按 $2+1$ 切开，用分块乘法求 KL。

3.6 ★★ 用 Gauss–Jordan 消元求 $M=\begin{bmatrix}0&1\\2&3\end{bmatrix}$ 的逆，并用一个乘积核对结果。

3.7 ★★ 取 $U=\begin{bmatrix}1&2&1\\0&1&3\\0&0&1\end{bmatrix}$，代入参数公式求逆，核对两个方向的乘积。

3.8 ★★ 判定 $S=\begin{bmatrix}1&2\\2&4\end{bmatrix}$ 是否可逆。给出一个非零齐次解，并指出一个无解右端。

3.9 ★★ 令 $C=\begin{bmatrix}1 & 0\end{bmatrix}$、$D=\begin{bmatrix}1\\0\end{bmatrix}$，比较 CD 与 DC。解释它们说明矩形单侧逆的什么限制。

3.10 ★★★ 证明：若 A 可逆且 $AB=AC$，则 $B=C$；再用一个奇异矩阵说明去掉可逆条件后结论会失败。

练习答案与提示

3.1 逐行乘列得

$$AB=\begin{bmatrix}3 & 11\\3 & 2\end{bmatrix}.$$

尺寸为 $(2\times 3)(3\times 2)=2\times 2$。第一行是 $1(1,3)+0(2,0)+2(1,4)=(3,11)$。

3.2 记 $A=[a_1\ a_2\ a_3]$，B 的三行为 $r_1=(1,3),r_2=(2,0),r_3=(1,4)$，则

$$AB=a_1r_1+a_2r_2+a_3r_3=\begin{bmatrix}1 & 3\\2 & 6\end{bmatrix}+\begin{bmatrix}0 & 0\\2 & 0\end{bmatrix}+\begin{bmatrix}2 & 8\\-1 & -4\end{bmatrix}.$$

三项相加给出上题结果，第二行逐项为 $(2,6)+(2,0)+(-1,-4)=(3,2)$。

3.3 先算 $AB=\begin{bmatrix}4 & 2\\6 & 0\end{bmatrix}$，所以 $(AB)C=(10,12)^T$。另一方面 $BC=(4,6)^T$，再乘 A 得 $A(BC)=(10,12)^T$。括号改变计算安排，未改变因子顺序。

3.4 $NR=\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$，$RN=\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$，$N^2=0$。前两项不同说明交换律失败；最后一项说明两个非零矩阵的乘积可以为零。

3.5 写 $K=[K_1\ K_2]$、$L=\begin{bmatrix}L_1\\L_2\end{bmatrix}$，其中 $K_1=\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$、$K_2=(0,1)^T$、$L_1=I_2$、$L_2=(2,3)$。因此

$$KL=K_1L_1+K_2L_2=\begin{bmatrix}1 & 2\\5 & 7\end{bmatrix}.$$

3.6 增广矩阵为

$$\left[\begin{array}{cc|cc}0 & 1 & 1 & 0\\2 & 3 & 0 & 1\end{array}\right]\xrightarrow{R_1\leftrightarrow R_2}\left[\begin{array}{cc|cc}2 & 3 & 0 & 1\\0 & 1 & 1 & 0\end{array}\right]\xrightarrow{R_1\leftarrow R_1-3R_2}\left[\begin{array}{cc|cc}2 & 0 & -3 & 1\\0 & 1 & 1 & 0\end{array}\right].$$

再作 $R_1\leftarrow R_1/2$，得到

$$\left[\begin{array}{cc|cc}1 & 0 & -3/2 & 1/2\\0 & 1 & 1 & 0\end{array}\right].$$

故 $M^{-1}=\begin{bmatrix}-3/2 & 1/2\\1 & 0\end{bmatrix}$；直接相乘 $MM^{-1}=I_2$。

3.7 代入 $a=2,b=1,c=3$，得

$$U^{-1}=\begin{bmatrix}1 & -2 & 5\\0 & 1 & -3\\0 & 0 & 1\end{bmatrix}.$$

例如 UU^{-1} 的三行分别为 $(1,0,0),(0,1,0),(0,0,1)$；右上角为 $5-6+1=0$。同样计算 $U^{-1}U$ 得 I_3，所以参数公式的两个方向都通过。

3.8 $S(2,-1)^T=0$，故由非零零输入方向判定 S 不可逆。其输出满足第二分量为第一分量的两倍，所以右端 $(1,0)^T$ 不可达，方程 $Sx=(1,0)^T$ 无解。

3.9 先把尺寸写开：C 将二维输入映到一维，D 再将一维数放回平面。$(1\times 2)(2\times 1)$ 得 $CD=[1]=I_1$；因此对每个标量 t，先用 D 再用 C 都能回到 t。

反向次序的尺寸是 $(2\times 1)(1\times 2)$，得到 $DC=\begin{bmatrix}1&0\\0&0\end{bmatrix}$。对一般输入 $(s,t)^T$，有 $DC(s,t)^T=(s,0)^T$；例如 $(0,1)^T$ 被送到零向量，第二坐标信息消失。因此 D 是 C 的右逆，C 的像虽覆盖整个一维目标空间，却不能恢复任意二维输入。矩形矩阵的单侧逆不能自动补成另一方向的单位阵。

3.10 在 $AB=AC$ 两边左乘 A^{-1}。结合律给出 $(A^{-1}A)B=(A^{-1}A)C$，即 $IB=IC$，所以 $B=C$。这里左乘是必要的，因为它让 A^{-1} 与等式左侧的 A 相邻。

去掉可逆条件时取 $A=\begin{bmatrix}1&0\\0&0\end{bmatrix}$、$B=\begin{bmatrix}0&0\\0&0\end{bmatrix}$、$C=\begin{bmatrix}0&0\\0&1\end{bmatrix}$。逐行相乘：$AB=\begin{bmatrix}0&0\\0&0\end{bmatrix}$，$AC=\begin{bmatrix}0&0\\0&0\end{bmatrix}$。左乘 A 只读取右侧矩阵的第一行，第二行无论是什么都会被变成零；而 B 的右下角为 0、C 的右下角为 1，故 $B\neq C$。

更一般地，若非零矩阵 A 有非零向量 u 满足 $Au=0$，取列向量 $v\neq 0$，并令 $B=0$、$C=uv^T$，则 $AC=(Au)v^T=0=AB$，但 $C\neq B$。消去公共左因子所需的关键条件，正是它没有会被压掉的非零方向；方阵可逆恰好保证这一点。

第 4 讲

LU 分解

Factorization into $A = LU$

对应材料：MIT 18.06SC Session 1.4（Strang 第 4 讲视频）课时摘要；习题 4.1、4.2 及其解答；参考 Strang《Introduction to Linear Algebra》§2.6（Elimination = Factorization: $A = LU$）与 §2.7（Transposes and Permutations）。

本讲目标

(1) 掌握乘积的逆与转置都要反序：$(AB)^{-1} = B^{-1}A^{-1}$，$(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$，并能说清为什么。

(2) 会把消元过程 $EA = U$ 改写成分解 $A = LU$，理解 L 是单位下三角矩阵、U 是上三角矩阵。

(3) 理解在不换行时，消元乘数 ℓ_{ij} 原样填进 L 的 (i, j) 位置，而 $E = L^{-1}$ 的元素会搅在一起。

(4) 会用 $A = LU$ 分两步（前代、回代）解 $Ax = b$，会写 $A = LDU$，知道分解在什么条件下唯一。

(5) 会估算消元的运算量：分解约 $\frac{1}{3}n^3$ 次，每个右端约 n^2 次，并能比较多个右端的主要运算量。

(6) 认识置换矩阵：共 $n!$ 个，$P^{-1} = P^{\mathsf{T}}$，乘积仍是置换矩阵；知道主元为零或相对过小时要考虑换行。

第 2 讲用初等矩阵 E_{ij} 描述了消元：一连串「第 i 行减去第 j 行的若干倍」把 A 变成上三角矩阵 U。第 3 讲讨论了逆矩阵和 Gauss–Jordan 方法。本讲把两者接起来：消元不只是一个求解过程，它还给出 A 本身的一个分解

$$A = LU,$$

其中 L 是下三角矩阵（lower），U 是上三角矩阵（upper）。这是数值线性代数里最常用的一个分解，计算机解线性方程组时做的几乎都是这件事。

要从 $EA = U$ 走到 $A = LU$，需要把左边的 E 移到右边，也就是求乘积的逆。所以我们先把乘积的逆和乘积的转置弄清楚。

4.1　乘积的逆与转置

4.1.1　乘积的逆：顺序要反过来

设 A、B 都是 $n\times n$ 可逆矩阵。一个很自然的猜测是 $(AB)^{-1}=A^{-1}B^{-1}$，可惜这是错的。正确的公式把顺序倒过来。

定理 4.1（乘积的逆）　若 A、B 都是 $n\times n$ 可逆矩阵，则 AB 可逆，并且

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

更一般地，若 $A_1,\dots,A_k$ 都可逆，则 $(A_1A_2\cdots A_k)^{-1}=A_k^{-1}\cdots A_2^{-1}A_1^{-1}$。

证明　只要验证 $B^{-1}A^{-1}$ 从左、右两边乘 AB 都得到 I。利用结合律，让中间相邻的两个因子先相遇：

$$(AB)(B^{-1}A^{-1})=A(BB^{-1})A^{-1}=AIA^{-1}=AA^{-1}=I,$$

$$(B^{-1}A^{-1})(AB)=B^{-1}(A^{-1}A)B=B^{-1}IB=B^{-1}B=I.$$

所以 $B^{-1}A^{-1}$ 就是 AB 的逆。证明中关键的一步是 B^{-1} 必须紧挨着 B，A^{-1} 必须紧挨着 A，这就决定了反序。k 个因子的情形对 k 归纳：$(A_1\cdots A_{k-1}A_k)^{-1}=A_k^{-1}(A_1\cdots A_{k-1})^{-1}$，再用归纳假设展开即可。∎

直观理解　早上先穿袜子（B），再穿鞋（A），合起来是 AB：矩阵乘向量时，离 x 最近的 B 先作用。晚上要撤销这件事，必须先脱鞋（A^{-1}），再脱袜子（B^{-1}），合起来是 $B^{-1}A^{-1}$。后做的操作要先撤销，这就是反序的全部原因。

■ **例 4.1　验证反序公式**　设 $A=\begin{bmatrix}1&2\\1&3\end{bmatrix}$，$B=\begin{bmatrix}2&1\\1&1\end{bmatrix}$。分别计算 $(AB)^{-1}$、$B^{-1}A^{-1}$ 和 $A^{-1}B^{-1}$，看哪一个对。

解　直接给出两个候选矩阵，并分别与原矩阵作双侧乘法核验：

$$A^{-1}=\begin{bmatrix}3&-2\\-1&1\end{bmatrix},\qquad B^{-1}=\begin{bmatrix}1&-1\\-1&2\end{bmatrix}.$$

乘积及其逆为

$$AB=\begin{bmatrix}1&2\\1&3\end{bmatrix}\begin{bmatrix}2&1\\1&1\end{bmatrix}=\begin{bmatrix}4&3\\5&4\end{bmatrix},\qquad (AB)^{-1}=\begin{bmatrix}4&-3\\-5&4\end{bmatrix}.$$

再按两种顺序相乘：

$$B^{-1}A^{-1}=\begin{bmatrix}1&-1\\-1&2\end{bmatrix}\begin{bmatrix}3&-2\\-1&1\end{bmatrix}$$

$$=\begin{bmatrix}4&-3\\-5&4\end{bmatrix},\qquad A^{-1}B^{-1}=\begin{bmatrix}3&-2\\-1&1\end{bmatrix}\begin{bmatrix}1&-1\\-1&2\end{bmatrix}$$

$$= \begin{bmatrix} 5 & -7 \\ -2 & 3 \end{bmatrix}.$$

$B^{-1}A^{-1}$ 与 $(AB)^{-1}$ 完全一致，而 $A^{-1}B^{-1}$ 连一个元素都对不上。顺序错了，结果就是另一个矩阵。 ♦

4.1.2 转置与乘积的转置

定义 4.2（转置） $m \times n$ 矩阵 A 的**转置**（transpose）A^{T} 是 $n \times m$ 矩阵，它的第 i 行就是 A 的第 i 列：

$$(A^{\mathsf{T}})_{ij} = A_{ji}.$$

满足 $A^{\mathsf{T}} = A$ 的方阵称为**对称矩阵**（symmetric matrix）。

例如 $\begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 4 \end{bmatrix}$。几何上看，转置就是沿主对角线把矩阵翻过去。转置和对称矩阵是第 5 讲的主角，这里只用到乘积的转置公式。

定理 4.3（乘积的转置） 若 A 是 $m \times n$ 矩阵，B 是 $n \times p$ 矩阵，则

$$(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}.$$

若 A 是可逆方阵，则 A^{T} 也可逆，并且 $(A^{\mathsf{T}})^{-1} = (A^{-1})^{\mathsf{T}}$。

思路　先看形状：$(AB)^{\mathsf{T}}$ 是 $p\times m$；$B^{\mathsf{T}}A^{\mathsf{T}}$ 是 $(p\times n)(n\times m)$，也是 $p\times m$；而 $A^{\mathsf{T}}B^{\mathsf{T}}$ 是 $(n\times m)(p\times n)$，当 $m \neq p$ 时根本乘不起来。形状已经告诉我们必须反序。接下来逐个元素比较即可。

证明　对 $1 \le i \le p$，$1 \le j \le m$，按矩阵乘法「行乘列」的定义，

$$\big((AB)^{\mathsf{T}}\big)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} A_{jk}B_{ki} = \sum_{k=1}^{n} (B^{\mathsf{T}})_{ik}(A^{\mathsf{T}})_{kj} = (B^{\mathsf{T}}A^{\mathsf{T}})_{ij}.$$

两边每个元素都相等，所以 $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$。

对第二个结论，把 $AA^{-1} = I$ 和 $A^{-1}A = I$ 两边转置，并用刚证明的公式和 $I^{\mathsf{T}} = I$：

$$(A^{-1})^{\mathsf{T}}A^{\mathsf{T}} = (AA^{-1})^{\mathsf{T}} = I, \qquad A^{\mathsf{T}}(A^{-1})^{\mathsf{T}} = (A^{-1}A)^{\mathsf{T}} = I.$$

所以 $(A^{-1})^{\mathsf{T}}$ 是 A^{T} 的逆。 ■

■ **例 4.2　形状说明一切**　设 $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{bmatrix}(2\times 3)$，$B = \begin{bmatrix} 2 & 1 \\ 0 & 1 \\ 1 & -1 \end{bmatrix}(3\times 2)$。验证 $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$，并说明为什么 $A^{\mathsf{T}}B^{\mathsf{T}}$ 不是答案。

解　先算乘积：

$$AB=\begin{bmatrix}1&2&0\\3&-1&4\end{bmatrix}\begin{bmatrix}2&1\\0&1\\1&-1\end{bmatrix}=\begin{bmatrix}2&3\\10&-2\end{bmatrix},\qquad (AB)^{\mathsf T}=\begin{bmatrix}2&10\\3&-2\end{bmatrix}.$$

再按反序相乘：

$$B^{\mathsf T}A^{\mathsf T}=\begin{bmatrix}2&0&1\\1&1&-1\end{bmatrix}\begin{bmatrix}1&3\\2&-1\\0&4\end{bmatrix}=\begin{bmatrix}2&10\\3&-2\end{bmatrix},$$

与 $(AB)^{\mathsf T}$ 一致。另一种顺序 $A^{\mathsf T}B^{\mathsf T}$ 是 $(3\times2)(2\times3)$，结果是 3×3 矩阵 $\begin{bmatrix}5&3&-2\\3&-1&3\\4&4&-4\end{bmatrix}$，连大小都不对。对于非方阵，形状就足以排除错误的顺序；对于方阵，形状检查失效，就要靠例 4.1 那样的计算或定理的证明。 ♦

想一想　设 A、B 都是 $n\times n$ 矩阵。已知 AB 可逆，能否推出 A 和 B 都可逆？

答　能。若 B 不可逆，按第 3 讲的判别法，存在 $x\neq0$ 使 $Bx=0$，于是 $(AB)x=A(Bx)=0$，与 AB 可逆矛盾，所以 B 可逆。再由 $A=(AB)B^{-1}$ 是两个可逆矩阵的乘积，用定理 4.1 知 A 可逆。「方阵」这个条件不能去掉：例 4.2 的 $AB=\begin{bmatrix}2&3\\10&-2\end{bmatrix}$ 有双侧逆 $\begin{bmatrix}1/17&3/34\\5/17&-1/17\end{bmatrix}$，但 A、B 都不是方阵，谈不上可逆。

易错提醒

- $(AB)^{-1}$ 一般不等于 $A^{-1}B^{-1}$，$(AB)^{\mathsf T}$ 一般不等于 $A^{\mathsf T}B^{\mathsf T}$。只有当 A、B 可交换（$AB=BA$）时，两种顺序才碰巧相同。
- 加法没有类似的公式：$(A+B)^{-1}$ 一般不等于 $A^{-1}+B^{-1}$。取 $A=B=I$，左边是 $(2I)^{-1}=\frac12 I$，右边是 $2I$。转置对加法倒是成立：$(A+B)^{\mathsf T}=A^{\mathsf T}+B^{\mathsf T}$。

4.2　从 $EA=U$ 到 $A=LU$

4.2.1　2×2 的情形

先看最简单的例子。对 $A=\begin{bmatrix}2&1\\8&7\end{bmatrix}$ 消元只需一步：第 2 行减去第 1 行的 $\ell_{21}=8/2=4$ 倍。写成矩阵就是

$$\underbrace{\begin{bmatrix}1&0\\-4&1\end{bmatrix}}_{E_{21}}\underbrace{\begin{bmatrix}2&1\\8&7\end{bmatrix}}_{A}=\underbrace{\begin{bmatrix}2&1\\0&3\end{bmatrix}}_{U}.$$

U 的对角元 2、3 就是两个主元。现在把 E_{21} 移到右边：两边左乘 E_{21}^{-1} 得 $A = E_{21}^{-1}U$。E_{21} 的作用是「第 2 行减去 4 倍第 1 行」，撤销它就是「第 2 行加回 4 倍第 1 行」，所以

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}}_{L=E_{21}^{-1}} \underbrace{\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}}_{U}.$$

乘数 4 出现在 L 的 $(2,1)$ 位置，符号由负变正。

定义 4.4（三角矩阵与 LU 分解）　方阵 L 称为**下三角矩阵**（lower triangular matrix），如果主对角线上方的元素全为零（$L_{ij} = 0$ 当 $i < j$）；若再要求对角元全为 1，称为**单位下三角矩阵**（unit lower triangular matrix）。主对角线下方全为零的方阵称为**上三角矩阵**（upper triangular matrix）。把方阵 A 写成

$$A = LU, \qquad L \text{ 单位下三角}, \quad U \text{ 上三角}$$

称为 A 的 *LU 分解*（LU factorization）。

L 的对角元是 1，U 的对角元是主元，两者看起来不对称。把主元从 U 中提出来，放进对角矩阵 D，剩下的上三角矩阵对角元也变成 1：

$$\begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix}.$$

这叫 *LDU 分解*（LDU factorization），其中 U 的每一行都除以了该行的主元。§4.4 会回到它。

■ **例 4.3　一般的 2×2 矩阵**　设 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 且 $a \neq 0$。求 A 的 LU 分解和 LDU 分解，并说明主元与 $ad - bc$ 的关系。

解　第一个主元是 a，乘数 $\ell_{21} = c/a$。第 2 行减去 $\frac{c}{a}$ 倍第 1 行，$(2,2)$ 元变成 $d - \frac{c}{a}b = \frac{ad-bc}{a}$，这是第二个主元。于是

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{c}{a} & 1 \end{bmatrix} \begin{bmatrix} a & b \\ 0 & \frac{ad-bc}{a} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{c}{a} & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & \frac{ad-bc}{a} \end{bmatrix} \begin{bmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{bmatrix}.$$

两个主元的乘积是 $a \cdot \frac{ad-bc}{a} = ad - bc$，正是第 3 讲 2×2 求逆公式里的分母。所以当 $a \neq 0$ 时，A 可逆当且仅当第二个主元不为零。用 $A = [2\ 1; 8\ 7]$ 检验：$\ell_{21} = 4$，第二个主元 $\frac{14-8}{2} = 3$，与上面一致。

若 $a = 0$ 而 $c \neq 0$，第一步就找不到主元，必须先交换两行。这种情形留到 §4.6。　♦

4.2.2　3×3 的情形

对 3×3 矩阵，不换行的消元有三步：先用 E_{21}、E_{31} 消去第 1 列主元下方的两个元素，再用 E_{32} 消去第 2 列主元下方的元素：

$$E_{32}E_{31}E_{21}\,A = U.$$

为了把 A 单独留在一边，依次左乘 E_{32}^{-1}、E_{31}^{-1}、E_{21}^{-1}。由定理 4.1，$(E_{32}E_{31}E_{21})^{-1} = E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}$，于是

$$A = \underbrace{E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}}_{L}\,U.$$

每个 E_{ij}^{-1} 都是单位下三角矩阵，只在 (i,j) 位置有一个非零的非对角元 $+\ell_{ij}$。下面的引理 4.7 会证明单位下三角矩阵的乘积仍是单位下三角矩阵，所以 L 配得上它的名字。

图 4.1 用后面例 4.4 的矩阵把两个方向画在一起：上面一行是消元，向右走；下面一行是撤销，向左走，先撤销最后一步。

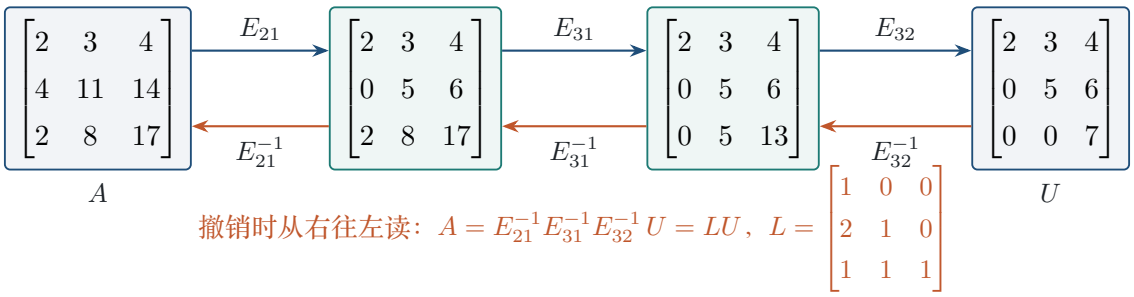


图 4.1　消元与撤销。蓝色箭头是消元：E_{21}、E_{31}、E_{32} 依次用乘数 2、1、1 消去主元下方的元素。橙色箭头是撤销：从 U 出发，先用 E_{32}^{-1} 撤销最后一步，最后用 E_{21}^{-1} 撤销第一步，回到 A。

核心思想　$A = LU$ 是消元过程的完整记录：U 记下消元的结果，L 记下每一步用了什么乘数。知道 L 和 U，就可以随时重建 A，也可以对任意右端重复消元，而不必再碰 A。

4.3　为什么乘子直接进入 L

4.3.1　E 与 L 的对比

$EA = U$ 和 $A = LU$ 含有同样的信息，因为 $L = E^{-1}$。但 Strang 在课上强调 L 比 E 好用得多。看一个例子：设消元时 $\ell_{21} = 2$、$\ell_{31} = 0$、$\ell_{32} = 5$（于是 $E_{31} = I$，可以省去）。

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{bmatrix},\quad E_{21}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad E_{32}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}.$$

先算 $E = E_{32}E_{21}$。左乘 E_{32} 的作用是「第 3 行减去 5 倍第 2 行」，作用在 E_{21} 上，第 3 行 $(0,0,1)$ 减去 $5\times(-2,1,0)$ 得 $(10,-5,1)$：

$$E = E_{32}E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 10 & -5 & 1 \end{bmatrix}.$$

再算 $L = E_{21}^{-1}E_{32}^{-1}$。左乘 E_{21}^{-1} 的作用是「第 2 行加上 2 倍第 1 行」，作用在 E_{32}^{-1} 上，只改动第 2 行，第 3 行原封不动：

$$L = E_{21}^{-1}E_{32}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}.$$

L 里只有两个乘数 2 和 5，各在各的位置；E 却多出了一个 10。

这个 10 从哪里来？E 回答的问题是「U 的各行怎样由 A 的各行组合出来」。记 A 的行为 $a_1^{\mathsf T}, a_2^{\mathsf T}, a_3^{\mathsf T}$。消元时第 3 行减去的是 5 倍**已经改过的**第 2 行 $a_2^{\mathsf T} - 2a_1^{\mathsf T}$，所以

$$u_3^{\mathsf T} = a_3^{\mathsf T} - 5\,(a_2^{\mathsf T} - 2a_1^{\mathsf T}) = 10\,a_1^{\mathsf T} - 5\,a_2^{\mathsf T} + a_3^{\mathsf T},$$

两步消元在这里搅在了一起。L 回答的是反方向的问题：「A 的各行怎样由 U 的各行组合出来」。撤销时第 2 行只要加回 $2u_1^{\mathsf T}$，第 3 行只要加回 $5u_2^{\mathsf T}$，各步互不干扰。表 4.1 把两者并排列出。

表 4.1　消元矩阵 $E = E_{32}E_{31}E_{21}$ 与 $L = E^{-1}$ 的对比（3×3，不换行）

	$E = E_{32}E_{31}E_{21}$	$L = E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}$
满足的等式	$EA = U$	$A = LU$
描述的方向	用 A 的行组合出 U 的行	用 U 的行组合出 A 的行
对角元	1	1
$(2,1)$ 元	$-\ell_{21}$	ℓ_{21}
$(3,2)$ 元	$-\ell_{32}$	ℓ_{32}
$(3,1)$ 元	$\ell_{32}\ell_{21} - \ell_{31}$	ℓ_{31}
上面例子	$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 10 & -5 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}$

表中 E 的 $(3,1)$ 元可以直接乘出来：$E_{31}E_{21}$ 的两个非对角元在不同的行，乘积就是把它们合在一起，$(2,1)$ 元为 $-\ell_{21}$，$(3,1)$ 元为 $-\ell_{31}$；再左乘 E_{32}，新的第 3 行是旧第 3 行减去 ℓ_{32} 倍旧第 2 行，$(3,1)$ 元变成 $-\ell_{31} - \ell_{32}(-\ell_{21}) = \ell_{32}\ell_{21} - \ell_{31}$。在上面的例子中它等于 $5\cdot 2 - 0 = 10$。

4.3.2　乘数原样填入 L

L 的这种整洁并非 3×3 的巧合，对任意 n 都成立。

定理 4.5（乘数原样进入 L）　设对 $n\times n$ 矩阵 A 逐列消元，过程中每个主元都非零且不需要换行：消第 j 列时，对每个 $i > j$，第 i 行减去第 j 行（主元行）的 ℓ_{ij} 倍。记消元结果为上三角矩阵 U，则 $A = LU$，其中

$$L = \begin{bmatrix} 1 & & & \\ \ell_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ \ell_{n1} & \cdots & \ell_{n,n-1} & 1 \end{bmatrix}.$$

思路　不去逐个乘 E_{ij}^{-1}，而是直接验证 A 的每一行等于 LU 的对应行。关键的观察是：第 j 行被当作主元行使用的时候，它已经是最终的样子，也就是 U 的第 j 行。

证明　记 A 的第 i 行为 $a_i^{\mathsf T}$，U 的第 i 行为 $u_i^{\mathsf T}$。

先说明第 j 行在消第 j 列时已是 u_j^{T}。消第 j 列及以后各列时，被修改的只有第 $j+1,\dots,n$ 行；第 j 行最后一次被修改发生在消第 $j-1$ 列的时候。所以第 j 行一旦当上主元行，就不再改变，它就是 U 的第 j 行 u_j^{T}。

再追踪第 i 行。它从 a_i^{T} 出发，消第 1 列时减去 $\ell_{i1}u_1^{\mathsf{T}}$，消第 2 列时减去 $\ell_{i2}u_2^{\mathsf{T}}$，……，消第 $i-1$ 列时减去 $\ell_{i,i-1}u_{i-1}^{\mathsf{T}}$，之后再也不变，最终成为 u_i^{T}：

$$a_i^{\mathsf{T}} - \ell_{i1}u_1^{\mathsf{T}} - \ell_{i2}u_2^{\mathsf{T}} - \cdots - \ell_{i,i-1}u_{i-1}^{\mathsf{T}} = u_i^{\mathsf{T}}.$$

移项得

$$a_i^{\mathsf{T}} = \ell_{i1}u_1^{\mathsf{T}} + \cdots + \ell_{i,i-1}u_{i-1}^{\mathsf{T}} + 1\cdot u_i^{\mathsf{T}}.$$

右边是 U 的各行以 $(\ell_{i1},\dots,\ell_{i,i-1},1,0,\dots,0)$ 为系数的组合，按「行乘矩阵」的观点，这正是 L 的第 i 行乘 U，即 LU 的第 i 行。对每个 i 都成立，所以 $A=LU$。 ■

证明中没有乘任何矩阵，也没有求逆。这就是实际计算中的做法：消元时把每个乘数记下来，L 就有了。

■ **例 4.4　完整的 3×3 分解**　求 $A=\begin{bmatrix}2&3&4\\4&11&14\\2&8&17\end{bmatrix}$ 的 LU 分解，并求出 $E=L^{-1}$。

解　*第 1 列*：主元 2，乘数 $\ell_{21}=4/2=2$，$\ell_{31}=2/2=1$。第 2 行减去 2 倍第 1 行得 $(0,5,6)$，第 3 行减去 1 倍第 1 行得 $(0,5,13)$。

第 2 列：主元 5，乘数 $\ell_{32}=5/5=1$。第 3 行减去第 2 行得 $(0,0,7)$。

三个主元是 $2,5,7$。由定理 4.5，把乘数 $2,1,1$ 填进去：

$$\underbrace{\begin{bmatrix}2&3&4\\4&11&14\\2&8&17\end{bmatrix}}_{A} = \underbrace{\begin{bmatrix}1&0&0\\2&1&0\\1&1&1\end{bmatrix}}_{L}\underbrace{\begin{bmatrix}2&3&4\\0&5&6\\0&0&7\end{bmatrix}}_{U}.$$

检验第 3 行：L 的第 3 行是 $(1,1,1)$，意思是 A 的第 3 行等于 U 三行之和，$(2,3,4)+(0,5,6)+(0,0,7)=(2,8,17)$，对。

再看 E。由表 4.1，它的 $(3,1)$ 元是 $\ell_{32}\ell_{21}-\ell_{31}=1\cdot 2-1=1$，所以

$$E=L^{-1}=\begin{bmatrix}1&0&0\\-2&1&0\\1&-1&1\end{bmatrix}.$$

可以验证 $EL=I$：例如 E 的第 3 行 $(1,-1,1)$ 乘 L 的第 1 列 $(1,2,1)$ 得 $1-2+1=0$。这里 E 的 $(3,1)$ 元是 1，而不是 $-\ell_{31}=-1$，单看 E 很难猜到乘数是多少。 ♦

■ **例 4.5　改编自习题 4.1**　设 $A=\begin{bmatrix}1&3&0\\2&4&0\\2&0&1\end{bmatrix}$。求矩阵 E 使 $EA=U$ 为上三角矩阵，再写出 $A=LU$。

解　*第 1 列*：主元 1，$\ell_{21}=2$，$\ell_{31}=2$。第 2 行变为 $(2,4,0)-2(1,3,0)=(0,-2,0)$，第 3 行变为 $(2,0,1)-2(1,3,0)=(0,-6,1)$。

第 2 列：主元 -2，$\ell_{32}=(-6)/(-2)=3$。第 3 行变为 $(0,-6,1)-3(0,-2,0)=(0,0,1)$。于是

$$U=\begin{bmatrix}1&3&0\\0&-2&0\\0&0&1\end{bmatrix},\qquad \text{主元}1,\ -2,\ 1.$$

求 $E=E_{32}E_{31}E_{21}$。$E_{31}E_{21}$ 只是把 -2、-2 放进第 1 列；再左乘 E_{32}（第 3 行减去 3 倍第 2 行），第 3 行 $(-2,0,1)$ 变为 $(-2,0,1)-3(-2,1,0)=(4,-3,1)$：

$$E=\begin{bmatrix}1&0&0\\-2&1&0\\4&-3&1\end{bmatrix}.$$

验算 EA 的第 3 行：$4(1,3,0)-3(2,4,0)+(2,0,1)=(0,0,1)$，正是 U 的第 3 行。$(3,1)$ 元 4 也符合公式 $\ell_{32}\ell_{21}-\ell_{31}=6-2$。

求 L 不必对 E 做 Gauss–Jordan，直接把乘数 $2,2,3$ 填进去：

$$\underbrace{\begin{bmatrix}1&3&0\\2&4&0\\2&0&1\end{bmatrix}}_{A}=\underbrace{\begin{bmatrix}1&0&0\\2&1&0\\2&3&1\end{bmatrix}}_{L}\underbrace{\begin{bmatrix}1&3&0\\0&-2&0\\0&0&1\end{bmatrix}}_{U}.$$

♦

4.3.3　列乘行：把 A 剥成秩一矩阵之和

第 3 讲介绍过矩阵乘法的「列乘行」观点：AB 等于 A 的第 k 列乘 B 的第 k 行，再对 k 求和。把它用在 LU 上，记 L 的第 k 列为 ℓ_k，U 的第 k 行为 $u_k^{\mathsf T}$，得到

$$A=LU=\ell_1u_1^{\mathsf T}+\ell_2u_2^{\mathsf T}+\cdots+\ell_nu_n^{\mathsf T}.$$

每一项都是一列乘一行，所有的行都是同一行的倍数，称为秩一矩阵。这个式子与消元一一对应：$\ell_1u_1^{\mathsf T}$ 的第 1 行和第 1 列与 A 完全相同，所以 $A-\ell_1u_1^{\mathsf T}$ 的第 1 行、第 1 列全是零，剩下的右下角就是第 1 步消元后的矩阵。之后每一步消元都再剥掉一个秩一矩阵。

■ **例 4.6**　**秩一分解**　对例 4.4 的 $A=LU$，写出三个秩一矩阵 $\ell_ku_k^{\mathsf T}$，并验证它们的和是 A。

解　L 的三列是 $(1,2,1)$、$(0,1,1)$、$(0,0,1)$，U 的三行是 $(2,3,4)$、$(0,5,6)$、$(0,0,7)$。

$$\ell_1u_1^{\mathsf T}=\begin{bmatrix}1\\2\\1\end{bmatrix}\begin{bmatrix}2&3&4\end{bmatrix}=\begin{bmatrix}2&3&4\\4&6&8\\2&3&4\end{bmatrix},\quad \ell_2u_2^{\mathsf T}=\begin{bmatrix}0&0&0\\0&5&6\\0&5&6\end{bmatrix},\quad \ell_3u_3^{\mathsf T}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&7\end{bmatrix}.$$

三者相加：$(2,2)$ 元 $6+5=11$，$(2,3)$ 元 $8+6=14$，$(3,2)$ 元 $3+5=8$，$(3,3)$ 元 $4+6+7=17$，其余元素一眼可见，和正是 A。

再看剥洋葱的过程：$A-\ell_1 u_1^{\mathsf{T}}=\begin{bmatrix}0&0&0\\0&5&6\\0&5&13\end{bmatrix}$，右下角 $\begin{bmatrix}5&6\\5&13\end{bmatrix}$ 就是消完第 1 列后剩下要处理的部分；再减去 $\ell_2 u_2^{\mathsf{T}}$ 只剩 $(3,3)$ 位置的 7，即最后一个主元。 ♦

想一想　对 3×3 矩阵，什么时候 E 的 $(3,1)$ 元恰好等于 $-\ell_{31}$，也就是 E 和 L 一样能直接读出乘数？

答　由表 4.1，E 的 $(3,1)$ 元是 $\ell_{32}\ell_{21}-\ell_{31}$。它等于 $-\ell_{31}$ 当且仅当 $\ell_{32}\ell_{21}=0$，即 $\ell_{21}=0$ 或 $\ell_{32}=0$。这时要么第 2 行在第 1 步没有被改动，要么第 2 行根本没有被用来改第 3 行，两步消元没有交织。例 4.5 中 $\ell_{21}=2$、$\ell_{32}=3$，所以 E 的 $(3,1)$ 元是 $6-2=4$，而不是 -2。

易错提醒

- 符号：E_{ij} 的 (i,j) 元是 $-\ell_{ij}$，L 的 (i,j) 元是 $+\ell_{ij}$。约定「第 i 行减去 ℓ_{ij} 倍第 j 行」，所以 ℓ_{ij} 是待消元素除以主元，不带负号；若需要加上某行，乘数就是负数。
- 乘数要用消元当时的矩阵计算，不能用 A 的原始元素。例 4.5 中 $\ell_{32}=(-6)/(-2)=3$ 用的是第 1 步之后的 -6 和 -2，而 A 的 $(3,2)$ 元是 0。

4.4　用 LU 解方程组，LDU 与唯一性

4.4.1　两个三角方程组

有了 $A=LU$，方程 $Ax=b$ 就是 $L(Ux)=b$。令 $c=Ux$，问题拆成两个三角方程组，各自都很好解。

算法 4.6（用 LU 解 $Ax=b$）

(1) 分解（只做一次，与 b 无关）：消元得 $A=LU$。

(2) 前代：解下三角方程组 $Lc=b$，从 c_1 往下依次求出 $c_1,c_2,\dots,c_n$。

(3) 回代：解上三角方程组 $Ux=c$，从 x_n 往上依次求出 $x_n,\dots,x_1$。

前代得到的 c 不是新东西：它正是第 2 讲对增广矩阵 $[A\ b]$ 消元后右端那一列。区别在于，LU 把「对 A 消元」和「对 b 做同样的事」分开了，换一个 b 时只需重做第 2、3 步。

■ **例 4.7　两个右端**　用例 4.4 的 $A=LU$ 解 $Ax=b$，其中 $b=(19,55,50)$；再解 $Ax=b'$，其中 $b'=(9,29,27)$。

解　前代 $Lc=b$：

$$\begin{bmatrix}1&0&0\\2&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix}=\begin{bmatrix}19\\55\\50\end{bmatrix}\quad\Longrightarrow\quad c_1=19,\quad c_2=55-2\cdot 19=17,\quad c_3=50-19-17=14.$$

回代 $Ux=c$：

$$\begin{bmatrix}2&3&4\\0&5&6\\0&0&7\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}19\\17\\14\end{bmatrix}\quad\Longrightarrow\quad x_3=\tfrac{14}{7}=2,\quad x_2=\tfrac{17-6\cdot2}{5}=1,\quad x_1=\tfrac{19-3\cdot1-4\cdot2}{2}=4.$$

所以 $x=(4,1,2)$。验算：$2\cdot4+3\cdot1+4\cdot2=19$，$4\cdot4+11\cdot1+14\cdot2=55$，$2\cdot4+8\cdot1+17\cdot2=50$。

换成 $b'=(9,29,27)$，L、U 不用重算。前代：$c_1=9$，$c_2=29-18=11$，$c_3=27-9-11=7$。回代：$x_3=1$，$x_2=(11-6)/5=1$，$x_1=(9-3-4)/2=1$。所以 $x=(1,1,1)$，这与 A 每行元素之和 $(9,29,27)$ 相符。 ♦

4.4.2 LDU 分解

在 $A=LU$ 中，L 的对角元是 1，U 的对角元是主元 $d_1,\dots,d_n$。若主元都不为零，把 U 的第 i 行除以 d_i，就把 U 写成 D 乘一个单位上三角矩阵：

$$A=LDU,\qquad D=\operatorname{diag}(d_1,\dots,d_n),$$

这里沿用 Strang 的记号，新的单位上三角矩阵仍记作 U。对例 4.4，U 的三行分别除以 2、5、7：

$$\begin{bmatrix}2&3&4\\0&5&6\\0&0&7\end{bmatrix}=\begin{bmatrix}2&0&0\\0&5&0\\0&0&7\end{bmatrix}\begin{bmatrix}1&\frac{3}{2}&2\\0&1&\frac{6}{5}\\0&0&1\end{bmatrix},$$

所以

$$A=\begin{bmatrix}1&0&0\\2&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}2&0&0\\0&5&0\\0&0&7\end{bmatrix}\begin{bmatrix}1&\frac{3}{2}&2\\0&1&\frac{6}{5}\\0&0&1\end{bmatrix}.$$

LDU 形式把「下」和「上」放在平等的地位，下面讨论对称矩阵时它更方便。

4.4.3 三角矩阵的运算与分解的唯一性

同一个 A 会不会有两种不同的 LU 分解？回答这个问题之前，先把三角矩阵的基本性质整理一下。

引理 4.7（三角矩阵的封闭性）　设 L、L' 是 $n\times n$ 下三角矩阵。

(1) LL' 是下三角矩阵，其对角元为 $(LL')_{ii}=L_{ii}L'_{ii}$。

(2) L 可逆当且仅当对角元全不为零；此时 L^{-1} 也是下三角矩阵，对角元为 $1/L_{ii}$。

(3) 特别地，单位下三角矩阵的乘积和逆仍是单位下三角矩阵。

(4) 把「下」换成「上」，以上结论同样成立。

证明　(1) $(LL')_{ij}=\sum_k L_{ik}L'_{kj}$。$L_{ik}\neq0$ 要求 $k\le i$，$L'_{kj}\neq0$ 要求 $k\ge j$。若 $i<j$，没有 k 同时满足两者，和为零；若 $i=j$，只剩 $k=i$ 一项。

(2) 若对角元全不为零，对每个 j 用前代解 $Lx=e_j$（e_j 是 I 的第 j 列）：前 $j-1$ 个方程右端为零，依次得 $x_1=\cdots=x_{j-1}=0$，第 j 个方程给出 $x_j=1/L_{jj}$。把这些解排成矩阵 X，则 $LX=I$，X 是下三角的，对角元为 $1/L_{jj}$；对方阵，第 3 讲说明了单边的逆就是逆。反过来，若某个 $L_{jj}=0$，取最大的这样的 j，可以构造 $x\neq 0$ 使 $Lx=0$：令 $x_j=1$、$x_1=\cdots=x_{j-1}=0$，再由第 $j+1,\dots,n$ 个方程依次确定 $x_{j+1},\dots,x_n$（这些方程的对角元不为零）。第 1 到第 j 个方程自动成立，所以 L 不可逆。

(3) 由 (1)、(2)，对角元为 $1\cdot 1=1$ 和 $1/1=1$。(4) 对上三角矩阵取转置化为下三角，再用 $(AB)^{\mathsf T}=B^{\mathsf T}A^{\mathsf T}$ 与 $(A^{-1})^{\mathsf T}=(A^{\mathsf T})^{-1}$。 ■

定理 4.8（LU 分解的唯一性）　设可逆矩阵 $A=L_1U_1=L_2U_2$，其中 L_1、L_2 是单位下三角矩阵，U_1、U_2 是上三角矩阵。则 $L_1=L_2$，$U_1=U_2$。因此 $A=LDU$ 中的 L、D、U 也唯一。

证明　A 可逆，$U_1=L_1^{-1}A$ 是可逆矩阵的乘积，所以 U_1 可逆，同理 U_2 可逆。由 $L_1U_1=L_2U_2$，左乘 L_2^{-1}、右乘 U_1^{-1} 得

$$L_2^{-1}L_1=U_2U_1^{-1}.$$

由引理 4.7，左边是单位下三角矩阵，右边是上三角矩阵。既是下三角又是上三角的矩阵只能是对角矩阵，而左边对角元都是 1，所以两边都等于 I。于是 $L_1=L_2$，$U_1=U_2$。D 由 U 的对角元确定，单位上三角部分由 $D^{-1}U$ 确定，因而也唯一。 ■

唯一性说明，不管用什么办法得到 $A=LU$（手算消元、按列计算、调用软件），只要 L 是单位下三角矩阵，结果都一样，L 里的数一定就是消元乘数。

想一想　定理 4.8 要求 L 的对角元是 1。去掉这个要求，分解还唯一吗？A 不可逆时又如何？
答　都不唯一。对任意可逆对角矩阵 D，$A=LU=(LD)(D^{-1}U)$，LD 仍是下三角、$D^{-1}U$ 仍是上三角。例如 $D=\mathrm{diag}(2,3)$ 时 $\begin{bmatrix}2&1\\8&7\end{bmatrix}=\begin{bmatrix}2&0\\8&3\end{bmatrix}\begin{bmatrix}1&\frac12\\0&1\end{bmatrix}$，这是让 U 的对角元为 1 的另一种约定（Crout 分解）。对不可逆的 A，取 $\begin{bmatrix}0&0\\0&1\end{bmatrix}=\begin{bmatrix}1&0\\t&1\end{bmatrix}\begin{bmatrix}0&0\\0&1\end{bmatrix}$，对任意 t 都成立：L 的第 2 行乘 U 的第 1 列时 t 乘的是 0。所以「L 对角元为 1」和「A 可逆」两个条件都不能少。

■ **例 4.8　改编自习题 4.2：一个对称矩阵**　求 $A=\begin{bmatrix}a&a&a&a\\a&b&b&b\\a&b&c&c\\a&b&c&d\end{bmatrix}$ 的 L 和 U。在什么条件下 $A=LU$ 有四个（非零）主元？

解　*第 1 列*（设 $a \neq 0$）：第 2、3、4 行的首元都是 a，乘数 $\ell_{21} = \ell_{31} = \ell_{41} = 1$。各行减去第 1 行：

$$\begin{bmatrix} a & a & a & a \\ 0 & b-a & b-a & b-a \\ 0 & b-a & c-a & c-a \\ 0 & b-a & c-a & d-a \end{bmatrix}.$$

第 2 列（设 $b \neq a$）：$\ell_{32} = \ell_{42} = 1$，第 3、4 行减去第 2 行，得到第 3 行 $(0, 0, c-b, c-b)$，第 4 行 $(0, 0, c-b, d-b)$。

第 3 列（设 $c \neq b$）：$\ell_{43} = 1$，第 4 行减去第 3 行得 $(0, 0, 0, d-c)$。于是

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \qquad U = \begin{bmatrix} a & a & a & a \\ 0 & b-a & b-a & b-a \\ 0 & 0 & c-b & c-b \\ 0 & 0 & 0 & d-c \end{bmatrix}.$$

四个主元是 $a,\ b-a,\ c-b,\ d-c$，所以条件是 $a \neq 0$，$b \neq a$，$c \neq b$，$d \neq c$。取 $a, b, c, d = 2, 5, 7, 11$，主元是 $2, 3, 2, 4$，直接消元验证无误。

再看 LDU：U 的每一行除以主元后，第 i 行在对角线及其右边全是 1，这个单位上三角矩阵正是 L^{T}。所以

$$A = LDL^{\mathsf{T}}, \qquad D = \operatorname{diag}(a,\ b-a,\ c-b,\ d-c).$$

♦

前后联系　$U = DL^{\mathsf{T}}$ 不是巧合。若 A 对称且 $A = LDU$（U 单位上三角），转置得 $A = A^{\mathsf{T}} = U^{\mathsf{T}}DL^{\mathsf{T}}$，这也是「单位下三角 × 对角 × 单位上三角」。由定理 4.8 的唯一性，$U^{\mathsf{T}} = L$，即 $U = L^{\mathsf{T}}$。对称矩阵的分解 $A = LDL^{\mathsf{T}}$ 是第 5 讲的内容，例 4.8 是一个提前的例子。

注 4.1　什么时候可以不换行一直消下去？记 A_k 为 A 左上角的 $k \times k$ 子矩阵。消前 k 列时，左上角 $k \times k$ 部分的变化只依赖于它自己，所以对 A_k 单独消元得到的主元，正是 A 的前 k 个主元。由此可证：**前导子矩阵全可逆**，当且仅当不换行时有 n 个非零主元。例如 $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}$ 可逆（交换第 2、3 行后主元是 $1, 1, 1$），但 $A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ 不可逆：消完第 1 列后第 2 行变为 $(0, 0, 1)$，第二个主元位置是 0，所以 A 没有 LU 分解，必须先换行。

4.5　消元要花多少代价

实际问题中的矩阵动辄有成千上万行，消元由计算机完成。这时要问两件事：要存什么，要算多少次。

4.5.1　原地存储

由定理 4.5，消元时只需记下乘数。乘数 ℓ_{ij} 恰好可以存在刚被消成零的 (i, j) 位置上：那个零不必存，反正我们知道它是零。这样整个消元在 A 原来的存储空间里完成，结束时下三角部分

是 L（对角线上的 1 不必存），上三角部分是 U。A 本身和各个 E_{ij} 都不再需要。图 4.2 画出了 4×4 的情形。

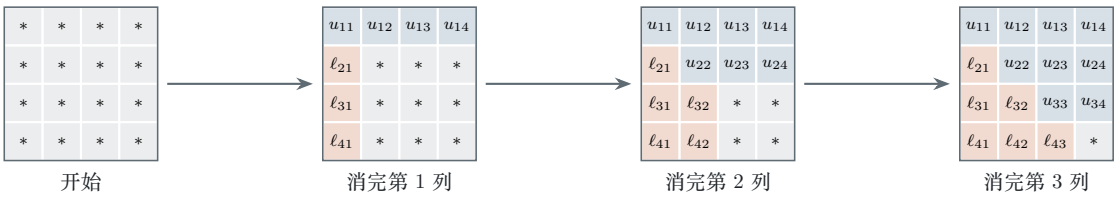


图 4.2　原地存储。灰色 $*$ 是仍在变化的活动块，橙色是已经记下的乘数（L 的下三角部分），蓝色是已经定型的 U 的行。每消一列，活动块缩小一行一列。

原地存储已经说明，计算机不必同时保存 A、每个消元矩阵和所有中间矩阵。每次把活动块中的一行减去另一行的倍数，并把乘数写入下三角位置，就能在同一块存储中逐步形成 L 与 U。

4.5.2　运算量的数量级

消去第 j 列时，活动块大约有 $n-j$ 行和 $n-j$ 个待更新的位置；一次乘法和减法的总数与 $(n-j)^2$ 同阶。因此整个分解的主要工作量可估为

$$1^2+2^2+\cdots+n^2=\sum_{i=1}^{n} i^2=\frac{n(n+1)(2n+1)}{6}\approx\frac{1}{3}n^3.$$

严格的计数会因是否把乘数计算、置零和加法分别计数而改变常数，但三次方数量级不变。分解完成后，对一个新的右端只需做一次前代和一次回代，工作量约为 n^2；若有许多右端，预先分解的成本可以被摊薄。

■ 例 4.9　**分解一次，求多个右端**　设 $n=1000$，需要求同一矩阵的 100 个右端。比较「先做一次 LU」与「每个右端重新消元」的主导运算量。

解　一次分解和 100 次三角求解约需

$$\frac{1000^3}{3}+100\cdot 1000^2\approx 3.333\times 10^8+1.0\times 10^8=4.333\times 10^8$$

个基本量级单位。若每个右端都重新消元，则约为

$$100\left(\frac{1000^3}{3}+1000^2\right)\approx 3.343\times 10^{10}.$$

后者约为前者的 77 倍。L,U 一旦保存，换右端只重做前代和回代，这正是分解法在多右端问题中的优势。 ♦

想一想　为什么不能把分解的 $\frac{1}{3}n^3$ 与每次求解的 n^2 混成同一个「单次复杂度」？

答　分解只对固定的 A 做一次，而求解成本随右端个数重复。若只有一个右端，二者都要付出；若有 r 个右端，总成本约为 $\frac{1}{3}n^3+rn^2$。把两项分开，才能判断何时值得先保存 L,U，也能避免把一次性的准备工作误报为每次查询都要承担的成本。

4.6　换行与置换矩阵

若当前主元为零而下方同列有非零元素，就必须先交换两行。浮点计算中，即使主元非零，若它相对同列其他候选值很小，除以它也可能放大舍入误差；实际计算常在当前列尚未处理的行中选绝对值较大的元素作主元。下面先用一个完整的分解展示这个选择，再单独看主元为零时怎样换行。

定义 4.9（置换矩阵）　把单位矩阵的行重新排列得到的矩阵称为**置换矩阵**（permutation matrix）。左乘置换矩阵会按同样的排列交换任意矩阵的行。

例如

$$P_{12}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}$$

把三行中的第 1、2 行交换。每一行、每一列恰有一个 1，其余位置为 0；所以 P 的列向量只是单位向量的排列。再交换一次就恢复原状，因而 $P^{-1}=P$。一般地，转置把行排列读成列排列，直接计算可得

$$P^{\mathsf T}P=I=PP^{\mathsf T},\qquad P^{-1}=P^{\mathsf T}.$$

■ **例 4.10　零主元时选较大的主元**　对 $A=\begin{bmatrix}0&2&1\\1&3&4\\2&1&0\end{bmatrix}$，第一列选绝对值较大的非零元素作主元，写出交换矩阵和第一列消元结果。

解　左上角为 0，第一列下方的候选值为 1 和 2，所以把第三行换到第一行，取

$$P_{13}=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix},\qquad P_{13}A=\begin{bmatrix}2&1&0\\1&3&4\\0&2&1\end{bmatrix}.$$

新的第一主元为 2。再作 $R_2\leftarrow R_2-\frac12 R_1$，得到

$$A^{(1)}=\begin{bmatrix}2&1&0\\0&\frac52&4\\0&2&1\end{bmatrix}.$$

第二列的候选主元是 $\frac52$ 和 2，所以不需再换行；该列乘数为 $\ell_{32}=\frac45$。本例展示了选较大主元的做法，下面先看它的完整 $PA=LU$，再看消元已经开始后才需要换行的情形。　♦

命题 4.10（置换矩阵的乘法性质）　对每个 n，恰有 $n!$ 个 $n\times n$ 置换矩阵。它们在矩阵乘法下封闭，单位矩阵是其中之一，每个置换矩阵的逆仍是置换矩阵，且 $P^{-1}=P^{\mathsf T}$。

证明　置换矩阵由 n 行的排列唯一确定，排列数是 $n!$。左乘 P 会重排矩阵的行；再左乘 Q 就按另

一个排列重排，合成仍是行的一个排列，因此 QP 仍是置换矩阵。单位矩阵不改变行序。逆操作是撤销该排列，也仍是一次排列；每列恰有一个 1，直接计算 $P^{\mathsf{T}}P = PP^{\mathsf{T}} = I$，故逆为转置。 ■

若先后左乘 Q_1、Q_2，作用结果是 Q_2Q_1A；矩阵乘法把先做的操作放在右边。不同的行交换一般不交换，因此累计置换必须按实际操作次序相乘。

■ **例 4.11　零主元时得到 $PA = LU$**　对前一例的 A，继续完成消元，写出 P, L, U 并核对等式 $PA = LU$。

解　第一列消元时 $\ell_{21} = \frac{1}{2}$、$\ell_{31} = 0$；第二列主元仍为 $\frac{5}{2}$，$\ell_{32} = \frac{4}{5}$。因此

$$PA = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 4 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & \frac{4}{5} & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & \frac{5}{2} & 4 \\ 0 & 0 & -\frac{11}{5} \end{bmatrix}.$$

右边乘积的第 2、3 行分别是 $\frac{1}{2}(2,1,0) + (0,\frac{5}{2},4) = (1,3,4)$ 和 $\frac{4}{5}(0,\frac{5}{2},4) + (0,0,-\frac{11}{5}) = (0,2,1)$，所以 $PA = LU$。这里的 P 左乘 A，表示对 A 的行作排列。 ♦

■ **例 4.12　消元中途换行**　对 $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}$，第一列消元后第二个主元为零。完成换行并写出 $PA = LU$。

解　第一列的乘数是 $\ell_{21} = 2$、$\ell_{31} = 1$，消元后

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

交换当前第 2、3 行。因为第一列乘数已记录在 L 中，也要把这两个记录 2、1 对调，得到

$$P = P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

于是

$$PA = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = LU.$$

若不交换 L 中已经记录的乘数，所得下三角因子就不能满足 $PA = LU$。 ♦

本讲小结

- **反序法则**：$(AB)^{-1}=B^{-1}A^{-1}$，$(AB)^{\mathsf{T}}=B^{\mathsf{T}}A^{\mathsf{T}}$；顺序由让相邻因子先抵消决定。
- **不换行的分解**：消元得到 $EA=U$ 时，$A=E^{-1}U=LU$；L 是单位下三角，U 是上三角。
- **乘数记录**：不换行时，消元乘数 ℓ_{ij} 直接填入 L_{ij}；它们在 E 中会因后续行操作混合。
- **三角求解**：$Ax=b$ 先解 $Lc=b$ 前代，再解 $Ux=c$ 回代；同一 A 的多个右端共用一次分解。
- **唯一性**：当 A 可逆且 L 规定为单位下三角时，$A=LU$ 唯一；不规定单位对角线会产生缩放自由度。
- **代价**：分解约需 $n^3/3$ 的量级，单个三角求解约需 n^2；原地存储把 L 和 U 放在同一矩阵中。
- **换行**：零主元或相对很小时可用行交换选取主元；$P^{-1}=P^{\mathsf{T}}$，累计置换左乘为 PA，求解时右端也同步置换，后续换行还要交换 L 中已记录的乘数。

下一讲将继续讨论转置、置换矩阵和实向量子空间。这里的 LDU 形式已经预告了对称情形的 LDL^{T}。

练习

4.1 ★ 对 $A=\begin{bmatrix}1&2\\1&3\end{bmatrix}$、$B=\begin{bmatrix}2&1\\1&1\end{bmatrix}$，求 AB、$(AB)^{-1}$、$B^{-1}A^{-1}$，并说明 $A^{-1}B^{-1}$ 为什么错误。

4.2 ★ 对 $A=\begin{bmatrix}1&2&0\\3&-1&4\end{bmatrix}$、$B=\begin{bmatrix}2&1\\0&1\\1&-1\end{bmatrix}$，验证 $(AB)^{\mathsf{T}}=B^{\mathsf{T}}A^{\mathsf{T}}$。

4.3 ★★ 对 $A=\begin{bmatrix}2&1\\8&7\end{bmatrix}$ 求 LU 分解与 LDU_0 分解，并列出两个主元。

4.4 ★★ 对 $A=\begin{bmatrix}2&3&4\\4&11&14\\2&8&17\end{bmatrix}$ 写出 L,U，并用它们解 $Ax=(19,55,50)^T$。

4.5 ★★ 对 $A=\begin{bmatrix}1&3&0\\2&4&0\\2&0&1\end{bmatrix}$ 求 E、$L=E^{-1}$ 和 $U=EA$，核对 $A=LU$。

4.6 ★★ 对上题的 L,U，写出 A 的三个列乘行外积之和，并指出每一项的矩阵尺寸。

4.7 ★★ 对 $A=\begin{bmatrix}a&a&a&a\\a&b&b&b\\a&b&c&c\\a&b&c&d\end{bmatrix}$，求 L,U，并给出有四个非零主元的充要条件。

4.8 ★ 说明 $n\times n$ 消元的主导运算量为何是 $\frac{1}{3}n^3$，并计算 $n=6$ 时 $\sum_{i=1}^{6} i^2$ 与 $6^3/3$。

4.9 ★★ 对 $A=\begin{bmatrix}0&2&1\\1&3&4\\2&1&0\end{bmatrix}$，取交换前两行的 P，求 PA 的前两步消元结果，并写出 $PA=LU$ 的 L,U。

4.10 ★★★ 证明单位下三角 L_1,L_2 的乘积仍为单位下三角，并用此证明单位下三角约束下 LU 分解的唯一性。

练习答案与提示

4.1 $AB=\begin{bmatrix}4&3\\5&4\end{bmatrix}$，$(AB)^{-1}=\begin{bmatrix}4&-3\\-5&4\end{bmatrix}$。由 $A^{-1}=\begin{bmatrix}3&-2\\-1&1\end{bmatrix}$、$B^{-1}=\begin{bmatrix}1&-1\\-1&2\end{bmatrix}$，得 $B^{-1}A^{-1}=(AB)^{-1}$，而 $A^{-1}B^{-1}=\begin{bmatrix}5&-7\\-2&3\end{bmatrix}$，因矩阵一般不交换而错误。

4.2 $AB=\begin{bmatrix}2&3\\10&-2\end{bmatrix}$，所以 $(AB)^{\mathsf T}=\begin{bmatrix}2&10\\3&-2\end{bmatrix}$。另一方面 $B^{\mathsf T}A^{\mathsf T}=\begin{bmatrix}2&0&1\\1&1&-1\end{bmatrix}\begin{bmatrix}1&3\\2&-1\\0&4\end{bmatrix}=\begin{bmatrix}2&10\\3&-2\end{bmatrix}$，两边相同。

4.3 $\ell_{21}=8/2=4$，故 $L=\begin{bmatrix}1&0\\4&1\end{bmatrix}$、$U=\begin{bmatrix}2&1\\0&3\end{bmatrix}$。把主元提出得 $D=\operatorname{diag}(2,3)$、单位上三角 $U_0=\begin{bmatrix}1&1/2\\0&1\end{bmatrix}$，于是 $A=LDU_0$；主元为 $2,3$。

4.4 $L=\begin{bmatrix}1&0&0\\2&1&0\\1&1&1\end{bmatrix}$、$U=\begin{bmatrix}2&3&4\\0&5&6\\0&0&7\end{bmatrix}$。前代给 $c=(19,17,14)^T$，回代给 $x_3=2$、$x_2=1$、$x_1=4$，所以 $x=(4,1,2)^T$。

4.5 $E=\begin{bmatrix}1&0&0\\-2&1&0\\4&-3&1\end{bmatrix}$，$U=EA=\begin{bmatrix}1&3&0\\0&-2&0\\0&0&1\end{bmatrix}$，$L=E^{-1}=\begin{bmatrix}1&0&0\\2&1&0\\2&3&1\end{bmatrix}$。逐行乘法给 $LU=A$；例如末行是 $2(1,3,0)+3(0,-2,0)+(0,0,1)=(2,0,1)$。

4.6 记 L 的列为 $\ell_1=(1,2,2)^T$、$\ell_2=(0,1,3)^T$、$\ell_3=(0,0,1)^T$，U 的行为 $u_1^T=(1,3,0)$、$u_2^T=(0,-2,0)$、$u_3^T=(0,0,1)$。则 $A=\ell_1u_1^T+\ell_2u_2^T+\ell_3u_3^T$，三项分别为 3×1 乘 1×3，结果均为 3×3。相加得 $\begin{bmatrix}1&3&0\\2&4&0\\2&0&1\end{bmatrix}$。

4.7 三个消元乘数均为 1，故 $L=\begin{bmatrix}1&0&0&0\\1&1&0&0\\1&1&1&0\\1&1&1&1\end{bmatrix}$，$U=\begin{bmatrix}a&a&a&a\\0&b-a&b-a&b-a\\0&0&c-b&c-b\\0&0&0&d-c\end{bmatrix}$。四个主元依次是 $a,b-a,c-b,d-c$，全部非零当且仅当 $a\neq0,b\neq a,c\neq b,d\neq c$。

4.8 $\sum_{i=1}^{6} i^2=1+4+9+16+25+36=91$，而 $6^3/3=72$。有限 n 的精确和是 $n(n+1)(2n+1)/6$；保留最高次项得到 $n^3/3$，所以这是渐近主导量级而非小 n 的精确计数。

4.9 取 $P = P_{12}$，则 $PA = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 2 & 1 \\ 2 & 1 & 0 \end{bmatrix}$。减去第一行的 2 倍后第三行变成 $(0, -5, -8)$；再减去第二行的 $-5/2$ 倍，末行成为 $(0, 0, -11/2)$。因此 $L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -5/2 & 1 \end{bmatrix}$、$U = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 2 & 1 \\ 0 & 0 & -11/2 \end{bmatrix}$，乘积末行为 $(2, 1, 0)$，从而 $PA = LU$。

4.10 设 $i > j$。$(L_1 L_2)_{ij} = \sum_k (L_1)_{ik}(L_2)_{kj}$；非零项必须同时有 $k \le i$ 和 $k \ge j$，所以当 $i < j$ 时为零。对角线上只有 $k = i = j$ 一项，等于 $1 \cdot 1 = 1$，故乘积仍是单位下三角。

若 $A = L_1 U_1 = L_2 U_2$ 且 A 可逆，三角因子均可逆。重排得 $L_2^{-1} L_1 = U_2 U_1^{-1}$。左侧既是单位下三角，右侧又是上三角，因此共同矩阵必须是对角矩阵；其对角线为 1，所以只能是 I。于是 $L_1 = L_2$，进而 $U_1 = U_2$。

第 5 讲

转置、置换与实向量子空间

Transposes, Permutations, and Subspaces

对应材料：MIT 18.06SC Session 1.5 课时摘要；习题 5.1–5.3 及解答；参考 Strang《Introduction to Linear Algebra》§2.7、§3.1。

本讲目标

(1) 会按元素定义转置，证明 $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$，并区分对称矩阵与斜对称矩阵。

(2) 理解置换矩阵左乘换行，能计算循环置换、逆和乘积。

(3) 掌握 $\mathbb{R}^n$、向量空间和子空间的定义，特别是对任意实数线性组合的封闭性。

(4) 能用零向量、加法封闭和数乘封闭判定一个集合是否为子空间。

(5) 理解矩阵的列空间是 $\mathbb{R}^m$ 的子空间，并把 $Ax = b$ 的可解性改写为 $b \in \mathrm{Col}(A)$。

第 4 讲中，转置只作为公式的一部分出现；置换矩阵也只是为了让主元离开零位置。现在把它们单独整理，并把「所有线性组合组成什么集合」变成严格的空间语言。先从一个可直接检查的 3×3 例子开始：

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \\ 2 & 5 & 1 \end{bmatrix}, \qquad A^{\mathsf{T}} = \begin{bmatrix} 1 & 3 & 2 \\ 2 & -1 & 5 \\ 0 & 4 & 1 \end{bmatrix}.$$

转置交换的是行和列的位置；它不改变元素，只改变我们读取元素的方向。后面会看到，转置把列空间的问题转成行空间的问题，也把下三角矩阵转成上三角矩阵。

5.1 转置与对称结构

定义 5.1（转置与对称矩阵） 设 A 为 $m \times n$ 矩阵，定义 A 的**转置**（transpose）A^{T} 为 $n \times m$ 矩阵，满足

$$(A^{\mathsf{T}})_{ij} = A_{ji}.$$

若 A 是方阵且 $A^{\mathsf{T}} = A$，称 A 为**对称矩阵**（symmetric matrix）；若 $A^{\mathsf{T}} = -A$，称为**斜对称矩阵**（skew-symmetric matrix）。

对称矩阵关于主对角线镜像相同，只需选择对角线及其一侧的元素。斜对称矩阵满足 $a_{ii} = -a_{ii}$，故对角元为零，镜像位置的非对角元互为相反数。零矩阵兼具两种性质。

例 5.1　按定义转置　求

$$R=\begin{bmatrix}1&3&4\\2&3&1\end{bmatrix}^{\mathsf T},\qquad S=\begin{bmatrix}2&-1&0\\-1&5&3\\0&3&4\end{bmatrix}^{\mathsf T},$$

并判断 S 是否对称。

解　交换行列得到 $R=\begin{bmatrix}1&2\\3&3\\4&1\end{bmatrix}$。矩阵 S 的 $(1,2)$ 与 $(2,1)$ 都为 -1，$(1,3)$ 与 $(3,1)$ 都为 0，$(2,3)$ 与 $(3,2)$ 都为 3，对角元不变，所以 $S^{\mathsf T}=S$，它是对称矩阵。每一对镜像位置都相等，而不是只检查一行就下结论。 ♦

定理 5.2（转置的基本运算）　对尺寸相容的矩阵，有

$$(A^{\mathsf T})^{\mathsf T}=A,\qquad (A+B)^{\mathsf T}=A^{\mathsf T}+B^{\mathsf T},\qquad (cA)^{\mathsf T}=cA^{\mathsf T},\qquad (AB)^{\mathsf T}=B^{\mathsf T}A^{\mathsf T}.$$

证明　前三式逐元素直接成立。对乘积，若 A 为 $m\times n$、B 为 $n\times p$，则

$$((AB)^{\mathsf T})_{ij}=(AB)_{ji}=\sum_{k=1}^{n}a_{jk}b_{ki}=\sum_{k=1}^{n}(B^{\mathsf T})_{ik}(A^{\mathsf T})_{kj}=(B^{\mathsf T}A^{\mathsf T})_{ij}.$$

两边尺寸都是 $p\times m$，且每个位置相同，所以矩阵相等。因子次序反过来是为了让原来求和中的 a_{jk} 与 b_{ki} 分别落在正确的行、列位置。 ■

例 5.2　乘积的转置　设 $A=\begin{bmatrix}1&2&0\\3&-1&4\end{bmatrix}$、$B=\begin{bmatrix}2&1\\0&1\\1&-1\end{bmatrix}$。验证 $(AB)^{\mathsf T}=B^{\mathsf T}A^{\mathsf T}$，并说明 $A^{\mathsf T}B^{\mathsf T}$ 为什么不对。

解　先算 $AB=\begin{bmatrix}2&3\\10&-2\end{bmatrix}$，所以 $(AB)^{\mathsf T}=\begin{bmatrix}2&10\\3&-2\end{bmatrix}$。另一方面

$$B^{\mathsf T}A^{\mathsf T}=\begin{bmatrix}2&0&1\\1&1&-1\end{bmatrix}\begin{bmatrix}1&3\\2&-1\\0&4\end{bmatrix}=\begin{bmatrix}2&10\\3&-2\end{bmatrix}.$$

错误顺序 $A^{\mathsf T}B^{\mathsf T}$ 的尺寸是 $(3\times 2)(2\times 3)$，虽然有定义，却产生 3×3 矩阵，尺寸已经与 $(AB)^{\mathsf T}$ 的 2×2 不同。方阵时尺寸检查不能排除错误顺序，仍需使用定理或具体核验。 ♦

命题 5.3（转置与逆）　若方阵 A 可逆，则 $A^{\mathsf T}$ 可逆，且

$$(A^{\mathsf T})^{-1}=(A^{-1})^{\mathsf T}.$$

证明　由 $AA^{-1}=I$ 与 $A^{-1}A=I$ 转置，得到 $(A^{-1})^{\mathsf T}A^{\mathsf T}=I$ 及 $A^{\mathsf T}(A^{-1})^{\mathsf T}=I$。这正是 $(A^{-1})^{\mathsf T}$ 为 $A^{\mathsf T}$ 双侧逆的定义。 ■

想一想　若 R 不是方阵，$R^{\mathsf T}R$ 与 $RR^{\mathsf T}$ 哪一个一定对称？

答　两个都在各自有定义时对称：$(R^{\mathsf T}R)^{\mathsf T}=R^{\mathsf T}(R^{\mathsf T})^{\mathsf T}=R^{\mathsf T}R$，同理 $(RR^{\mathsf T})^{\mathsf T}=RR^{\mathsf T}$。但它们尺寸不同：若 R 为 $m\times n$，前者为 $n\times n$，后者为 $m\times m$。不能因为都对称就把两个乘积当成同一个矩阵。

转置乘积必须反序　错误写法是 $(AB)^{\mathsf T}=A^{\mathsf T}B^{\mathsf T}$。例如上例中右侧会得到一个 3×3 矩阵，而左侧为 2×2；正确公式是 $(AB)^{\mathsf T}=B^{\mathsf T}A^{\mathsf T}$。即使 A,B 都是方阵，错误式也通常只在偶然交换时成立。

■ **例 5.3　四阶矩阵有多少个自由元素**　分别写出一般的 4×4 对称矩阵和斜对称矩阵，并数出可以独立选择的实数。

解　先选主对角线和上三角元素，其余位置由转置关系决定。对称矩阵的一般形式为

$$\begin{bmatrix} a & e & f & g \\ e & b & h & i \\ f & h & c & j \\ g & i & j & d \end{bmatrix}.$$

对角线上有 4 个自由元素，上三角有 $3+2+1=6$ 个，共 10 个。斜对称矩阵则为

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

对角线全是零，只剩 6 个自由元素。两个计数相加为 16，恰好是任意四阶矩阵的元素数；后面会把这个现象解释为一个分解。 ♦

一般的 n 阶对称矩阵有 $n+n(n-1)/2=n(n+1)/2$ 个自由元素；斜对称矩阵有 $n(n-1)/2$ 个。这里「自由」表示可以任意独立指定，不要求这些数非零。例如所有自由元素都取零，就得到两类集合共同包含的零矩阵。

■ **例 5.4　两个方向的对称乘积**　对 $R=\begin{bmatrix} 1 & 3 & 4 \\ 2 & 3 & 1 \end{bmatrix}$，分别计算 $R^{\mathsf T}R$ 与 $RR^{\mathsf T}$。

解　把列与列的点积填进第一个乘积，把行与行的点积填进第二个乘积，得到

$$R^{\mathsf T}R=\begin{bmatrix} 5 & 9 & 6 \\ 9 & 18 & 15 \\ 6 & 15 & 17 \end{bmatrix},\qquad RR^{\mathsf T}=\begin{bmatrix} 26 & 15 \\ 15 & 14 \end{bmatrix}.$$

例如第一个矩阵的 $(1,3)$ 元素为 $1\cdot 4+2\cdot 1=6$，其 $(3,1)$ 元素为 $4\cdot 1+1\cdot 2=6$；第二个矩阵的 $(1,2)$ 元素为 $1\cdot 2+3\cdot 3+4\cdot 1=15$。点积交换两个向量不改变结果，因此每一对镜像位置相等。

尺寸分别是 3×3 和 2×2。它们都对称，却记录了不同的点积表。前者比较三列，后者比较两行；在实际计算前先写下尺寸，可以避免把两种表混合。 ♦

转置运算只改变索引位置，但它同时联系了三个层次：元素级别是 $(A^{\mathsf{T}})_{ij}=a_{ji}$；乘法级别是反序公式；结构级别是对称性。证明结构时优先用反序公式，算具体数值时用行乘列，最后再对照尺寸。三种检查彼此补充。

5.2　置换矩阵与行的重排

把三行按 $1,2,3$ 改成 $3,1,2$，连续做三次会恢复原次序。这里需要记录的是每一条原行最后到了哪里。矩阵乘法提供一种统一的记录方法，也让多次重排可以合成为一个矩阵。

定义 5.4（置换矩阵）　把单位矩阵的行重新排列所得矩阵称为**置换矩阵**（permutation matrix）。左乘置换矩阵会按同一排列交换任意矩阵的行。

例如

$$P_{12}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}$$

交换第 1、2 行。每一行和每一列恰有一个 1，其余元素为 0。再交换一次恢复原矩阵，所以 $P_{12}^{-1}=P_{12}=P_{12}^{\mathsf{T}}$。

■ **例 5.5**　三循环置换　找一个 3×3 置换矩阵 $P\neq I$，使 $P^3=I$，并写出 P^2。

解　取循环 $1\to 2\to 3\to 1$ 的矩阵

$$P=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix},\qquad P^2=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}.$$

左乘 P 把原第 1 行放到第 2 行、第 2 行放到第 3 行、第 3 行放到第 1 行。连续三次每行回到原位，因此 $P^3=I$，而 $P\neq I$。 ♦

命题 5.5（置换矩阵的逆与乘积）　任意 $n\times n$ 置换矩阵 P 满足 $P^{\mathsf{T}}P=PP^{\mathsf{T}}=I$，故 $P^{-1}=P^{\mathsf{T}}$。两个同阶置换矩阵的乘积仍是置换矩阵；共有 $n!$ 个这样的矩阵。

证明　P 的列是单位向量的排列，列之间正交且长度为 1，所以 $P^{\mathsf{T}}P=I$；同理行的排列给出 $PP^{\mathsf{T}}=I$。先用 P 重排行，再用 Q 重排一次，合成仍然只是原行的一个排列，故 QP 仍为置换矩阵。n 行的排列数是 $n!$。 ■

■ **例 5.6　四阶置换的四次幂未必是单位矩阵**　构造一个 4×4 置换矩阵 Q，使 $Q^4\neq I$。

解　保留第一行，让其余三行作三循环，取

$$Q=\begin{bmatrix}1&0&0&0\\0&0&0&1\\0&1&0&0\\0&0&1&0\end{bmatrix}.$$

三次重排使后三行各自回到原位，第一行一直不动，所以 $Q^3=I$。再乘一次得到 $Q^4=Q\neq I$。矩阵的阶数是行列数，恢复原排列需要的次数由循环长度决定；两者不能混为一谈。 ♦

图 5.1 把三循环中原行的位置变化画出。箭头表示一次左乘 P 后的去向；沿箭头走三步，才会回到起点。

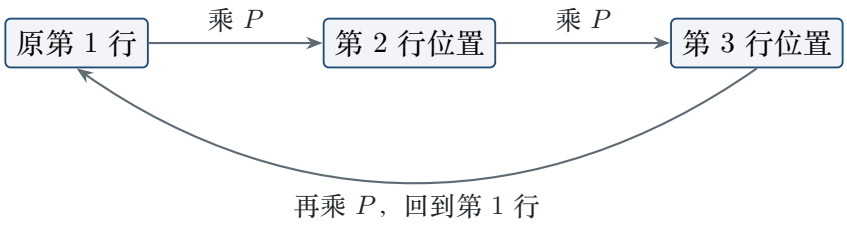


图 5.1　三循环中的每一条原行经过三次重排回到原位，因此三次幂为单位矩阵。

原第 2 行和原第 3 行也走同样的循环，只是出发点不同。于是所有行同时复位，矩阵才恢复为单位矩阵。若排列含多个互不相交的循环，需要让每个循环都复位。

5.3　实向量空间与子空间

向量 $(1,2)^T$ 的所有实数倍组成一条直线；如果只允许正倍数，就只剩一条射线。两者都能画在平面上，但后者乘以 -1 会离开集合。这说明，判断一个集合能否当作向量空间使用，必须检查允许哪些运算。

$\mathbb{R}^n$（n-dimensional real space）是所有含 n 个实分量的列向量组成的集合，向量按分量相加，数乘也按分量进行。它包含零向量，且对加法、数乘封闭；所以任取有限个向量和任意实数系数，所得线性组合仍在 $\mathbb{R}^n$ 中。

> **定义 5.6（子空间）**　向量空间 V 的非空子集 W 若本身也是向量空间，称为 V 的**子空间**（subspace）。对非空的 $W\subseteq\mathbb{R}^n$，这等价于：对任意 $u,v\in W$ 与任意 $c,d\in\mathbb{R}$，都有 $cu+dv\in W$。

证明　若 W 是子空间，它对加法和数乘封闭，故 $cu+dv$ 仍在 W。反过来，若对所有 c,d 都封闭，取 $c=d=0$ 得 $0\in W$；取 $(c,d)=(1,1)$ 得加法封闭，取 $d=0$ 得数乘封闭。继承自 $\mathbb{R}^n$ 的其余运算规则仍成立，因此 W 是子空间。 ■

■ **例 5.7　检验三类集合**　判断 $W_1=\{(x,2x):x\in\mathbb{R}\}$、$W_2=\{(x,2x+1):x\in\mathbb{R}\}$ 与 $W_3=\{(x,y):x\geq 0,\ y\geq 0\}$ 是否为 $\mathbb{R}^2$ 的子空间。

解　W_1 中向量是 $(x,2x)=x(1,2)$，所以任意线性组合仍是某个实数倍的 $(1,2)$，是子空间。W_2 不含零向量，因为 $(0,0)$ 不满足 $0=2\cdot 0+1$。W_3 对加法封闭，但 $(-1)(1,0)=(-1,0)$ 不在其中，故都不是子空间。图 5.2 对比了过原点直线、平移直线与第一象限。 ♦

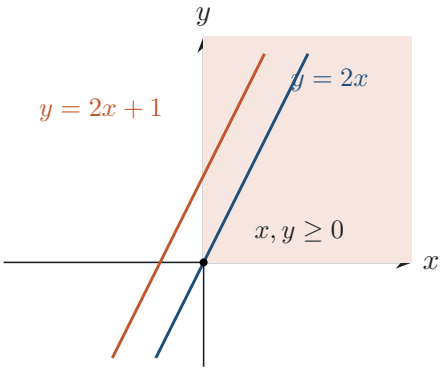


图 5.2　阴影是当前视窗内的第一象限。子空间中的直线必须经过原点；平移直线和第一象限分别因不含零向量、对负数乘法不封闭而失败。

想一想　为什么「过原点」只是必要条件，还不是子空间的完整判据？

答　过原点只保证零向量可能在集合中，还需检查加法和任意实数数乘。例如集合 $\{(x,y):y=x^2\}$ 经过原点，但 $(1,1)+(1,1)=(2,2)$ 不满足 $2=4$，所以不是子空间。直线 $y=2x$ 则对任意线性组合封闭，才是子空间。

命题 5.7（平面中的全部子空间）　$\mathbb{R}^2$ 的子空间恰好分为三类：$\{0\}$、过原点的直线、整个 $\mathbb{R}^2$。

证明　设 W 为子空间。若其中没有非零向量，则 $W=\{0\}$。若有非零向量 u，数乘封闭保证整条直线 $\{cu:c\in\mathbb{R}\}$ 都在 W 中。若 W 没有这条直线以外的向量，它就等于该直线。

否则取直线外的 v。任意平面向量都可以写成 $cu+dv$：矩阵 $[u\ v]$ 的两列不成比例，对应二元方程消元时有两个主元，因此每个右端都有唯一解。由封闭性，这些组合都在 W 中，故 $W=\mathbb{R}^2$。 ■

在 $\mathbb{R}^3$ 中对应的四类是：$\{0\}$、过原点的直线、过原点的平面、整个 $\mathbb{R}^3$。取一个非零向量得到直线；再取直线外的向量，其组合构成平面；若再加入平面外的向量，三个方向可以产生每个三维向量。这里的完整分类将在第 9 讲用维数加以论证。

检查子空间的顺序　先检查零向量；失败就可以停止。零向量在集合中，再对任意两个元素写出一般形式，检查它们的任意线性组合。要证明封闭，有限个试算不够；要否定封闭，一个反例就够。

■ **例 5.8　一个平面与两个坐标平面的并**　判断 $U=\{(x,y,z)^T:x+2y-z=0\}$ 和 $V=\{(x,y,z)^T:x=0\ \text{或}\ y=0\}$ 是否为子空间。

解　对 $u, v \in U$，分别有 $u_1 + 2u_2 - u_3 = 0$、$v_1 + 2v_2 - v_3 = 0$。于是对任意实数 c, d，

$$(cu_1 + dv_1) + 2(cu_2 + dv_2) - (cu_3 + dv_3) = c \cdot 0 + d \cdot 0 = 0.$$

零向量也满足方程，因此 U 是子空间。写成参数形式则为 $(x, y, z)^T = x(1, 0, 1)^T + y(0, 1, 2)^T$，可看出它是过原点的平面。

V 包含零向量，且数乘不改变某个坐标为零的事实。但 $(1, 0, 0)^T$ 和 $(0, 1, 0)^T$ 都在 V 中，其和 $(1, 1, 0)^T$ 不在 V 中。它对加法不封闭，因而不是子空间。 ♦

5.4　张成、列空间与线性方程

矩阵 $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 1 & 3 \end{bmatrix}$ 有两个列向量，却只能产生一条直线上的输出：第二列是第一列的 3 倍，无论怎样选择输入，输出方向都没有增加。列的数量与能够到达的向量集合是两回事，需要把后者单独命名。

设矩阵 $A = [a_1 \ \cdots \ a_n]$ 的每一列都在 $\mathbb{R}^m$。所有列的线性组合

$$c_1 a_1 + \cdots + c_n a_n = A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

组成 A 的**列空间**（column space），记为 $\mathrm{Col}(A)$。它是 $\mathbb{R}^m$ 的子空间：零组合给出零向量，两个列组合相加仍是列组合，乘以实数也仍是列组合。

> **命题 5.8（列空间与可达右端）**　方程 $Ax = b$ 有解，当且仅当 $b \in \mathrm{Col}(A)$。若 A 的列空间等于 $\mathbb{R}^m$，则每个 m 维右端都可达；若列空间只是一个更小的子空间，则只有其中的右端有解。

证明　写 $x = (x_1, \ldots, x_n)^T$，则 $Ax = x_1 a_1 + \cdots + x_n a_n$，所以任意解给出 b 的列组合。反过来，若 b 是列组合 $b = c_1 a_1 + \cdots + c_n a_n$，取 $x = (c_1, \ldots, c_n)^T$ 即得 $Ax = b$。两边互相推出，条件等价。 ■

■ **例 5.9　三维中的一个列空间**　设 $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 1 & 3 \end{bmatrix} = [a_1 \ a_2]$。描述 $\mathrm{Col}(A)$，并判定 $b = (4, 8, 4)^T$ 和 $c = (1, 2, 3)^T$ 是否可达。

解　$a_2 = 3a_1$，所以所有列组合都是 $(s, 2s, s)^T = s(1, 2, 1)^T$，列空间是 $\mathbb{R}^3$ 中过原点的一条直线。$b = 4a_1$ 可达，取 $x = (4, 0)^T$；而 c 的第三分量为 3、第一分量为 1，不满足第三分量等于第一分量，故不在列空间中，方程 $Ax = c$ 无解。 ♦

■ **例 5.10　原课程中的平面列空间**　设 $C=\begin{bmatrix}3&1\\3&2\\1&4\end{bmatrix}$。描述其列空间，判断 $(5,4,-2)^T$ 与 $(1,0,0)^T$ 是否属于其中。

解　设输出为 $(p,q,r)^T$，则列组合给出 $p=3s+t$、$q=3s+2t$、$r=s+4t$。前两式相减得 $t=q-p$，继而 $s=(2p-q)/3$；代入第三式得

$$10p-11q+3r=0.$$

反过来，若这个等式成立，以上的 s,t 确实产生三个指定分量。因此列空间恰是这个过原点的平面，不能只由一次消元推出必要条件就结束。

对 $(5,4,-2)^T$，取 $s=2,t=-1$，有 $2(3,3,1)^T-(1,2,4)^T=(5,4,-2)^T$，所以可达。对 $(1,0,0)^T$，平面方程左侧为 10，不等于零，所以不可达。两个不成比例的列张成一个平面，仍然无法覆盖整个三维空间。 ♦

直观理解　输入 $(s,t)^T$ 允许任意取值，输出却始终受一个齐次线性等式约束。这个约束描述了列空间的位置。把右端向量放进约束中检验，就能先判断方程是否可能有解；确认可达后，再求组合系数。

5.5　矩阵组成的向量空间

两个 2×2 矩阵相加仍是 2×2 矩阵，乘以实数也不改变尺寸。因此向量空间中的元素可以是整个矩阵。这里的加法与数乘仍按元素进行；矩阵乘法不属于子空间判据需要检查的运算。

定义 5.9（矩阵空间）　所有 $m\times n$ 实矩阵组成的集合记为 $M_{m\times n}(\mathbb{R})$，在通常的矩阵加法和实数数乘下是向量空间。方阵情形简记为 $M_n(\mathbb{R})$。

例如零矩阵是它的零元素，$-A$ 是 A 的加法逆元。矩阵尺寸必须固定：一个 2×2 矩阵与一个 2×3 矩阵不能按通常定义相加，所以所有不同尺寸的矩阵放在一起不构成这里讨论的矩阵空间。

命题 5.10（对称与斜对称矩阵的子空间）　在 $M_n(\mathbb{R})$ 中，对称矩阵的集合和斜对称矩阵的集合分别都是子空间。

证明　零矩阵同时满足 $0^{\mathsf T}=0$ 和 $0^{\mathsf T}=-0$。若 $A^{\mathsf T}=A$、$B^{\mathsf T}=B$，则对任意实数 c,d，

$$(cA+dB)^{\mathsf T}=cA^{\mathsf T}+dB^{\mathsf T}=cA+dB.$$

因此对称矩阵对线性组合封闭。若 $A^{\mathsf T}=-A$、$B^{\mathsf T}=-B$，同样计算得到

$$(cA+dB)^{\mathsf T}=-cA-dB=-(cA+dB),$$

所以斜对称矩阵也对线性组合封闭。两类集合都非空，故都是子空间。 ■

例 5.11　非对称矩阵为什么不构成子空间　设 $N=\{A\in M_2(\mathbb{R}):A^{\mathsf{T}}\neq A\}$。给出一个违反加法封闭性的反例。

解　选取两块互补的矩阵

$$A=\begin{bmatrix}1&1\\0&0\end{bmatrix},\qquad B=\begin{bmatrix}0&0\\1&1\end{bmatrix}.$$

两者的非对角元素不相等，因而都在 N 中。但是 $A+B=\begin{bmatrix}1&1\\1&1\end{bmatrix}$ 对称，不在 N 中。加法封闭失败已经足以否定子空间。此外，N 不含零矩阵，数乘 0 也会离开集合。 ♦

补集通常不继承子空间性质　对称矩阵构成子空间，不能据此断言其补集也构成子空间。上例中两个非对称矩阵相加就变成对称矩阵。检查集合时，要把定义中的每一项直接用于该集合。

命题 5.11（对称部分与斜对称部分）　任意实方阵 A 都能唯一写成 $A=S+K$，其中 $S^{\mathsf{T}}=S$、$K^{\mathsf{T}}=-K$。具体公式为

$$S=\frac{A+A^{\mathsf{T}}}{2},\qquad K=\frac{A-A^{\mathsf{T}}}{2}.$$

证明　先验证存在性。将上式转置，得到 $S^{\mathsf{T}}=(A^{\mathsf{T}}+A)/2=S$，以及 $K^{\mathsf{T}}=(A^{\mathsf{T}}-A)/2=-K$；二者相加为 A。

再验证唯一性。若 $A=S_1+K_1=S_2+K_2$，则 $D=S_1-S_2=K_2-K_1$。左式说明 D 对称，右式说明 D 斜对称。因此 $D=D^{\mathsf{T}}=-D$，从而 $2D=0$，即 $D=0$。所以 $S_1=S_2$、$K_1=K_2$。 ■

例 5.12　把一个矩阵分成两部分　把 $A=\begin{bmatrix}1&5\\1&3\end{bmatrix}$ 写成一个对称矩阵与一个斜对称矩阵之和。

解　直接应用公式，但仍要检查两部分的性质：

$$S=\frac{1}{2}\begin{bmatrix}2&6\\6&6\end{bmatrix}=\begin{bmatrix}1&3\\3&3\end{bmatrix},\qquad K=\frac{1}{2}\begin{bmatrix}0&4\\-4&0\end{bmatrix}=\begin{bmatrix}0&2\\-2&0\end{bmatrix}.$$

S 的镜像位置相等，K 的镜像位置互为相反数且对角线为零。相加恢复 A。每对非对角元素的平均数进入 S，差的一半进入 K；这种分配同时解释了公式和唯一性。 ♦

想一想　两个对称矩阵的乘积一定对称吗？这是否影响它们构成子空间？

答　不一定。取 $A=\begin{bmatrix}1&0\\0&2\end{bmatrix}$、$B=\begin{bmatrix}0&1\\1&0\end{bmatrix}$，则 $AB=\begin{bmatrix}0&1\\2&0\end{bmatrix}$ 不对称。由 $(AB)^{\mathsf{T}}=BA$ 可知，对称矩阵的乘积对称当且仅当 $AB=BA$。这不影响子空间结论，因为子空间只检查加法与实数数乘。

5.6　把判据用于新的集合

集合 $W=\{(x,y,z)^T:x+y+z=0\}$ 的方程右边为零，而平行集合 $W_1=\{(x,y,z)^T:x+y+z=1\}$ 的右边为一。这一个常数的变化决定了封闭性。把定义直接代入，可以看清每一个条件分别起什么作用。

> **命题 5.12（齐次线性条件的封闭性）**　设 A 为实 $m\times n$ 矩阵。集合 $Z=\{x\in\mathbb{R}^n:Ax=0\}$ 是 $\mathbb{R}^n$ 的子空间。对固定非零向量 b，集合 $Z_b=\{x:Ax=b\}$ 若非空，则不是子空间。

证明　零向量满足 $A0=0$，所以 Z 非空。若 $Au=0$、$Av=0$，由矩阵乘法的分配律，对任意实数 c,d 都有

$$A(cu+dv)=cAu+dAv=c0+d0=0.$$

因此 Z 对任意线性组合封闭。对第二个集合，由 $b\neq0$，零向量不满足 $A0=b$，所以它不含零向量，不可能是子空间。 ■

非空条件使第二个集合确实描述一组解，不过即使它为空，也仍不是子空间。齐次方程总有零解，因此第一部分没有额外的存在性问题。第 6 讲会把 Z 命名为零空间；这里先用本讲的判据验证其结构。

■ **例 5.13　同一个左端的两个解集**　描述 $x+y+z=0$ 的所有解，再说明 $x+y+z=1$ 的所有解与它有什么关系。

解　把 $y=s,z=t$ 作为任意实数，齐次方程给出 $x=-s-t$，所以全部解为

$$(x,y,z)^T=s(-1,1,0)^T+t(-1,0,1)^T.$$

两个固定向量的所有组合构成过原点的平面。非齐次方程则给出 $x=1-s-t$，全部解为

$$(x,y,z)^T=(1,0,0)^T+s(-1,1,0)^T+t(-1,0,1)^T.$$

它是前一个平面的平移。每个向量都满足坐标和为一，特别是零向量不在其中。两个解相减，坐标和变为零，因此差向量回到齐次解集；两个解相加，坐标和为二，一般离开原非齐次解集。 ♦

同一个矩阵的列空间和齐次解集属于不同的周围空间。若 A 为 $m\times n$，列空间由 m 维列向量的组合构成，位于 $\mathbb{R}^m$；解向量有 n 个分量，齐次解集位于 $\mathbb{R}^n$。即使 $m=n$，两者也按不同的定义产生，不能直接认作同一个集合。

■ **例 5.14　包含关系与列的增删**　设 $u=(1,2,1)^T$，比较矩阵 $A=[u\ 3u]$、$B=[u]$、$C=[u\ 0\ -2u]$ 的列空间。

解　A 的输出为 $su+3tu=(s+3t)u$。系数 $s+3t$ 可以取任意实数，因为可取 $t=0$，所以列空间恰是所有实数倍的 u。B 的输出直接是 su，得到同一个集合。C 的输出为 $su+t0-2ru=(s-2r)u$，系数同样任意。因此三者的列空间相等，尽管列数分别为二、一、三。

删除一个能由其余列组合出来的列，不会减少列空间；增加这样的列，也不会扩大列空间。这个判断需要两个包含方向：新列的每个组合都能改写成旧列组合；旧列仍保留，因此每个旧组合仍能产生。 ♦

证明两个描述相同　把一个集合写成方程形式，再写成参数形式时，需检查两个方向。先证明每个参数向量都满足方程，再证明每个满足方程的向量都有相应参数。只做第一步，可能遗漏一部分解；只做第二步，可能把不符合条件的参数也收入集合。

子空间的交仍是子空间，因为任意线性组合会同时保留两组限制。并集的情况不同：两个向量可能分别来自不同的子空间，它们的和未必属于任何一个。这解释了坐标轴并集的失败，也给出了练习最后一题的思路。这里的「同时满足」对应交，「满足其中任一组」对应并，两种逻辑不能交换。

最后回到矩阵空间。对称条件 $a_{ij}-a_{ji}=0$ 是一组关于元素的齐次线性等式；斜对称条件 $a_{ij}+a_{ji}=0$ 也是如此。它们与上面的齐次方程采用相同的封闭性机制。非对称条件要求至少一对元素不相等，这个不等关系在相加时可能抵消，所以无法用相同论证得到封闭性。

■ **例 5.15**　**只检查加法会遗漏什么**　设 W 是所有坐标均为整数的二维向量组成的集合。它包含零向量，且任意两个元素相加仍在其中。它是否为实向量空间 $\mathbb{R}^2$ 的子空间？

解　整数坐标相加仍为整数，所以加法确实封闭。可是 $(1,0)^T\in W$，乘以实数 $1/2$ 得 $(1/2,0)^T\notin W$。因此实数数乘不封闭，不能成为 $\mathbb{R}^2$ 的子空间。问题中的系数范围有实质作用：要求对所有实数成立，不能只验证整数系数。相反，整条直线 $\{t(1,0)^T:t\in\mathbb{R}\}$ 允许全部实数倍，便满足这一条件。 ♦

空集也提醒我们注意量词。对空集中的任意两个元素作线性组合，这句话没有可选元素，形式上不会遇到反例；但空集没有零向量，因此不是子空间。定义中要求非空，正是为了排除这种情形。单元素集合只有在该元素为零时才可能是子空间：若元素为非零 u，数乘零要求 $0u=0$ 也在集合中，与集合只有 u 矛盾。

■ **例 5.16**　**同一条直线的不同写法**　比较 $L_1=\{t(1,2)^T:t\in\mathbb{R}\}$、$L_2=\{s(3,6)^T:s\in\mathbb{R}\}$ 与 $L_3=\{r(-1,-2)^T:r\in\mathbb{R}\}$。

解　L_2 中任意向量为 $3s(1,2)^T$，所以属于 L_1；反过来，给定 t，取 $s=t/3$ 就产生 $t(1,2)^T$。因此 $L_1=L_2$。对 L_3，取 $r=-t$ 即产生同样的向量，所以三者相等。更换非零方向向量的长度或符号，不改变所有实数倍组成的直线。

若用零向量代替方向向量，集合则退化为 $\{0\}$。因此「非零方向」是直线描述中的必要条件。零空间仍是子空间，却不能称为一条直线。集合相等讨论的是全部元素，不能只比较写在花括号里的一个生成向量。 ♦

判定方程是否可解时，要区分一个右端与所有右端。找到某个 b 的解，只能说明 $b\in\mathrm{Col}(A)$。

要证明每个右端都有解，需有 $\mathrm{Col}(A)=\mathbb{R}^m$。例如只有一个非零列的三行矩阵仍可解许多右端，但这些右端都在同一条直线上。一个不可达的右端足以否定「每个右端都可解」，却不能说明其他右端也无解。

向量的坐标个数同样不能省略。$\{0\}$ 写法会随周围空间改变：在 $\mathbb{R}^2$ 中零向量有两个分量，在 $\mathbb{R}^3$ 中有三个分量。两者都满足子空间判据，但元素的尺寸不同。矩阵空间中的零矩阵也必须与其他元素同尺寸，否则加法无定义。每次使用子空间判据前，先写出周围空间，再验证零元素和封闭性，可以让论证中的对象保持一致。

上述每一种验证都只使用加法和数乘，因此适用于向量集合，也适用于固定尺寸的矩阵集合。

判断过程还可以反向用于设计集合。若希望一个集合稳定地保留线性组合，就用齐次线性方程描述它，或直接取固定向量的全部实系数组合。前一种写法方便检验成员资格，后一种写法方便构造成员。两种写法表达同一个集合时，它们的优势互补：方程告诉我们哪些向量被排除，参数告诉我们允许的向量怎样产生。

实际书写中，一个完整结论最好包含三个部分：集合属于哪个周围空间；它是否为子空间；支持结论的闭包计算或具体反例。例如只写「这是平面」不足以完成判定，因为平面可能经过原点，也可能平移后离开原点；只写「包含零向量」也不足够，因为坐标轴的并仍会在加法处失败。把失败的具体运算写出来，读者就可以直接核查结论。

本讲小结

- **转置**交换行列：$(A^{\mathsf{T}})_{ij}=a_{ji}$，且 $(AB)^{\mathsf{T}}=B^{\mathsf{T}}A^{\mathsf{T}}$、$(A^{\mathsf{T}})^{\mathsf{T}}=A$。
- **对称与斜对称**分别满足 $A^{\mathsf{T}}=A$ 与 $A^{\mathsf{T}}=-A$；后者的对角元全为零。自由元素数分别为 $n(n+1)/2$ 与 $n(n-1)/2$。
- **置换矩阵**记录行的重排，满足 $P^{-1}=P^{\mathsf{T}}$；复合重排对应矩阵乘积，右侧矩阵先作用。
- **子空间**非空且对任意线性组合封闭：$u,v\in W$、$c,d\in\mathbb{R}$ 推出 $cu+dv\in W$。含零向量只是必要条件。
- **列空间**是矩阵全部列的所有实系数线性组合；$Ax=b$ 有解当且仅当 $b\in\mathrm{Col}(A)$。
- **矩阵空间**中的加法与数乘按元素进行。对称矩阵和斜对称矩阵分别构成子空间，非对称矩阵的集合不构成子空间。
- **唯一分解**：$A=(A+A^{\mathsf{T}})/2+(A-A^{\mathsf{T}})/2$，两项分别对称与斜对称。

第 6 讲将继续研究列空间，并把齐次方程 $Ax=0$ 的所有解作为另一个子空间来处理。这里对封闭性的逐项验证会继续使用。

练习

5.1 ★ 求 $A=\begin{bmatrix}2&0&-1\\3&4&5\end{bmatrix}$ 的转置，并写出 $A^{\mathsf{T}}A$、AA^{T} 的尺寸。

5.2 ★ 写出一般的四阶对称矩阵与斜对称矩阵，各有多少自由元素？推广到 n 阶。

5.3 ★★ 构造 $P\neq I_3$ 满足 $P^3=I_3$，再构造四阶置换矩阵 Q 满足 $Q^4\neq I_4$，说明理由。

5.4 ★★ 设 S_{12}、S_{23} 分别交换三行中的前两行、后两行。求 $S_{23}S_{12}$ 与 $S_{12}S_{23}$，解释作用次序。

5.5 ★ 判断下列集合是否为 $\mathbb{R}^2$ 的子空间：$x+y=0$；$x+y=1$；$xy=0$；$x\ge 0$。成立的写闭包证明，失败的给反例。

5.6 ★★ 判断 $\mathbb{R}^3$ 中集合 $U=\{x: x_1+x_2=0\}$ 与 $V=\{x: x_2+x_3=0\}$ 的交是否为子空间，并写出交集中全部向量。

5.7 ★★ 对 $A=\begin{bmatrix}1&0\\1&1\\0&1\end{bmatrix}$，描述列空间，并判断 $b=(2,5,3)^T$、$c=(2,4,3)^T$ 是否可达；有解时求解。

5.8 ★★ 在 $M_2(\mathbb{R})$ 中分别判断：对称矩阵、斜对称矩阵、非对称矩阵、可逆矩阵是否构成子空间。

5.9 ★★ 把 $A=\begin{bmatrix}1&5\\1&3\end{bmatrix}$ 分解成对称部分与斜对称部分，证明二者只有零矩阵可能同时具有两种性质。

5.10 ★★★ 设 U,V 是同一向量空间的子空间。证明 $U\cap V$ 是子空间；证明 $U\cup V$ 是子空间当且仅当 $U\subseteq V$ 或 $V\subseteq U$。

练习答案与提示

5.1 $A^{\mathsf T}=\begin{bmatrix}2&3\\0&4\\-1&5\end{bmatrix}$。$A^{\mathsf T}A$ 为 3×3、$AA^{\mathsf T}$ 为 2×2。

5.2 四阶对称、斜对称矩阵分别有 10、6 个自由元素；一般分别为 $n(n+1)/2$ 与 $n(n-1)/2$，计数来自对角线及严格上三角。

5.3 取三循环 $P=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}$，则 $P\neq I_3$、$P^3=I_3$。令 $Q=\operatorname{diag}(1,P)$，有 $Q^3=I_4$、$Q^4=Q\neq I_4$。

5.4 $S_{23}S_{12}=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$，$S_{12}S_{23}=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}$。行序分别为 $2,3,1$ 与 $3,1,2$；乘积从右向左作用。

5.5 $x+y=0$ 对任意线性组合封闭；$x+y=1$ 不含零向量。$xy=0$ 中 $(1,0)^T+(0,1)^T=(1,1)^T$ 不封闭；$x\ge 0$ 对负数数乘不封闭。

5.6 $U\cap V=\{t(1,-1,1)^T: t\in\mathbb{R}\}$，由 $x_1=x_3=-x_2$ 得到；该集合含零且对线性组合封闭。

5.7 $Ax=(s,s+t,t)^T$，故列空间满足 $b_2=b_1+b_3$，反过来取 $s=b_1,t=b_3$。b 的解为 $(2,3)$；c 不满足 $4=2+3$，无解。

5.8 对称、斜对称矩阵集对线性组合封闭。非对称矩阵集不含零且加法不封闭；可逆矩阵集不含零，且 $I+(-I)=0$。

5.9 $S=\begin{bmatrix}1&3\\3&3\end{bmatrix}$、$K=\begin{bmatrix}0&2\\-2&0\end{bmatrix}$ 满足 $A=S+K$、$S^{\mathsf T}=S$、$K^{\mathsf T}=-K$。若 D 兼具两性质，则 $D=-D$，故 $D=0$，分解唯一。

5.10 交含零且对线性组合封闭。若一方包含另一方，并等于较大空间。若互不包含，取 $u\in U\setminus V$、$v\in V\setminus U$；$u+v$ 属于并会分别推出 $v\in U$ 或 $u\in V$，矛盾。故并为子空间当且仅当一方包含另一方。

第 6 讲

列空间与零空间

Column Space and Nullspace

对应材料：MIT 18.06SC Session 1.6 课时摘要（对应 Strang 第 6 讲视频 Column Space and Nullspace）；Session 1.6 习题 6.1–6.3 及解答；Strang《Introduction to Linear Algebra》第 3.1–3.2 节。

本讲目标

(1) 会用「对加法和数乘封闭」判断一个集合是不是子空间，知道「含零向量」是必要条件。

(2) 知道两个子空间的交一定是子空间而并一般不是，会求交 $S \cap T$ 与和 $S + T$。

(3) 掌握列空间 $C(A)$ 的定义，理解 $Ax = b$ 有解当且仅当 $b \in C(A)$，会用消元写出 b 的可解条件。

(4) 掌握零空间 $N(A)$ 的定义，会证明它是子空间，会对小矩阵求出 $N(A)$。

(5) 理解 $b \neq 0$ 时 $Ax = b$ 的解集不是子空间，而是零空间平移后的副本：全部解 $= x_p + N(A)$。

(6) 体会构造子空间的两种方式：用向量张成（列空间）与用方程刻画（零空间）。

本讲围绕 Strang 在课上提出的一个「大问题」展开：

> 给定矩阵A，对哪些右端b，方程$Ax = b$ 有解？

第 2 讲的消元法针对一个具体的 b 回答「有没有解」。现在我们要一次性描述所有能解的 b。答案是一个子空间，叫列空间。紧接着还有第二个问题：有解的时候，解有多少个？答案由另一个子空间控制，叫零空间。

进入这两个空间之前，先把第 5 讲的子空间概念再磨一遍，特别是两个子空间的「交」和「并」。本讲把列向量写成 $(1, 2, 3)$ 这样的横排形式以节省篇幅，它表示的仍是列向量。

6.1 子空间回顾

第 5 讲说过，向量空间（vector space）是一族对线性组合封闭的向量：空间中任取两个向量 v, w 和任意两个实数 c, d，组合 $cv + dw$ 仍在这个空间里。包含在一个向量空间里、自身也是向量空间的集合，叫子空间（subspace）。

定义 6.1（子空间）　设 V 是 $\mathbb{R}^n$ 的非空子集。如果

(1)（加法封闭）对任意 $v, w \in V$，有 $v + w \in V$；

(2)（数乘封闭）对任意 $v \in V$ 和任意实数 c，有 $cv \in V$，

就称 V 是 $\mathbb{R}^n$ 的**子空间**。两条合起来等价于：对任意 $v, w \in V$ 和实数 c, d，都有 $cv + dw \in V$。

两条要求里藏着第三条：在数乘封闭中取 $c = 0$，得到 $0 \cdot v$，也就是零向量，所以零向量必须在 V 里。

> 子空间一定含零向量。

这是最快的排除手段：一个集合不含零向量，立刻知道它不是子空间。反过来，含零向量还不够，下面的例子里就有含零向量却不是子空间的集合。

你可能会问：向量空间的定义（Strang 书 3.1 节）还有八条运算律，例如 $v + w = w + v$、$c(v + w) = cv + cw$，为什么 $\mathbb{R}^n$ 的子集只验证两条？因为那些运算律对 $\mathbb{R}^n$ 中的全部向量都成立，对其中一部分向量当然也成立。真正可能出问题的只有一件事：运算的结果会不会跑出这个集合。

■ **例 6.1　判断子空间**　下列集合中哪些是子空间？不是的，指出违反了哪一条。

(1) $\mathbb{R}^2$ 中的直线 $y = 3x$，即 $\{\,(x, 3x) \mid x \in \mathbb{R}\,\}$；

(2) $\mathbb{R}^2$ 中的直线 $y = 3x + 1$；

(3) $\mathbb{R}^2$ 的第一象限 $\{\,(x, y) \mid x \geq 0,\ y \geq 0\,\}$；

(4) $\mathbb{R}^2$ 中两条坐标轴的并 $\{\,(x, y) \mid xy = 0\,\}$；

(5) $\mathbb{R}^3$ 中的平面 $x + 2y - z = 0$。

解　逐个检查两条封闭性。

(1) 是。$(x, 3x) + (x', 3x') = (x + x',\, 3(x + x'))$ 仍在直线上；$c(x, 3x) = (cx,\, 3(cx))$ 也在直线上。

(2) 不是。代入 $(0, 0)$：$0 \neq 3 \cdot 0 + 1$，不含零向量。这条直线与上一条平行，只是没有过原点。

(3) 不是。加法封闭：两个分量都非负的向量相加，分量仍非负。数乘不封闭：$(1, 1)$ 在其中，$(-1) \cdot (1, 1) = (-1, -1)$ 不在。

(4) 不是。数乘封闭：若 $xy = 0$，则 $(cx)(cy) = c^2 xy = 0$。加法不封闭：$(1, 0)$ 和 $(0, 1)$ 都在其中，和 $(1, 1)$ 满足 $1 \cdot 1 = 1 \neq 0$，不在其中。

(5) 是。设 (x, y, z) 与 (x', y', z') 都满足方程，则

$$(x + x') + 2(y + y') - (z + z') = (x + 2y - z) + (x' + 2y' - z') = 0 + 0 = 0,$$

所以和也在平面上；又 $cx + 2(cy) - (cz) = c(x + 2y - z) = c \cdot 0 = 0$，数乘也封闭。

(3) 与 (4) 说明两条封闭性缺一不可：(3) 只满足加法，(4) 只满足数乘，它们都含零向量，都不是子空间。(5) 的论证只用到了方程右端是 0，这正是 §6.5 零空间的雏形。　♦

第 5 讲列出过 $\mathbb{R}^3$ 的全部子空间，它们都是「过原点的平直对象」：

- 只含零向量的集合 $Z=\{0\}$；
- 过原点的直线；
- 过原点的平面；
- $\mathbb{R}^3$ 本身。

为什么没有别的（比如过原点的曲线或锥面）？要等第 9 讲有了维数才能严格证明。本讲的列空间和零空间都会是这几类之一。

6.2　两个子空间的交、并与和

已有两个子空间，能否拼出新的子空间？最自然的两种拼法是取并和取交。摘要里的例子是：$\mathbb{R}^3$ 中过原点的平面 P 和过原点的直线 L。

例 6.2　平面与直线的并　设 P 是 $\mathbb{R}^3$ 中的平面 $x+y+z=0$，L 是 z 轴 $\{(0,0,z)\mid z\in\mathbb{R}\}$。$P$ 和 L 都是子空间（P 的论证同例 6.1(5)）。问：$P\cup L$ 是子空间吗？$P\cap L$ 是什么？

解　$P\cup L$ 含零向量，对数乘也封闭（P 中向量的倍数仍在 P 中，L 中向量的倍数仍在 L 中）。问题出在加法。取

$$p=(1,-1,0)\in P\quad(1-1+0=0),\qquad l=(0,0,1)\in L.$$

它们的和 $p+l=(1,-1,1)$：代入平面方程得 $1-1+1=1\neq 0$，不在 P 中；第一个分量是 $1\neq 0$，不在 z 轴上。所以 $p+l\notin P\cup L$，$P\cup L$ **不是子空间**。

再求交。L 中的向量形如 $(0,0,z)$，它在 P 中当且仅当 $0+0+z=0$，即 $z=0$。所以 $P\cap L=\{0\}$，只含零向量，是最小的那个子空间。 ♦

例子中的并失败，原因可以说得很具体：从 P 里取一个不在 L 中的向量，从 L 里取一个不在 P 中的向量，二者之和哪边都不在。这个论证对任意两个子空间都适用。

命题 6.2（并）　设 S,T 是 $\mathbb{R}^n$ 的子空间。$S\cup T$ 是子空间，当且仅当 $S\subseteq T$ 或 $T\subseteq S$。

思路　「当」的方向是平凡的：一个包含另一个，并集就是较大的那个。「只当」的方向用反证：两个都不包含对方，就各取一个「出界」的向量，证明它们的和既不在 S 中也不在 T 中。

证明　若 $S\subseteq T$，则 $S\cup T=T$ 是子空间；$T\subseteq S$ 同理。

反过来，假设 $S\not\subseteq T$ 且 $T\not\subseteq S$。于是可以取 $s\in S$ 但 $s\notin T$，取 $t\in T$ 但 $t\notin S$。考察 $s+t$：

- 若 $s+t\in S$，由于 $s\in S$ 且 S 对线性组合封闭，$t=(s+t)-s\in S$，与 $t\notin S$ 矛盾。
- 若 $s+t\in T$，同理 $s=(s+t)-t\in T$，与 $s\notin T$ 矛盾。

所以 $s+t\notin S\cup T$，而 $s,t\in S\cup T$，$S\cup T$ 对加法不封闭，不是子空间。 ■

证明的关键一步是「$t=(s+t)-s$」：子空间里两个向量的差还在子空间里，所以一个向量和它的「伙伴」只要有一个在 S 外，和就一定在 S 外。

图 6.1 在 $\mathbb{R}^2$ 中画出了同一现象：两条过原点的不同直线 S 和 T，各取一个向量相加，和落在两条直线之外。

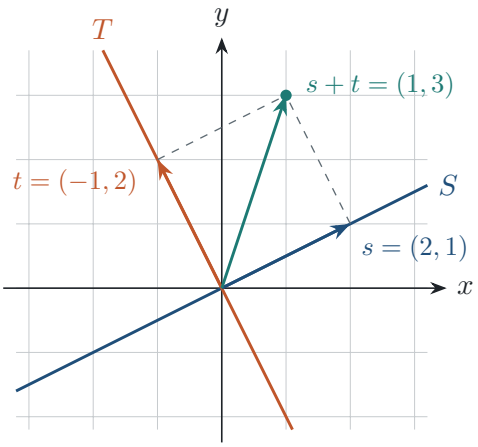


图 6.1　S、T 是 $\mathbb{R}^2$ 中过原点的两条直线。$s+t=(1,3)$ 不在任何一条上，所以 $S\cup T$ 不是子空间。两条直线上向量的所有和 $s+t$ 铺满了整个平面，即 $S+T=\mathbb{R}^2$。

并失败了，交却总是成功的。

> **定理 6.3（交）**　若 S,T 是 $\mathbb{R}^n$ 的子空间，则 $S\cap T$ 也是子空间。

证明　零向量既在 S 中又在 T 中，所以 $0\in S\cap T$，交集非空。设 $v,w\in S\cap T$，c,d 是实数。因为 $v,w\in S$ 且 S 是子空间，$cv+dw\in S$；因为 $v,w\in T$ 且 T 是子空间，$cv+dw\in T$。所以 $cv+dw\in S\cap T$。 ■

同样的论证说明任意多个子空间的交仍是子空间：一个组合只要在每个子空间里，就在它们的交里。

■ **例 6.3　两平面之交**　求 $\mathbb{R}^3$ 中平面 P_1：$x+y+z=0$ 与 P_2：$x-y=0$ 的交。

解　$(x,y,z)\in P_1\cap P_2$ 当且仅当它同时满足两个方程：

$$\begin{cases} x+y+z=0,\\ x-y=0. \end{cases}$$

由第二个方程 $y=x$，代入第一个得 $2x+z=0$，即 $z=-2x$。于是交中的向量都形如

$$(x,y,z)=(x,\,x,\,-2x)=x\,(1,1,-2),\qquad x\in\mathbb{R}.$$

检验：$(1,1,-2)$ 满足 $1+1-2=0$ 与 $1-1=0$。所以 $P_1\cap P_2$ 是方向为 $(1,1,-2)$ 的过原点直线。

两个不同的过原点平面交成一条过原点的直线，这符合几何直觉。代数上，求交就是把两组方程合起来联立。§6.5 会看到，这个交正是矩阵 $\begin{bmatrix}1&1&1\\1&-1&0\end{bmatrix}$ 的零空间。 ♦

并集不是子空间，那么「包含 S 和 T 的最小子空间」是什么？图 6.1 已经给了提示：把 S 中的向量与 T 中的向量相加。这就是习题 6.1 引入的和。先补一个术语。

定义 6.4（张成） 一组向量的全部（有限）线性组合构成的集合，叫这组向量**张成**（span）的空间，记作 span。例如 $\operatorname{span}\{v,w\}=\{\,cv+dw \mid c,d\in\mathbb{R}\,\}$。

$\operatorname{span}\{v,w\}$ 总是子空间：两个组合相加 $(cv+dw)+(c'v+d'w)=(c+c')v+(d+d')w$ 仍是组合，组合的倍数 $k(cv+dw)=(kc)v+(kd)w$ 也是组合。Strang 的说法是：任何一组向量都张成一个向量空间。

定义 6.5（子空间的和） 设 S,T 是 $\mathbb{R}^n$ 的子空间。它们的**和**（sum）是

$$S+T=\{\,s+t \mid s\in S,\ t\in T\,\},$$

即 S 中一个向量与 T 中一个向量之和的全体。

定理 6.6（和；改编自习题 6.1） 设 S,T 是 $\mathbb{R}^n$ 的子空间。则

(1) $S+T$ 是子空间，且 $S\cup T\subseteq S+T$；

(2) $S+T=\operatorname{span}(S\cup T)$；

(3) 任何包含 S 和 T 的子空间 W 都包含 $S+T$。

所以 $S+T$ 是包含 $S\cup T$ 的**最小**子空间。

证明 (1) 取 $S+T$ 中两个元素 $s+t$ 与 $s'+t'$（$s,s'\in S$，$t,t'\in T$）和实数 c。则

$$(s+t)+(s'+t')=(s+s')+(t+t'),\qquad c(s+t)=cs+ct.$$

因为 S,T 是子空间，$s+s'$ 与 cs 在 S 中，$t+t'$ 与 ct 在 T 中，所以两个结果都还是「S 中向量 $+$ T 中向量」的形式，属于 $S+T$。又 $s=s+0$，$t=0+t$，所以 S 与 T 都含于 $S+T$。

(2) 每个 $s+t$ 是 $S\cup T$ 中两个向量的组合（系数都是 1），所以 $S+T\subseteq\operatorname{span}(S\cup T)$。反过来，$S\cup T$ 中有限个向量的组合，可以把来自 S 的项归为一组、来自 T 的项归为另一组（同时在两边的向量随意归入一组），写成

$$\underbrace{(c_1s_1+\cdots+c_ks_k)}_{\in S}+\underbrace{(d_1t_1+\cdots+d_lt_l)}_{\in T},$$

这正是 $S+T$ 中的元素。

(3) 若 W 是子空间且 $S,T\subseteq W$，则任意 $s\in S$、$t\in T$ 都在 W 中，由 W 对加法封闭得 $s+t\in W$。 ■

■ **例 6.4　两条直线之和；习题 6.1(b)**　设 $S=\operatorname{span}\{u\}$，$T=\operatorname{span}\{w\}$ 是 $\mathbb{R}^3$ 中的两条直线，$u=(1,1,0)$，$w=(0,1,1)$。求 $S+T$，并比较 $S+T$ 与 $S\cup T$。

解　S 中的向量是 cu，T 中的向量是 dw，所以

$$S+T=\{\,cu+dw\mid c,d\in\mathbb{R}\,\}=\{\,(c,\ c+d,\ d)\mid c,d\in\mathbb{R}\,\}.$$

设 $(x,y,z)=(c,c+d,d)$，则 $x=c$，$z=d$，$y=c+d=x+z$。所以 $S+T$ 中每个向量都满足

$$x-y+z=0.$$

反过来，若 (x,y,z) 满足 $x-y+z=0$，取 $c=x$，$d=z$，就有 $cu+dw=(x,\,x+z,\,z)=(x,y,z)$。因此 $S+T$ **恰好是平面** $x-y+z=0$。检验：u 给出 $1-1+0=0$，w 给出 $0-1+1=0$。

$S\cup T$ 只是两条直线。例如 $u+w=(1,2,1)$ 在平面 $S+T$ 中（$1-2+1=0$），但它不是 u 的倍数（第三个分量 $1\neq 0$），也不是 w 的倍数（第一个分量 $1\neq 0$），所以不在 $S\cup T$ 中。这就是习题 6.1(b) 的答案：两条不同直线的并只是两条线，和却是它们张成的整个平面；$S\cup T$ 张成的空间正是 $S+T$。 ♦

三种运算的对比如表 6.1。

	交 $S\cap T$	并 $S\cup T$	和 $S+T$
是不是子空间	总是	仅当一个含于另一个	总是
与 S,T 的关系	含于 S 和 T	含 S 和 T	含 S 和 T
$\mathbb{R}^3$ 中两条不同直线	$\{0\}$	两条直线	它们张成的平面
$\mathbb{R}^3$ 中两个不同平面	一条直线	两个平面	$\mathbb{R}^3$
怎样算	联立两组方程	不必算	合并两组生成向量

表 6.1　两个子空间的交、并与和。交是包含于二者的最大子空间，和是包含二者的最小子空间。

想一想　设直线 L 落在平面 P 里（二者都过原点）。$P\cup L$ 是子空间吗？$P\cap L$ 与 $P+L$ 各是什么？

答　$L\subseteq P$，所以 $P\cup L=P$，是子空间，这是命题 6.2 中「当」的情形。$P\cap L=L$，$P+L=P$。摘要说「$P\cup L$ 一般不是子空间」，「一般」二字排除的就是这种一个含于另一个的情形。

6.3　列空间 $C(A)$

第 1 讲的「列图像」说：矩阵乘向量，就是用 x 的分量作系数，把 A 的列组合起来。设 A 是 $m\times n$ 矩阵，列为 $a_1,\dots,a_n\in\mathbb{R}^m$，则

$$Ax=\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}=x_1a_1+x_2a_2+\cdots+x_na_n.$$

让 x 取遍 $\mathbb{R}^n$ 中的所有向量，Ax 就取遍 A 的列的所有组合。

定义 6.7（列空间）　$m \times n$ 矩阵 A 的**列空间**（column space）$C(A)$ 是 A 的列的全部线性组合构成的集合：

$$C(A) = \operatorname{span}\{a_1, \dots, a_n\} = \{\, Ax \mid x \in \mathbb{R}^n \,\} \subseteq \mathbb{R}^m.$$

命题 6.8　$C(A)$ 是 $\mathbb{R}^m$ 的子空间。

证明　$C(A)$ 是一组向量张成的空间，由 §6.2 的讨论已是子空间。也可以直接验证：设 $b = Ax$，$b' = Ay$ 是 $C(A)$ 中两个向量，c 是实数。由矩阵乘法的线性性，

$$b + b' = Ax + Ay = A(x + y), \qquad cb = c(Ax) = A(cx),$$

二者都形如「A 乘某个向量」，都在 $C(A)$ 中。 ■

易错提醒　$C(A)$ 住在 $\mathbb{R}^m$ 里，m 是**行数**，不是列数。原因是每一列有 m 个分量。4×3 矩阵有 3 列，每列是 $\mathbb{R}^4$ 中的向量，所以它的列空间是 $\mathbb{R}^4$ 的子空间。

直观理解　把 A 看成一台机器：输入 $x \in \mathbb{R}^n$，输出 $Ax \in \mathbb{R}^m$。列空间就是这台机器**所有可能的输出**。线性代数的其他书常把它叫作 A 的**值域**（range）或**像**（image）。「$Ax = b$ 有没有解」换个说法就是：b 是不是这台机器能造出来的输出。

■ **例 6.5　几个列空间**　描述下列矩阵的列空间：

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad Z = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

解

(1) $Ix = x$，任何 $b \in \mathbb{R}^3$ 都是 I 的输出（取 $x = b$），所以 $C(I) = \mathbb{R}^3$。

(2) A_1 可逆（第 3 讲）：消元的两个主元都是 1，逆为 $A_1^{-1} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$。对任意 $b = (b_1, b_2)$，取 $x = A_1^{-1} b = (7b_1 - 2b_2,\ -3b_1 + b_2)$，验证：

$$A_1 x = \big((7b_1 - 2b_2) + 2(-3b_1 + b_2),\ 3(7b_1 - 2b_2) + 7(-3b_1 + b_2)\big) = (b_1, b_2).$$

所以 $C(A_1) = \mathbb{R}^2$。一般地，**可逆方阵的列空间是整个空间**。

(3) A_2 的第二列 $(2, 4)$ 是第一列 $(1, 2)$ 的 2 倍，于是

$$x_1(1, 2) + x_2(2, 4) = (x_1 + 2x_2)\,(1, 2).$$

所有输出都是 $(1, 2)$ 的倍数。反过来 $c(1, 2) = A_2(c, 0)$。所以 $C(A_2) = \operatorname{span}\{(1, 2)\}$，是 $\mathbb{R}^2$ 中的直线 $b_2 = 2b_1$。两列都在，却只张成一条线，因为第二列没有带来新方向。

(4) $A_3x = x_1(1,0,1) + x_2(0,1,1) = (x_1,\ x_2,\ x_1+x_2)$。输出的第三个分量总是前两个之和，所以 $C(A_3)$ 含于平面 $b_3 = b_1 + b_2$；反过来，平面上任一点 $(b_1, b_2, b_1+b_2) = A_3(b_1, b_2)$。所以 $C(A_3)$ **恰好是 $\mathbb{R}^3$ 中的平面 $b_3 = b_1 + b_2$**（图 6.2）。例如 $(1,1,0)$ 不在其中，因为 $0 \neq 1+1$。

(5) $Zx = 0$ 对一切 x 成立，$C(Z) = \{0\}$。

♦

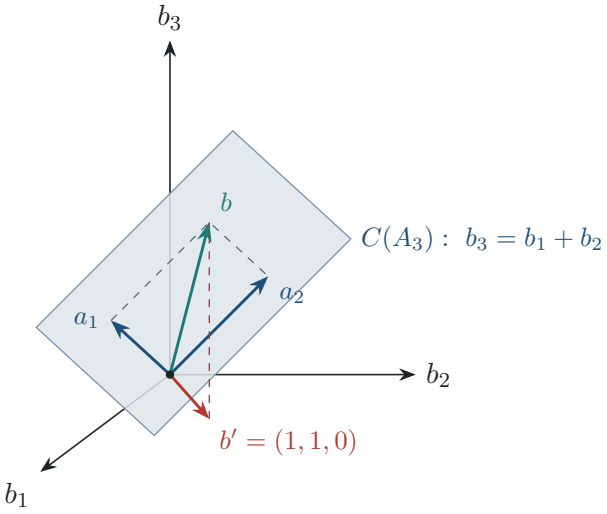


图 6.2　A_3 的列空间是 $\mathbb{R}^3$ 中过原点的平面。$b = a_1 + a_2 = (1,1,2)$ 在平面上，$A_3x = b$ 有解 $x = (1,1)$；$b' = (1,1,0)$ 在 b 的正下方，离开了平面，$A_3x = b'$ 无解。

6.4　什么时候 $Ax = b$ 有解

上一节的例子已经把答案摆在那里了。把它写成定理。

定理 6.9（可解性）　方程组 $Ax = b$ 有解，当且仅当 $b \in C(A)$。

证明　$Ax = b$ 有解 $\iff$ 存在 x 使 $x_1a_1 + \cdots + x_na_n = b$ $\iff$ b 是 A 的列的线性组合 $\iff$ $b \in C(A)$。 ■

证明只有一行，因为它不过是把定义换个说法。它的价值在于换了观点：「方程组有没有解」是一个**代数**问题，「b 在不在 $C(A)$ 这个子空间里」是一个**几何**问题。从此我们可以对所有 b 一起回答，而不必一个一个去试。解 x 的分量，就是把 b 写成列的组合时用的系数。

6.4.1　摘要中的主例子

Strang 在课上用的矩阵是

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}, \qquad a_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}.$$

$Ax=b$ 是 4 个方程、3 个未知数。直觉上，方程比未知数多，不是每个 b 都能解：三根列向量不太可能「填满」四维空间 $\mathbb{R}^4$。

Strang 建议反过来想：先挑 x，再看得到哪个 b。

- $x=(0,0,0)$ 给出 $b=0$。零向量总在列空间里。
- $x=(1,0,0)$ 给出 $b=a_1=(1,2,3,4)$；$x=(0,1,0)$ 给出 $b=a_2=(1,1,1,1)$。每一列本身都可达。
- $x=(1,1,0)$ 给出 $b=a_1+a_2=(2,3,4,5)$，这恰好是 a_3。

最后一条暴露了这个矩阵的关键事实：

$$\boxed{a_3=a_1+a_2.}$$

第三列不给列空间带来任何新东西。任何组合都可以改写成只用前两列：

$$x_1a_1+x_2a_2+x_3a_3=x_1a_1+x_2a_2+x_3(a_1+a_2)=(x_1+x_3)\,a_1+(x_2+x_3)\,a_2.$$

所以 $C(A)=\operatorname{span}\{a_1,a_2\}$。又 a_1,a_2 不共线（a_2 的分量全相等，a_1 的不全相等，二者不成比例），所以 $C(A)$ 是 $\mathbb{R}^4$ 中一个过原点的平面。用第 9 讲的语言，$C(A)$ 是 $\mathbb{R}^4$ 的二维子空间。

■ **例 6.6　求列空间的方程**　对上面的 4×3 矩阵 A，求 $b=(b_1,b_2,b_3,b_4)$ 属于 $C(A)$ 的条件。

解　*方法一：参数化。*$C(A)=\operatorname{span}\{a_1,a_2\}$，其中的向量是

$$ca_1+da_2=(c+d,\ 2c+d,\ 3c+d,\ 4c+d).$$

相邻分量之差都等于 c，所以 b 是一个等差数列。反过来，若 $b_2-b_1=b_3-b_2=b_4-b_3$，记公差为 c，取 $d=b_1-c$，则 $ca_1+da_2=(b_1,\,b_1+c,\,b_1+2c,\,b_1+3c)=b$。因此

$$b\in C(A)\iff b_2-b_1=b_3-b_2=b_4-b_3\iff\begin{cases}b_1-2b_2+b_3=0,\\ b_2-2b_3+b_4=0.\end{cases}$$

*方法二：对增广矩阵消元。*这是对任何矩阵都适用的做法。把 b 带在最后一列，按第 2 讲的步骤消元。先用第 1 行消去下面三行的第一个元素（乘数 $2,3,4$）：

$$\left[\begin{array}{ccc|c}1&1&2&b_1\\2&1&3&b_2\\3&1&4&b_3\\4&1&5&b_4\end{array}\right]\longrightarrow\left[\begin{array}{ccc|c}1&1&2&b_1\\0&-1&-1&b_2-2b_1\\0&-2&-2&b_3-3b_1\\0&-3&-3&b_4-4b_1\end{array}\right].$$

再用第 2 行消去第 3、4 行的第二个元素（乘数 $2,3$）：

$$\longrightarrow\left[\begin{array}{ccc|c}1&1&2&b_1\\0&-1&-1&b_2-2b_1\\0&0&0&b_1-2b_2+b_3\\0&0&0&2b_1-3b_2+b_4\end{array}\right].$$

这里第 3 行右端是 $(b_3-3b_1)-2(b_2-2b_1)=b_1-2b_2+b_3$，第 4 行右端是 $(b_4-4b_1)-3(b_2-2b_1)=2b_1-3b_2+b_4$。后两行左边全是 0，它们读作 $0=b_1-2b_2+b_3$ 与 $0=2b_1-3b_2+b_4$。若这两个数不为零，方程组就出现「$0=$ 非零数」的矛盾；若它们都为零，前两行总能回代求解（x_3 随便取，再解出 x_2,x_1）。所以

$$b\in C(A)\iff b_1-2b_2+b_3=0\ \text{且}\ 2b_1-3b_2+b_4=0.$$

两种方法的条件一致：第二组的第二个方程减去第一个方程的 2 倍，得 $(2b_1-3b_2+b_4)-2(b_1-2b_2+b_3)=b_2-2b_3+b_4$，正是方法一的第二个方程。♦

■ **例 6.7　逐个检验**　对同一个 A，判断 $Ax=b$ 对下列 b 是否有解，有解时求出一个解：$b=(2,5,8,11)$，$(1,0,0,0)$，$(1,1,1,1)$，$(0,0,0,1)$。

解　用例 6.6 的两个条件 $b_1-2b_2+b_3=0$，$2b_1-3b_2+b_4=0$：

b	$b_1-2b_2+b_3$	$2b_1-3b_2+b_4$	是否可解
$(2,5,8,11)$	$2-10+8=0$	$4-15+11=0$	可解
$(1,0,0,0)$	1	2	不可解
$(1,1,1,1)$	$1-2+1=0$	$2-3+1=0$	可解
$(0,0,0,1)$	0	1	不可解

$(1,1,1,1)=a_2$，取 $x=(0,1,0)$ 即可。对 $b=(2,5,8,11)$ 回代：消元后前两行是

$$x_1+x_2+2x_3=2,\qquad -x_2-x_3=5-2\cdot 2=1.$$

取 $x_3=0$，得 $x_2=-1$，$x_1=2-x_2=3$，即 $x=(3,-1,0)$。检验：$3a_1-a_2=(3-1,\,6-1,\,9-1,\,12-1)=(2,5,8,11)$。

取 $x_3=-1$ 也行：$-x_2+1=1$ 给出 $x_2=0$，$x_1=2-0+2=4$，即 $x=(4,0,-1)$。检验：$4a_1-a_3=(4-2,\,8-3,\,12-4,\,16-5)=(2,5,8,11)$。同一个 b 有不止一个解，这把我们引向下一节。

$(0,0,0,1)$ 的不可解也可以直接看出：它不是等差数列（差依次为 $0,0,1$）。♦

想一想　换一个 4×3 矩阵，让三列「尽量独立」，它的列空间有没有可能是整个 $\mathbb{R}^4$？

答　不可能。消元时每个主元占一列，3 列至多产生 3 个主元，而有 4 行，所以消元后（必要时换行）至少有一行左边全为 0。这一行的右端是 b 各分量的某个组合，可解要求它等于 0。这个组合的系数不全为零（它来自对 b 做的可逆行变换），所以总有 b 违反它，那样的 b 就不在列空间中。第 9 讲会用维数把这个论证说得更简洁：三个向量张成的空间至多三维。

消元会改变列空间　求可解条件时，读的是增广矩阵的最后一列，不能用消元后矩阵 U 的列空间代替 $C(A)$。上例中

$$U=\begin{bmatrix}1&1&2\\0&-1&-1\\0&0&0\\0&0&0\end{bmatrix}$$

的每一列后两个分量都是 0，所以 $C(U)$ 里的向量后两个分量都是 0。而 $a_1=(1,2,3,4)$ 在 $C(A)$ 中却不在 $C(U)$ 中。行变换把每一列都变了样，列空间也跟着变了。下一节会看到，零空间在消元下保持不变。

6.5　零空间 $N(A)$

主例子里的关系 $a_1+a_2-a_3=0$ 可以写成矩阵形式：

$$\begin{bmatrix}1&1&2\\2&1&3\\3&1&4\\4&1&5\end{bmatrix}\begin{bmatrix}1\\1\\-1\end{bmatrix}=1\cdot a_1+1\cdot a_2+(-1)\cdot a_3=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}.$$

也就是说，$x=(1,1,-1)$ 是 $Ax=0$ 的一个非零解。这样的 x 组成第二个重要的子空间。

定义 6.10（零空间）　$m\times n$ 矩阵 A 的零空间（nullspace）$N(A)$ 是方程 $Ax=0$ 的全部解构成的集合：

$$N(A)=\{\,x\in\mathbb{R}^n\mid Ax=0\,\}\subseteq\mathbb{R}^n.$$

其他书也叫 A 的核（kernel），记作 $\ker A$。

定理 6.11　$N(A)$ 是 $\mathbb{R}^n$ 的子空间。

证明　$A0=0$，所以 $0\in N(A)$。设 $x,y\in N(A)$，即 $Ax=0$，$Ay=0$；设 c 是实数。由矩阵乘法的线性性，

$$A(x+y)=Ax+Ay=0+0=0,\qquad A(cx)=c(Ax)=c\cdot 0=0.$$

所以 $x+y$ 与 cx 都是 $Ax=0$ 的解，都在 $N(A)$ 中。 ■

证明里起作用的是两件事：A 对加法和数乘的线性性，以及右端是 0。右端换成别的 b，第二件事就失效了，§6.6 专门讨论这一点。

两个空间住在不同的地方　对 $m\times n$ 矩阵：$C(A)$ 由输出组成，是 $\mathbb{R}^m$ 的子空间；$N(A)$ 由输入组成，是 $\mathbb{R}^n$ 的子空间。主例子中 $C(A)\subseteq\mathbb{R}^4$，$N(A)\subseteq\mathbb{R}^3$。$m\neq n$ 时，问「$C(A)$ 与 $N(A)$ 是否相交」没有意义，因为它们根本不在同一个空间里。

■ **例 6.8　主例子的零空间**　求 $N(A)$，其中 A 是 §6.4 的 4×3 矩阵。

解　解 $Ax=0$。右端为 0 时，行变换不改变右端（零的组合还是零），所以只需对 A 消元。例 6.6 已经算出 $A\to U$，后两行为零行，$Ax=0$ 化为

$$\begin{cases} x_1+x_2+2x_3=0,\\ -x_2-x_3=0. \end{cases}$$

两个方程、三个未知数。第二个方程给出 $x_2=-x_3$，代入第一个：$x_1=-x_2-2x_3=x_3-2x_3=-x_3$。于是

$$x=(x_1,x_2,x_3)=(-x_3,\,-x_3,\,x_3)=-x_3\,(1,1,-1).$$

x_3 可以任取，所以

$$N(A)=\{\,c\,(1,1,-1)\mid c\in\mathbb{R}\,\}=\operatorname{span}\{(1,1,-1)\},$$

是 $\mathbb{R}^3$ 中过原点的一条直线。这与摘要的结论一致：第 1 列加第 2 列减第 3 列等于零向量，而且本质上只有这一种关系（其余的都是它的倍数）。

「x_3 可以任取」这样的未知数叫自由变量，第 7 讲会系统地讨论怎样用它们写出零空间。 ♦

■ **例 6.9　几个零空间**　求下列矩阵的零空间，并说出它的几何形状：

$$B_1=\begin{bmatrix}1&2\\3&7\end{bmatrix},\quad B_2=\begin{bmatrix}1&2&3\end{bmatrix},\quad B_3=\begin{bmatrix}1&2\\2&4\end{bmatrix},\quad B_4=\begin{bmatrix}1&2&3\\2&4&6\end{bmatrix},\quad Z=\begin{bmatrix}0&0\\0&0\end{bmatrix}.$$

解

(1) B_1 可逆（例 6.5）。若 $B_1x=0$，两边左乘 B_1^{-1} 得 $x=B_1^{-1}0=0$。所以 $N(B_1)=\{0\}$，只有一个点。

(2) $B_2x=0$ 是一个方程 $x_1+2x_2+3x_3=0$，即 $x_1=-2x_2-3x_3$，x_2,x_3 任取：

$$x=(-2x_2-3x_3,\;x_2,\;x_3)=x_2(-2,1,0)+x_3(-3,0,1).$$

所以 $N(B_2)=\operatorname{span}\{(-2,1,0),(-3,0,1)\}$，是 $\mathbb{R}^3$ 中过原点的平面 $x_1+2x_2+3x_3=0$。检验：$-2+2=0$，$-3+3=0$。

(3) $B_3x=0$ 是 $x_1+2x_2=0$ 与 $2x_1+4x_2=0$。第二个方程是第一个的 2 倍，只剩 $x_1=-2x_2$：

$$N(B_3)=\operatorname{span}\{(-2,1)\},$$

是 $\mathbb{R}^2$ 中的直线。检验：$B_3(-2,1)=(-2+2,\,-4+4)=(0,0)$。

(4) B_4 的第二行是第一行的 2 倍，第二个方程不带来新约束，$N(B_4)=N(B_2)$，是同一个平面。

(5) $Zx=0$ 对一切 x 成立，$N(Z)=\mathbb{R}^2$。

对比例 6.5：B_1 的列空间最大（$\mathbb{R}^2$）而零空间最小（$\{0\}$）；Z 恰好相反；$B_3=A_2$ 的列空间与零空间都是直线。对这三个同为 2×2 的矩阵，列空间与零空间的维数之和都是 2。一般地，A 有 n 列时，$\dim C(A)+\dim N(A)=n$；第 9 讲会证明这个秩—零度公式。 ♦

零空间中的每个非零向量都记录了列之间的一个关系：$x\in N(A)$ 说的正是 $x_1a_1+\cdots+x_na_n=0$。$N(A)=\{0\}$ 意味着列之间没有任何非平凡关系，第 9 讲把这叫做「列线性无关」。

6.5.1 叠起来的矩阵：零空间与交

习题 6.3 问：把两个列数相同的矩阵上下叠起来，零空间怎样变化？

命题 6.12（习题 6.3）　设 A 是 $m\times n$ 矩阵，B 是 $p\times n$ 矩阵，$C=\begin{bmatrix}A\\B\end{bmatrix}$ 是 $(m+p)\times n$ 矩阵。则

$$N(C)=N(A)\cap N(B).$$

证明　按分块乘法，$Cx=\begin{bmatrix}Ax\\Bx\end{bmatrix}$。这个向量为零，当且仅当上半截 Ax 与下半截 Bx 都为零。所以 $x\in N(C)\iff x\in N(A)$ 且 $x\in N(B)$。 ■

这就是例 6.3 背后的结构。平面 $P_1:\ x+y+z=0$ 是 1×3 矩阵 $\begin{bmatrix}1&1&1\end{bmatrix}$ 的零空间，平面 $P_2:\ x-y=0$ 是 $\begin{bmatrix}1&-1&0\end{bmatrix}$ 的零空间。把两行叠成

$$C=\begin{bmatrix}1&1&1\\1&-1&0\end{bmatrix},$$

命题 6.12 就给出 $N(C)=P_1\cap P_2=\operatorname{span}\{(1,1,-2)\}$。验算：$C(1,1,-2)=(1+1-2,\ 1-1+0)=(0,0)$。

核心思想　每加一行（一个方程），零空间就和一个新的子空间取交，因而只会变小或不变。方程越多，解越少。

消元不改变零空间　与列空间（§6.4 末的警告）相反，行变换不改变零空间。每一步消元都是可逆的：$Ax=0$ 的解满足消元后的方程，反过来，把消元步骤倒着做，消元后方程的解也满足原方程。例如

$$A=\begin{bmatrix}1&2\\2&4\end{bmatrix}\longrightarrow U=\begin{bmatrix}1&2\\0&0\end{bmatrix},\qquad N(A)=N(U)=\operatorname{span}\{(-2,1)\},$$

但 $C(A)$ 是直线 $b_2=2b_1$，$C(U)$ 是直线 $b_2=0$，两者不同。所以求零空间可以放心在 U 上做，求列空间却不行。

想一想　设 A 是 $n\times n$ 可逆矩阵。$N(A)$ 是什么？$C(A)$ 又是什么？

答　$N(A)=\{0\}$：若 $Ax=0$，左乘 A^{-1} 得 $x=0$。$C(A)=\mathbb{R}^n$：对任意 b，$x=A^{-1}b$ 满足 $Ax=b$。可逆矩阵的零空间最小、列空间最大，这正是例 6.9 中 B_1 的情形。

6.6 $Ax=b$ 的全部解

零空间回答了「$Ax=0$ 有哪些解」。右端 $b\neq 0$ 时，解的集合长什么样？

6.6.1 解集不是子空间

仍用主例子，取 $b = a_1 = (1,2,3,4)$。容易找到两个解：

$$x = (1,0,0)\ (\text{只用第 1 列}), \qquad x' = (0,-1,1)\ (-a_2 + a_3 = a_1).$$

但它们的和 $x + x' = (1,-1,1)$ 给出

$$A(1,-1,1) = a_1 - a_2 + a_3 = (2,4,6,8) = 2b \neq b,$$

不再是解。更直接的理由：$A0 = 0 \neq b$，零向量根本不是解。所以**$b \neq 0$ 时，$Ax = b$ 的解集不是子空间**。

两个解的和不行，两个解的差却很有用：

$$A(x - x') = Ax - Ax' = b - b = 0, \qquad x - x' = (1,1,-1) \in N(A).$$

任意两个解相差零空间中的一个向量。这就是下面命题的核心。

> **命题 6.13（解集的结构）**　设 x_p 是 $Ax = b$ 的一个解（称为**特解**（particular solution））。则 $Ax = b$ 的全部解是
>
> $$x_p + N(A) = \{\, x_p + x_n \mid x_n \in N(A) \,\}.$$

证明　若 $x_n \in N(A)$，则 $A(x_p + x_n) = Ax_p + Ax_n = b + 0 = b$，所以 $x_p + x_n$ 是解。反过来，若 $Ax = b$，令 $x_n = x - x_p$，则 $Ax_n = b - b = 0$，即 $x_n \in N(A)$，而 $x = x_p + x_n$。 ■

$$\boxed{\text{全部解} = \text{一个特解} + \text{零空间}}$$

几何上，$x_p + N(A)$ 是把子空间 $N(A)$ 平移到经过 x_p 的位置：它与 $N(A)$ 平行、形状相同，但一般不过原点。用哪个特解都可以，平移后得到的是同一个集合。

■ **例 6.10　主例子的全部解**　写出主例子中 $Ax = (1,2,3,4)$ 与 $Ax = (2,5,8,11)$ 的全部解。

解　例 6.8 已经求出 $N(A) = \operatorname{span}\{(1,1,-1)\}$。

对 $b = (1,2,3,4)$，取 $x_p = (1,0,0)$。由命题 6.13，

$$x = (1,0,0) + t\,(1,1,-1) = (1+t,\ t,\ -t), \qquad t \in \mathbb{R}.$$

这是 $\mathbb{R}^3$ 中一条**不过原点**的直线（$t = 0$ 时 $x_1 = 1$，但要过原点需 $1 + t = t = 0$，不可能）。$t = -1$ 给出 $x' = (0,-1,1)$，$t = 1$ 给出 $(2,1,-1)$，都可以代回验证：例如 $2a_1 + a_2 - a_3 = (2+1-2,\ 4+1-3,\ 6+1-4,\ 8+1-5) = (1,2,3,4)$。

对 $b = (2,5,8,11)$，例 6.7 求出了 $x_p = (3,-1,0)$，于是全部解是 $(3,-1,0) + t\,(1,1,-1)$。$t = 1$ 给出例 6.7 的另一个解 $(4,0,-1)$。两条解直线互相平行，也都平行于 $N(A)$。 ♦

图 6.3 用 2×2 矩阵 $A_2 = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ 把两个子空间放在一起看。左边是输入空间 $\mathbb{R}^2$：$N(A_2)$ 是直线 $x_1 + 2x_2 = 0$（例 6.9），$A_2x = (3,6)$ 的解集是与它平行的直线 $x_1 + 2x_2 = 3$，上面的

$(3,0)$、$(1,1)$、$(-1,2)$ 都被 A_2 送到同一个 $b=(3,6)$。右边是输出空间 $\mathbb{R}^2$：$C(A_2)$ 是直线 $b_2=2b_1$，$b=(3,6)$ 在上面，$b'=(3,2)$ 不在上面，$A_2x=b'$ 无解。

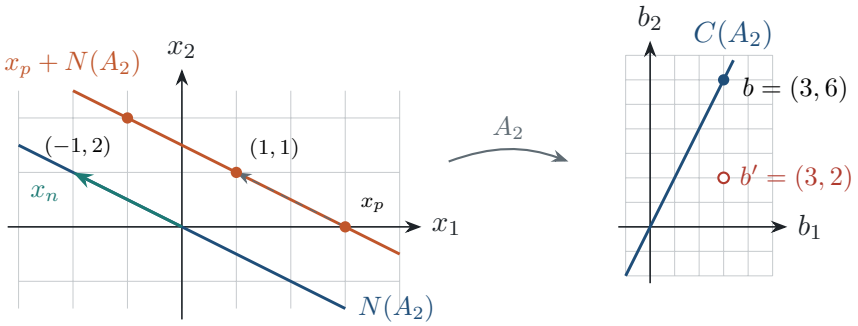


图 6.3　左：$N(A_2)$（蓝）与 $A_2x=(3,6)$ 的解集（橙）是两条平行直线，后者由前者平移 x_p 得到。虚线箭头 $(3,0)\to(1,1)$ 正是加上 $x_n=(-2,1)$。右：A_2 把整条橙线压到一个点 b 上，所有输出都落在 $C(A_2)$ 这条直线上。

■ **例 6.11　一个平移的平面；习题 6.2**　平面 $x-3y-z=12$ 与过原点的平面 $x-3y-z=0$ 平行，$(12,0,0)$ 是前者上的一点。把前者上的所有点写成

$$\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}12\\0\\0\end{bmatrix}+y\begin{bmatrix}*\\1\\0\end{bmatrix}+z\begin{bmatrix}*\\0\\1\end{bmatrix}$$

的形式，填出 $*$。（原题印成了「$x-3y-x=0$」，是 $x-3y-z=0$ 的笔误。）

解　这是 1×3 矩阵 $A=\begin{bmatrix}1 & -3 & -1\end{bmatrix}$ 的方程 $Ax=12$。把方程解成 $x=12+3y+z$，y,z 任取：

$$\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}12+3y+z\\y\\z\end{bmatrix}=\begin{bmatrix}12\\0\\0\end{bmatrix}+y\begin{bmatrix}3\\1\\0\end{bmatrix}+z\begin{bmatrix}1\\0\\1\end{bmatrix}.$$

两个 $*$ 依次是 3 和 1。对照命题 6.13：$x_p=(12,0,0)$ 是特解（$12-0-0=12$）；$(3,1,0)$ 与 $(1,0,1)$ 满足 $3-3-0=0$ 与 $1-0-1=0$，在 $N(A)$ 中，它们张成的正是过原点的平面 $x-3y-z=0$。例如 $y=1,z=2$ 给出点 $(17,1,2)$，验算 $17-3-2=12$。 ♦

想一想　设 x_1,x_2 都是 $Ax=b$ 的解。x_1-x_2 和 x_1+x_2 分别满足什么方程？中点 $\frac{1}{2}(x_1+x_2)$ 呢？

答　$A(x_1-x_2)=b-b=0$，所以 $x_1-x_2\in N(A)$。$A(x_1+x_2)=2b$，所以 x_1+x_2 解的是 $Ax=2b$，只有 $b=0$ 时才是原方程的解。中点满足 $A\cdot\frac{1}{2}(x_1+x_2)=\frac{1}{2}(b+b)=b$，仍是解。一般地，系数和为 1 的组合 $sx_1+(1-s)x_2$ 都是解，它们正好组成过 x_1,x_2 的直线。

命题 6.13 立刻回答了「解有几个」：

推论 6.14（唯一性）　若 $Ax=b$ 有解，则解唯一当且仅当 $N(A)=\{0\}$。

证明　解集是 $x_p + N(A)$。它只含一个点，当且仅当 $N(A)$ 只含一个点，即 $N(A) = \{0\}$。 ■

于是 $Ax = b$ 的解的情况由两个子空间分工决定：b 是否在 $C(A)$ 中，决定有没有解；$N(A)$ 是否为 $\{0\}$，决定有几个解。$N(A) \neq \{0\}$ 时，$N(A)$ 至少含一条直线，有解就有无穷多个解。

前后联系　线性常微分方程的通解是同一个结构。$y'' + y = 1$ 的一个特解是常数 $y_p = 1$；对应的齐次方程 $y'' + y = 0$ 的解是 $c_1 \cos t + c_2 \sin t$，它们构成函数空间的一个子空间，正是算子 $D^2 + 1$ 的「零空间」。通解是

$$y = \underbrace{1}_{y_p} + \underbrace{c_1 \cos t + c_2 \sin t}_{\text{齐次解}}.$$

证明也一样：两个解之差满足齐次方程。这里的关键只是线性：$L(y_1 - y_2) = Ly_1 - Ly_2$。

6.7　构造子空间的两种方式

列空间与零空间代表了描述子空间的两种根本不同的方式。

- **给出生成向量，取全部组合**：$C(A) = \{\, Ax \mid x \in \mathbb{R}^n \,\}$。这是显式描述，元素可以直接写出来。
- **给出方程，取全部解**：$N(A) = \{\, x \mid Ax = 0 \,\}$。这是隐式描述，只说明元素要满足的条件。

同一个子空间往往两种写法都有。例如 $\mathbb{R}^3$ 中的平面 $x + y + z = 0$ 是 1×3 矩阵 $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ 的零空间；把方程解成 $x = -y - z$，又得到

$$(x, y, z) = y\,(-1, 1, 0) + z\,(-1, 0, 1), \qquad \text{即} \qquad \{\,(x, y, z) \mid x + y + z = 0\,\} = C\left(\begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\right).$$

两种写法各有长处：要举出平面上的点，用列空间的写法，随便取 y, z；要判断一个点在不在平面上，用零空间的写法，代入方程即可。

■ **例 6.12　列空间写成零空间**　把主例子的 $C(A) \subseteq \mathbb{R}^4$ 写成某个矩阵 B 的零空间。

解　例 6.6 已经把 $C(A)$ 写成了两个方程 $b_1 - 2b_2 + b_3 = 0$，$2b_1 - 3b_2 + b_4 = 0$。把系数排成矩阵：

$$B = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 2 & -3 & 0 & 1 \end{bmatrix}, \qquad C(A) = N(B).$$

验算：$C(A) \subseteq N(B)$ 等价于 B 消灭 A 的每一列，即 $BA = 0$。逐个算：$Ba_1 = (1-4+3,\ 2-6+4) = (0, 0)$，$Ba_2 = (1 - 2 + 1,\ 2 - 3 + 1) = (0, 0)$，$Ba_3 = (2 - 6 + 4,\ 4 - 9 + 5) = (0, 0)$。

这个 B 不唯一。方法一得到的 $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix}$ 也行，它的第二行是 B 的第二行减去第一行的 2 倍，零空间不变（见 §6.5 的警告）。 ♦

■ **例 6.13　零空间写成列空间**　把 $N\left(\begin{bmatrix} 1 & -3 & -1 \end{bmatrix}\right)$ 写成某个矩阵的列空间。

解　这就是例 6.11 中过原点的那个平面。方程 $x-3y-z=0$ 解成 $x=3y+z$：

$$N\left(\begin{bmatrix}1 & -3 & -1\end{bmatrix}\right)=\{\,y\,(3,1,0)+z\,(1,0,1)\mid y,z\in\mathbb{R}\,\}=C\left(\begin{bmatrix}3 & 1\\ 1 & 0\\ 0 & 1\end{bmatrix}\right).$$

两列满足 $3-3-0=0$ 与 $1-0-1=0$。从方程出发、让一部分未知数自由取值、写出生成向量，这正是第 7 讲求零空间的标准程序。 ♦

表 6.2 把两个子空间并排比较。

	列空间 $C(A)$	零空间 $N(A)$
住在哪里（A 为 $m\times n$）	$\mathbb{R}^m$（输出）	$\mathbb{R}^n$（输入）
怎样构造	列的全部组合	$Ax=0$ 的全部解
描述方式	显式：给出生成向量	隐式：给出方程
写出元素	容易：任取 x，算 Ax	需要解方程
判断元素	需要解方程	容易：代入看 Ax 是否为 0
对 $Ax=b$ 的意义	决定有没有解	决定解有几个
消元 $A\to U$ 之后	会改变	不变

表 6.2　列空间与零空间的对照。

核心思想　两种描述之间来回转换靠的都是消元：对 $[A\mid b]$ 消元，把「生成向量」变成「方程」（例 6.12）；解 $Ax=0$，把「方程」变成「生成向量」（例 6.13）。

事实上，$\mathbb{R}^n$ 的每个子空间都既是某个矩阵的列空间，又是某个矩阵的零空间。证明要用到后面几讲的基与维数，这里不展开。

本讲小结

- 子空间对加法和数乘封闭，必含零向量。两个子空间的交总是子空间；并一般不是，仅当一个含于另一个时才是；和 $S+T$ 是包含二者的最小子空间，等于 $\operatorname{span}(S\cup T)$。
- $m\times n$ 矩阵的列空间 $C(A)=\{\,Ax\mid x\in\mathbb{R}^n\,\}$ 是 $\mathbb{R}^m$ 的子空间。$Ax=b$ 有解 $\iff b\in C(A)$。
- 零空间 $N(A)=\{\,x\mid Ax=0\,\}$ 是 $\mathbb{R}^n$ 的子空间。其中的非零向量记录了列之间的关系。$N\left(\begin{bmatrix}A\\B\end{bmatrix}\right)=N(A)\cap N(B)$。
- 消元不改变零空间，但会改变列空间。
- $Ax=b$ 无解时解集为空；有特解 x_p 时，解集为 $x_p+N(A)$，且解唯一 $\iff N(A)=\{0\}$。齐次情形 $b=0$ 的解集是 $N(A)$；$b\neq 0$ 时解集不含零向量，因而不是子空间。
- 列空间是「组合」构造的子空间，零空间是「方程」构造的子空间；消元在两种描述之间转换。

练习

6.1 ★ 判断 $\mathbb{R}^3$ 的下列子集是不是子空间；不是的，指出哪条性质失败并给出反例：
(a) $\{\,(x,y,z)\mid x=2y\,\}$；(b) $\{\,(x,y,z)\mid x+y+z=1\,\}$；
(c) $\{\,(x,y,z)\mid z\geq 0\,\}$；(d) $\{\,(x,y,z)\mid x^2=y^2\,\}$；
(e) $\{\,(x,y,z)\mid x=y=z\,\}$。

6.2 ★ 在 $\mathbb{R}^3$ 中，记 S 为 x 轴，T 为 y 轴，P 为 xy 平面，Q 为 xz 平面。求 $S\cap T$，$S+T$，$P\cap Q$，$P+Q$，$S+P$，并判断 $S\cup P$ 与 $P\cup Q$ 是不是子空间。

6.3 ★ 设 $E=\begin{bmatrix}1&2\\2&5\\0&1\end{bmatrix}$。用消元求 $b\in C(E)$ 的条件。$b=(1,3,1)$ 与 $b=(1,0,0)$ 是否在 $C(E)$ 中？可解时求出解，并说明解是否唯一。

6.4 ★ 求 $A=\begin{bmatrix}1&2&-1\\2&-1&1\end{bmatrix}$ 的零空间，并用两个过原点平面的交描述它。

6.5 ★★ (a) 对主例子的 4×3 矩阵 A，判断 $b=(0,1,2,3)$ 是否在 $C(A)$ 中；若在，写出 $Ax=b$ 的全部解。(b) 写出平面 $x+2y+3z=6$ 上全部点的形式 $x_p+y\,v+z\,w$，并说明 v,w 张成的是什么。

6.6 ★★ 构造一个 3×3 矩阵 A，使 $C(A)$ 含 $(1,0,0)$ 与 $(0,1,0)$，且 $N(A)$ 含 $(1,1,1)$。你的 A 的 $C(A)$ 与 $N(A)$ 各是什么？

6.7 ★★ 设 A、B 行数相同。证明 $C\left(\begin{bmatrix}A&B\end{bmatrix}\right)=C(A)+C(B)$。（这是命题 6.12 的「列版本」：左右并排对应列空间的和，上下叠放对应零空间的交。）

6.8 ★★ 证明：$N(A)\subseteq N(BA)$，$C(AB)\subseteq C(A)$（假定乘积有定义）。各举一例说明包含可以是严格的。

6.9 ★★ 判断真假，假的给出反例：(a) 存在 $b \neq 0$，使 $Ax = b$ 的解集是子空间；(b) 若 $C(A) = \{0\}$，则 A 是零矩阵；(c) $N(2A) = N(A)$；(d) $C(A+B) = C(A) + C(B)$。

6.10 ★★★ 设 $N = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$。(a) 求 $C(N)$ 与 $N(N)$，验证 $N^2 = 0$。(b) 证明：对 $n \times n$ 矩阵 A，$A^2 = 0$ 当且仅当 $C(A) \subseteq N(A)$。(c) 能否有 2×2 矩阵满足 $C(A) = N(A) = \mathbb{R}^2$？

练习答案与提示

6.1 (a) 是，条件 $x - 2y = 0$ 对任意线性组合保持成立。(b) 不是，零向量不满足坐标和为一。(c) 不是，$(0,0,1)$ 在其中，乘 -1 后不在。(d) 不是，$(1,1,0)$ 和 $(1,-1,0)$ 都满足条件，其和 $(2,0,0)$ 不满足。(e) 是，全部向量为 $t(1,1,1)$，任意线性组合仍是这个形式。

6.2 $S \cap T = \{0\}$，$S + T = P$。$P \cap Q = S$，因为交中同时有 $z = 0, y = 0$；$P + Q = \mathbb{R}^3$，因为 $(x,y,z) = (x,y,0) + (0,0,z)$。$S + P = P$。$S \cup P = P$ 是子空间，$P \cup Q$ 不是：$(0,1,0)$ 与 $(0,0,1)$ 的和不属于任一平面。

6.3 方程为 $x_1 + 2x_2 = b_1$、$2x_1 + 5x_2 = b_2$、$x_2 = b_3$。消去 x_1 得 $x_2 = b_2 - 2b_1$，所以可解当且仅当 $b_3 = b_2 - 2b_1$；满足时 $x = (b_1 - 2b_3, b_3)$。第一个右端满足条件，解为 $(-1,1)$；第二个不满足 $0 = 0 - 2$，无解。齐次时两个未知数都为零，所以可解时解唯一。

6.4 联立 $x_1 + 2x_2 - x_3 = 0$ 与 $2x_1 - x_2 + x_3 = 0$。两式相加得 $3x_1 + x_2 = 0$，故 $x_2 = -3x_1$，第一式再给 $x_3 = -5x_1$。所以 $N(A) = \operatorname{span}\{(1,-3,-5)\}$，是两个过原点平面的交线。代回两式分别为 $1 - 6 + 5 = 0$ 和 $2 + 3 - 5 = 0$。

6.5 (a) 右端为等差数列，取特解 $(1,-1,0)$。全部解为 $(1,-1,0) + t(1,1,-1)$，$t \in \mathbb{R}$。(b) $x = 6 - 2y - 3z$，所以全部点为 $(6,0,0) + y(-2,1,0) + z(-3,0,1)$。两个方向向量张成齐次平面 $x + 2y + 3z = 0$，原平面是其平移。

6.6 可取 $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$。前两列就是所需向量，三列之和为零。所有输出为 $(x_1 - x_3, x_2 - x_3, 0)$，故 $C(A)$ 是 xy 平面。齐次方程要求 $x_1 = x_3, x_2 = x_3$，所以 $N(A) = \operatorname{span}\{(1,1,1)\}$。

6.7 把输入分块为 $(u,v)^T$，则 $[A\ B](u,v)^T = Au + Bv$。每个输出是 $C(A)$ 中向量与 $C(B)$ 中向量之和，故属于右侧。反过来，每个 $a + b$ 可选 u, v 满足 $a = Au, b = Bv$，由同一分块输入产生，故属于左侧。两个包含方向给出相等。

6.8 若 $Ax = 0$，则 $BAx = B0 = 0$，所以 $N(A) \subseteq N(BA)$。若 $y = ABx$，则 $y = A(Bx) \in C(A)$。第一式可取 $A = I_2, B = 0$，从 $\{0\}$ 严格包含到 $\mathbb{R}^2$；第二式也取 $A = I_2, B = 0$，从 $C(AB) = \{0\}$ 严格包含于 $C(A) = \mathbb{R}^2$。

6.9 (a) 假，$b \neq 0$ 时零向量从来不是解。(b) 真，每一列都是列空间中的向量，列空间只有零向量就要求各列全零。(c) 真，$2Ax = 0$ 当且仅当 $Ax = 0$，因为实数 2 非零。(d) 假，取 $A = I_2, B = -I_2$，则左侧为 $\{0\}$，右侧为 $\mathbb{R}^2$。相加时列可能抵消，不能把两个独立输入强制为同一个输入后仍断言输出相等。

6.10 (a) 两列互为相反数，所以 $C(N) = \operatorname{span}\{(1,1)\}$；齐次方程为 $x_1 - x_2 = 0$，故 $N(N)$ 也是同一直线。直接乘得 $N^2 = 0$。(b) 若 $A^2 = 0$，对任意 $y = Ax$ 有 $Ay = A^2x = 0$，所以 $y \in N(A)$。反过来若每个 Ax 都属于 $N(A)$，则对每个 x 有 $A^2x = 0$；特别取各单位向量，说明 A^2 的每列为零，故 $A^2 = 0$。(c) 不可能。$N(A) = \mathbb{R}^2$ 要求每个输入都输出零，所以 $A = 0$，进而列空间只有 $\{0\}$。

第 7 讲

齐次方程：主元变量、自由变量与特殊解

Solving $Ax = 0$: Pivot Variables and Special Solutions

对应材料：MIT 18.06SC Session 1.7 课时摘要 Solving Ax = 0: pivot variables, special solutions；习题 7.1、7.2 及解答。

本讲目标

(1) 会在非方阵和奇异矩阵中继续消元，辨认主元列与自由列。

(2) 能逐个给自由变量赋值，回代得到全部解，说明为什么没有遗漏。

(3) 会构造特殊解，证明它们能张成零空间且彼此线性无关。

(4) 会从简化行阶梯形直接读出零空间，正确处理列重排后的变量次序。

(5) 能用特殊解解释矩阵列之间的关系，并判断矩阵乘积何时可能为零。

第 6 讲把零空间定义为 $N(A) = \{\, x \in \mathbb{R}^n \mid Ax = 0 \,\}$，并证明它是子空间。现在的问题更具体：给出一个矩阵，怎样用有限次计算，把这个集合中的全部向量写出来？只找到一个非零解还不够，因为别的方向可能没有被找到；把四个未知数逐一猜值也没有系统性。本讲用消元确定哪些未知数可以任意选，再把这几项选择变成一组固定的生成向量。

全文默认实数矩阵，x 为列向量。为了节省横向空间，(x_1, x_2, x_3, x_4) 在文字中仍表示列向量，不表示行向量。

7.1 第二列没有主元，还能消元吗？

考虑原课摘要中的矩阵

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}.$$

这是一个 3×4 矩阵，方程 $Ax = 0$ 有四个未知数。第 2 讲用过的消元仍能进行：先用第一行消去第一列下面的数，作 $R_2 \leftarrow R_2 - 2R_1$、$R_3 \leftarrow R_3 - 3R_1$，得到

$$A \longrightarrow \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix}.$$

此时第二列在第二行及以下全为零。若坚持寻找第二个对角元作主元，就会停住；但第三列还有数 2，它能继续消去下一行。作 $R_3 \leftarrow R_3 - R_2$，有

$$U = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

最后一行是方程 $0 = 0$，不会增加限制。原来的第三行本来就是第一行与第二行之和，所以它在消元后变成零行有明确的代数原因。

定义 7.1（阶梯形、主元列与自由列）　一个矩阵处于**行阶梯形**（row echelon form），指零行都在非零行下方，每个非零行的第一个非零元比上一非零行的第一个非零元更靠右。每行第一个非零元称为**主元**（pivot）；包含主元的列称为**主元列**（pivot column），其余列称为**自由列**（free column）。消元中主元的个数称为矩阵的**秩**（rank），记作 $r = \operatorname{rank} A$。

这里的「主元列」说的是消元确定的列位置：主例的位置为第 $1, 3$ 列，因而原矩阵对应的第 $1, 3$ 列也叫主元列。原矩阵各列是否与别列相关，将由齐次解进一步解释。第 9 讲会用基与维数说明秩与消元路线无关；本讲使用从左到右选主元的消元程序。矩阵 U 的主元是位置 $(1,1)$ 的 1 与位置 $(2,3)$ 的 2，所以 $r = 2$。第 $2, 4$ 列是自由列。对比矩阵 $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$，它虽有零行在底部，却不满足主元向右移动的要求，交换前两行后才是阶梯形。零矩阵没有非零行，满足定义，其主元个数为零。

图 7.1 把主例消元后每个变量的角色标出。蓝色格中的主元随着行号增加向右移动，橙色列不含主元。

x_1	x_2	x_3	x_4
①	2	2	2
0	0	②	4
0	0	0	0
主元	自由	主元	自由

图 7.1　主元可以跳过一列，未选为主元列的位置对应自由变量。

这张图也说明，四个未知数中只有两个被主元方程直接解出。主元列未必相邻，也未必都位于左上角；消元程序依据每行第一个尚未消去的非零元决定下一步。寻找主元时，需要查看当前行及其下方的整列：当前格为零时还可能通过换行找到主元，只有下方全零才跳过这列。

■ **例 7.1　从阶梯形读出约束**　对上面的 U，写出 $Ux = 0$ 的两个非零方程，并说明哪个变量可以先任选。

解　思路是先看每一行的主元位置，再从最后一个主元行向上回代。两行给出

$$2x_3 + 4x_4 = 0, \qquad x_1 + 2x_2 + 2x_3 + 2x_4 = 0.$$

第 2 列、第 4 列是自由列，因此先取 x_2, x_4。第一式给 $x_3 = -2x_4$；代回第二式得 $x_1 = -2x_2 + 2x_4$。所以

$$x = (-2x_2 + 2x_4,\ x_2,\ -2x_4,\ x_4).$$

这里没有把 x_1, x_3 也任意指定，因为它们的值由两个主元方程确定。代入 U 的第一行和第二行都得到 0，所以写出的每个向量确实是解。 ♦

想一想　如果某一列全为零，它对应的变量是否一定自由？

答　是。该变量在每个方程中的系数都为零，改变它不会改变 Ux。但「自由」不等于「原矩阵的这一列为零」：行消元会改变行的数值，判断自由性要看阶梯形中的主元位置，而列之间的关系要回到 $Ax = 0$ 检查。

易错提醒　「有 n 个变量，就有 n 个主元」是错误的。上面的矩阵有 4 个变量，却只有 2 个主元；而 $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ 有 2 个主元、3 行。正确的上界是 $r \le \min(m, n)$，没有主元的列给出自由变量。

7.2　自由变量如何产生全部解？

上一节只写出了一个含参数的向量。现在把「全部」这件事拆开：任取自由变量的数值，会得到一个解；反过来，每个解的自由坐标本身就是那组参数。两句话合在一起，才是完整的解集描述。

定理 7.2（行操作保持齐次解集）　若矩阵 B 由 A 经过有限次初等行变换得到，则 $Ax = 0$ 与 $Bx = 0$ 有完全相同的解集。

直观理解　一次行变换只是把原方程换成等价的方程组合。例如把第二行换成「第二行减去第一行」，相当于把两个等式 $0 = 0$ 做相同的线性组合，等号两侧仍同时变化；可逆的行变换还可以恢复原方程。

思路　证明只需检查三种初等行变换。每一种变换都能由同类变换或非零倍数恢复，所以先证明 $Ax = 0$ 推出 $Bx = 0$，再用逆变换得到反向包含。

证明　交换两行只改变方程的排列；某行乘非零数后，0 仍是右端；某行加上另一行的倍数，是把两个等式相加。三种操作都把一个成立的齐次方程组变成成立的方程组，因此 $Ax = 0 \Rightarrow Bx = 0$。交换行可逆，非零倍数可除回去，加倍行可减回去；沿逆操作得到 $Bx = 0 \Rightarrow Ax = 0$。 ■

■ **例 7.2　第一列也可能是自由列**　求 $Zx = 0$ 的全部解，其中 $Z = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 4 & 7 \end{bmatrix}$。

解　第二行减去第一行的 2 倍，得到 $[0\ 0\ 0\ 1]$；再把第一行减去 3 倍的新第二行，得到 $[0\ 1\ 2\ 0]$。于是主元列为第 2、4 列，自由变量为 x_1, x_3，方程为 $x_4 = 0$、$x_2 + 2x_3 = 0$。令 $x_1 = s, x_3 = t$，全部解是

$$x = (s, -2t, t, 0) = s(1, 0, 0, 0) + t(0, -2, 1, 0).$$

直接乘回原矩阵：$Z(1,0,0,0)^T = 0$，$Z(0,-2,1,0)^T = 0$。这个例子提醒我们，主元列的编号不能靠「第 1 列、第 2 列」猜，而要在实际消元后读取。 ♦

图 7.2 用最简单的方程 $x_1 + 2x_2 = 0$ 展示同一种机制。这里 $x_2 = t$ 可以任取，$x_1 = -2t$ 随之确定；每一条水平虚线与解直线只有一个交点。

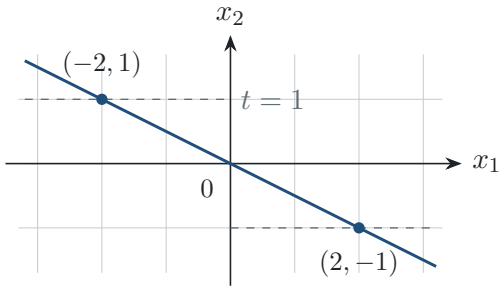


图 7.2　每个自由坐标 $x_2 = t$ 唯一确定直线上的一个解 $(-2t, t)$。

图中的自由变量并不表示一个独立于方程之外的新量。它就是原向量的第二个坐标，只是我们把它选为描述整条直线的参数。若错误地让 x_1, x_2 同时任取，会得到整个平面；若把 x_2 限制为零，只留下原点。正确的参数选择既不能多加独立选择，也不能漏掉本来可以自由变化的坐标。

从下往上读取解　先在阶梯形中标出主元列；给所有非主元坐标任意实数；从最后一个非零行开始回代，每行只解出一个主元变量。遇到零行 $0 = 0$ 时不增加条件。

对齐次方程而言，不会产生 $0 = c \neq 0$ 的矛盾行，因为行变换始终把零右端变成零。第 8 讲处理非零右端时必须额外检查可解性；本讲只需把已有的自由度完整保留下来。

7.3　特殊解为什么足以描述零空间？

回到主例，把 $x_2 = s, x_4 = t$ 代入得到

$$x = (-2s + 2t, s, -2t, t) = s\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}.$$

右侧只有两个固定向量；任意解都通过改变前面的两个系数得到。它们是如何被挑出来的？第一个把自由坐标设为 $(s, t) = (1, 0)$，第二个设为 $(s, t) = (0, 1)$。每次只打开一个自由坐标，其他自由坐标全部设为零。

定义 7.3（特殊解）　求齐次方程 $Ax=0$ 时，选定一个自由变量令其等于 1，其余自由变量均令其等于 0，再回代求出主元变量，所得向量称为一个**特殊解**（special solution）。每个自由变量给出一个特殊解。

原课的 special solution 在这里统一译作「特殊解」，以免与非齐次方程的「一个特解」混淆。非齐次特解只要求满足给定右端，未必采用上述自由坐标单位化的规则。

主例的两个特殊解记为 $s_1=(-2,1,0,0)$、$s_2=(2,0,-2,1)$。它们在自由坐标 x_2,x_4 上分别留下 $(1,0)$、$(0,1)$，所以 s_1 不可能是 s_2 的倍数。任取 s,t 后，组合 ss_1+ts_2 的自由坐标恰好为 (s,t)；这组坐标既能恢复系数，又能识别得到的是哪个解。

定理 7.4（特殊解给出全部解）　设 A 为 $m\times n$ 矩阵，消元后有 r 个主元。按自由列从左到右构造 $n-r$ 个特殊解 $s_1,\dots,s_{n-r}$。则

$$N(A)=\{\,c_1s_1+\cdots+c_{n-r}s_{n-r}\mid c_1,\dots,c_{n-r}\in\mathbb{R}\,\}.$$

而且每个解只有一种这样的表示；等价地，$s_1,\dots,s_{n-r}$ 线性无关。若没有自由列，结论解释为 $N(A)=\{0\}$。

直观理解　每个特殊解负责一个自由坐标。给这些特殊解乘上所需的自由坐标再相加，就逐个填好了全部自由位置；主元位置随后由方程唯一确定，因此不可能还差另一个解。

思路　先证任意线性组合仍是解，再证每个解都能由这组组合得到，最后读取自由坐标证明系数唯一。这里无需预先引用零空间维数公式。

证明　**组合仍为解。**每个 s_j 满足 $As_j=0$，所以由矩阵乘法的线性性，$A(\sum_j c_js_j)=\sum_j c_jAs_j=0$。

每个解都在组合中。设 x 是任意解，其自由坐标按顺序为 $c_1,\dots,c_{n-r}$。令 $y=\sum_j c_js_j$。由特殊解的构造，y 的各自由坐标与 x 相同。于是 $z=x-y$ 是齐次解，且所有自由坐标为零。从最后一个主元行开始回代：这个主元之后的自由坐标为零，更靠后的主元坐标也已依次解成零，所以当前主元乘以当前坐标为零。主元非零，当前坐标只能为零。向上逐行重复，得到所有主元坐标也为零，即 $z=0$、$x=y$。

系数唯一。若 $\sum_j d_js_j=0$，看第 k 个自由位置，只有 s_k 在此位置为 1，其他特殊解在此位置为零，因此这个位置的坐标为 $d_k=0$。每个 k 都如此，全部系数为零。若两个组合相等，移项后用同一论证可得两组系数逐个相等。 ■

定理中「一个自由变量取 1，其余取 0」保证了自由坐标构成单位阵。若只随意选两个齐次解，例如 s_1 与 $2s_1$，虽然每个都正确，却只能得到一条直线，不能描述含有 s_2 的全部零空间。零向量也永远是解，但把它当作唯一生成向量，只能描述零空间恰为 $\{0\}$ 的情形。

■ **例 7.3　从自由坐标恢复系数**　对主例 A，判断 $v=(-8,3,2,-1)$ 是否为零空间中的向量，并求它关于 s_1,s_2 的表示。

解　直接读取自由坐标可知，若表示存在，其系数必须是 $x_2 = 3$、$x_4 = -1$。计算

$$3s_1 - s_2 = 3(-2,1,0,0) - (2,0,-2,1) = (-8,3,2,-1) = v.$$

所以 v 确实在零空间中。原方程三行给出 $-8+6+4-2 = 0$、$-16+12+12-8 = 0$、$-24+18+16-10 = 0$，再次验证了这一点。唯一性来自自由坐标，不能再有另一对系数表示同一个向量。 ♦

图 7.3 将特殊解的组合关系画在二维参数平面中。它所表示的不是原空间中向量的所有坐标，而是自由坐标 (x_2, x_4) 如何决定整个解向量。

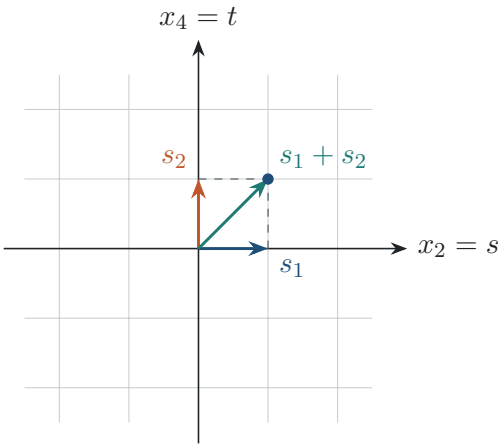


图 7.3　自由坐标平面中点 (s,t) 对应零空间向量 $ss_1 + ts_2$。

图中的水平、竖直方向表示两个自由坐标各改变一单位，映到解空间就是分别加上 s_1、s_2。若画原来的四维向量空间，几何对象会是一张二维平面；坐标平面直接标出「系数就是自由坐标」，更适合检查定理中的存在性和唯一性。

想一想　若只知道 s_1, s_2 都满足 $As = 0$，能否立刻断定它们生成全部零空间？

答　不能。还需要检查所有解都能表示出来；两向量本身是解只保证它们的线性组合仍是解。主例通过自由坐标恢复系数来证明张成；若两个解成比例，它们甚至不能生成二维零空间。还需要论证表示唯一，才能得出线性无关。

独立性与张成性各看一件事　要证特殊解线性无关，读取自由坐标：$\sum c_j s_j = 0$ 立即给出每个 $c_j = 0$。要证它们张成全部解，读取任意解的自由坐标，用这些坐标作系数，再靠回代唯一性补齐主元坐标。

7.4　简化行阶梯形怎样直接给出零空间？

在主例中继续消元：第二行除以 2，第一行减去第二行的 2 倍，得到

$$R = \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

每个主元为 1，主元上方与下方均为零，称为**简化行阶梯形**（reduced row echelon form）。由此直接读出 $x_1 = -2x_2 + 2x_4$、$x_3 = -2x_4$，与此前回代结果一致。

把主元列排在前面、自由列排在后面，会更容易看清公式。主例应同时重排变量：

$$y = (x_1, x_3, x_2, x_4)^T.$$

这样方程的非零部分变为

$$\begin{bmatrix} 1 & 0 & 2 & -2 \\ 0 & 1 & 0 & 2 \end{bmatrix} y = 0.$$

分块记作 $[I\ F]y = 0$，其中 $I = I_2$、$F = \left[\begin{smallmatrix} 2 & -2 \\ 0 & 2 \end{smallmatrix}\right]$。令自由坐标 $z = (x_2, x_4)^T$，主元坐标必须是 $-Fz$，所以

$$y = \begin{bmatrix} -F \\ I_2 \end{bmatrix} z = \begin{bmatrix} -2 & 2 \\ 0 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} z.$$

这个矩阵的各行仍按 x_1, x_3, x_2, x_4 排列。恢复原变量次序后，才得到主例的零空间矩阵

$$N = \begin{bmatrix} -2 & 2 \\ 1 & 0 \\ 0 & -2 \\ 0 & 1 \end{bmatrix}, \qquad x = Nz, \quad AN = 0.$$

易错提醒　列重排改变了变量次序。把原矩阵的列换位后，若仍把坐标按原顺序读出，就会把主元变量与自由变量混在一起。例如上面的分块矩阵第一列是 $(-2, 0, 1, 0)$，按原变量顺序代回 A 得不到零；它必须先恢复成 $(-2, 1, 0, 0)$。行变换保持原变量顺序，列重排则必须同步记录坐标换位。

命题 7.5（分块公式）　将简化行阶梯形的列与变量同步重排，得到 $\widetilde{R} = \left[\begin{smallmatrix} I_r & F \\ 0 & 0 \end{smallmatrix}\right]$。若 $q = n - r > 0$，则全部解为

$$y = \widetilde{N} z, \qquad \widetilde{N} = \begin{bmatrix} -F \\ I_q \end{bmatrix}, \qquad z \in \mathbb{R}^q.$$

证明　把变量分成主元坐标 $y_p \in \mathbb{R}^r$ 与自由坐标 $z \in \mathbb{R}^q$，非零行给出 $y_p + Fz = 0$，即 $y_p = -Fz$，零行不加限制。因此任取 z 都给出解，任意解也必有这个形式。矩阵乘法验证 $\widetilde{R}\widetilde{N} = [-F + F; 0] = 0$，其列正是逐个令自由坐标为单位向量所得的特殊解。 ∎

公式的维数需要对齐：F 为 $r \times q$，$-F$ 上面有 r 行，I_q 下面有 q 行，因此 $\widetilde{N}$ 总计 n 行、q 列。它将自由坐标向量映到原来的解向量空间。恢复变量次序只重排这些行，不改变特殊解的个数。

■ **例 7.4　换一组系数检验特殊解**　对 $Q=\begin{bmatrix}1&2&1&4\\2&4&3&9\end{bmatrix}$，求简化行阶梯形和全部特殊解。

解　先作 $R_2\leftarrow R_2-2R_1$，再作 $R_1\leftarrow R_1-R_2$，得到

$$R_Q=\begin{bmatrix}1&2&0&3\\0&0&1&1\end{bmatrix}.$$

主元列是第 1、3 列，令自由变量 $x_2=s,x_4=t$，则 $x_1=-2s-3t$、$x_3=-t$。所以

$$x=s(-2,1,0,0)+t(-3,0,-1,1).$$

两条特殊解代回 Q 都得到零向量；每个解的自由坐标又唯一给出系数 s,t。 ♦

■ **例 7.5　从指定生成向量反过来造方程**　构造一个 2×4 矩阵 C，使其零空间恰好为

$$\operatorname{span}\{(2,2,1,0),(3,1,0,1)\}.$$

解　先看第三、第四坐标：给出的两个向量在那里组成单位阵，所以可把它们设为自由坐标。若 $x=s(2,2,1,0)+t(3,1,0,1)$，则 $x_3=s,x_4=t$，主元坐标满足 $x_1=2x_3+3x_4$、$x_2=2x_3+x_4$。因此取

$$C=\begin{bmatrix}1&0&-2&-3\\0&1&-2&-1\end{bmatrix}.$$

任意给定组合都满足 $Cx=0$，反过来 $Cx=0$ 恰好要求上述两条关系，所以每个齐次解都是给定的组合。两向包含说明零空间恰为所需集合。单独验算两个向量被 C 消灭只证明一个包含方向，仍需反向解方程才能证明没有额外的解。 ♦

■ **例 7.6　单个方程给出一张平面**　求 $x_1+2x_2+3x_3=0$ 的全部解，并用零空间矩阵表示。

解　系数矩阵 $H=[1\ 2\ 3]$ 已经是简化行阶梯形。主元变量为 x_1，自由变量为 $x_2=s,x_3=t$，所以

$$x=s(-2,1,0)+t(-3,0,1),\qquad N_H=\begin{bmatrix}-2&-3\\1&0\\0&1\end{bmatrix}.$$

检查 $HN_H=[-2+2\ -3+3]=[0\ 0]$。更完整地说，任取 s,t 都给出解，每个解的第二、第三坐标又把 s,t 唯一恢复。这是一张经过原点的二维平面。 ♦

前后联系　第 6 讲把零空间写成某个矩阵的列空间，本节给出了计算这个矩阵的通法：$N(A)=C(N)$。这里 $N(A)$ 是集合，N 是以特殊解为列的矩阵，不能把两个同名符号当作同一个对象。第 9 讲会把这组独立且张成全部解的特殊解称为零空间的一组基，并用 $\dim N(A)=n-r$ 记录其维数。

7.5　零矩阵、满列主元与宽矩阵

如果一个矩阵的自由变量很多，零空间就需要很多特殊解来描述。相反，主元占满每一列时，没有可以任取的坐标，回代只能得到零解。把这两个极端先算清楚，可以检查公式中的 $n-r$ 是否使用正确。

■ **例 7.7　两个极端**　分别求 $O=\begin{bmatrix}0&0\\0&0\end{bmatrix}$ 与 $D=\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}$ 的零空间。

解　对 O，两个方程都是 $0=0$，没有主元，$r=0$。两个坐标都自由，所以特殊解是 $(1,0)$、$(0,1)$，所有解为 $s(1,0)+t(0,1)$，即 $N(O)=\mathbb{R}^2$。

对 D，第二行减去第一行的 3 倍，得到 $[0\ -2]$；第三行减去第一行的 5 倍，得到 $[0\ -4]$。再用第二行消去第三行，得到两列都有主元的阶梯形。齐次方程从下往上给出 $-2x_2=0$、$x_1+2x_2=0$，所以 $x_2=0, x_1=0$。因此 $N(D)=\{0\}$，没有特殊解需要列出。这里零空间包含一个向量，但生成它需要的向量个数为零。　♦

「没有特殊解」并不表示方程无解。零向量永远满足齐次方程；没有自由变量表示没有非零解。前一个矩阵则让所有输入都被送到零，零空间占满整个输入空间。这个对照还能用来区分方程条数与有效约束数：D 虽有三条方程，第三条在消元后消失；只需要两个主元就能限制住两个输入坐标。

> **命题 7.6（齐次解的三个判断）**　设 A 为 $m\times n$ 矩阵，r 为消元中的主元个数。
>
> (1) $Ax=0$ 只有零解，当且仅当 $r=n$，即每列都有主元。
>
> (2) $Ax=0$ 有非零解，当且仅当 $r<n$，即至少有一个自由变量。
>
> (3) 若 $n>m$，则 $Ax=0$ 必有非零解。

证明　若 $r=n$，没有自由坐标，从最后一个主元行回代，所有坐标依次为零。若 $r<n$，任选一个自由变量置 1，得到的特殊解在该坐标非零，所以它是非零解。这证明前两条。若 $n>m$，主元每行至多一个，因此 $r\le m<n$，再用第二条得到第三条。　■

第三条只适用于未知数多于方程的齐次系统。若右端非零，一个零行可能变成矛盾方程，连一个解都没有；若 $n\le m$，则是否有非零齐次解需要实际看主元，不能只数矩阵的行列。例如 I_2 没有非零齐次解，而 $\left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]$ 有非零解 $(-1,1)$，两者都是方阵。

■ **例 7.8　只有一个有效方程**　求 $G=\begin{bmatrix}1&3&-2&4\\2&6&-4&8\\-1&-3&2&-4\end{bmatrix}$ 的秩和全部齐次解。

解　第二行是第一行的 2 倍，第三行是第一行的 -1 倍，消去后只剩 $x_1+3x_2-2x_3+4x_4=0$。

所以 $r=1$，自由变量为 x_2,x_3,x_4。依次让它们取单位坐标，得

$$s_1=(-3,1,0,0),\quad s_2=(2,0,1,0),\quad s_3=(-4,0,0,1).$$

全部解为 $x=as_1+bs_2+cs_3$。直接读取自由坐标得 $a=x_2,b=x_3,c=x_4$，而方程要求 $x_1=-3a+2b-4c$，所以没有别的解。验算第一行分别给出 $-3+3=0$、$2-2=0$、$-4+4=0$；其他行是它的倍数，乘积也为零。矩阵有三行，却只有一个有效约束，自由变量个数是 $4-1=3$，不能用 $4-3$ 代替。 ♦

想一想　若一个 3×5 矩阵的秩为 2，特殊解有几个？是否所有解都能用两个向量描述？

答　有 $5-2=3$ 个特殊解，每个自由变量给出一个。两向量至多对应自由坐标空间中的两个独立方向，无法产生三个自由坐标的任意选择；例如三个位于不同自由位置的单位坐标不可能都由两个固定坐标向量生成。准确的维数论证在第 9 讲给出，本讲已通过特殊解的自由坐标单位阵证明了三个生成方向彼此独立。

主元数、变量数与零行数分别回答不同的问题。对于 $m\times n$ 矩阵，非零行数是 r，零行数是 $m-r$；自由变量数是 $n-r$。主例中 $m=3,n=4,r=2$，因此只有一个零行，却有两个自由变量。把零行个数当成特殊解个数，是把输出方程所在的 $\mathbb{R}^m$ 与输入变量所在的 $\mathbb{R}^n$ 混淆了。

7.6　原课习题：消元结果与列关系互相核对

原课习题 7.1 把相同的算法换成分数主元。它适合检查两件事：分数是否算对，以及第三行消失后是否仍把它误当作一个约束。

■ **例 7.9　原课习题 7.1**　求矩阵 P 的简化行阶梯形、秩与特殊解，其中

$$P=\begin{bmatrix}1&5&7&9\\0&4&1&7\\2&-2&11&-3\end{bmatrix}.$$

解　先用第一行消去第三行第一列：

$$R_3\leftarrow R_3-2R_1:\qquad \begin{bmatrix}1&5&7&9\\0&4&1&7\\0&-12&-3&-21\end{bmatrix}.$$

现在第三行是第二行的 -3 倍，可直接作 $R_3\leftarrow R_3+3R_2$ 得到零行。第二行除以 4，再用它消去第一行第二列的 5，得到

$$R_P=\begin{bmatrix}1&0&23/4&1/4\\0&1&1/4&7/4\\0&0&0&0\end{bmatrix}.$$

检验两个分数：第一行第三列是 $7-5/4=23/4$，第四列是 $9-35/4=1/4$。主元列为第 1、2 列，所以 $r=2$；自由列为第 3、4 列。

令 $x_3 = s, x_4 = t$，由简化阶梯形得到

$$x_1 = -\frac{23}{4}s - \frac{1}{4}t, \qquad x_2 = -\frac{1}{4}s - \frac{7}{4}t.$$

特殊解为

$$u_1 = \begin{bmatrix} -23/4 \\ -1/4 \\ 1 \\ 0 \end{bmatrix}, \qquad u_2 = \begin{bmatrix} -1/4 \\ -7/4 \\ 0 \\ 1 \end{bmatrix}.$$

全部解是 $su_1 + tu_2$。若只想避免分数，也可以取整数向量

$$4u_1 = (-23, -1, 4, 0), \qquad 4u_2 = (-1, -7, 0, 4).$$

它们张成同一个零空间。但按本讲定义的「特殊解」要求自由坐标为单位坐标，所以整数倍向量是另一组生成向量，不能在不说明的情况下替换特殊解的标准化结果。

验算整数倍向量更方便：第一向量乘原矩阵三行给 $-23-5+28=0$、$-4+4=0$、$-46+2+44=0$；第二向量分别给 $-1-35+36=0$、$-28+28=0$、$-2+14-12=0$。这些验算还说明分数结果没有在回代时写错符号。 ♦

一个齐次解也记录一条列关系。设 P 的列为 p_1, p_2, p_3, p_4，$Pu_1 = 0$ 可写成

$$-\frac{23}{4}p_1 - \frac{1}{4}p_2 + p_3 = 0, \qquad p_3 = \frac{23}{4}p_1 + \frac{1}{4}p_2.$$

另一个特殊解给出 $p_4 = \frac{1}{4}p_1 + \frac{7}{4}p_2$。这两条关系使自由列能由主元列产生。主元列自身没有非平凡关系：若 $ap_1 + bp_2 = 0$，对应解的自由坐标全为零，由回代可得 $a = b = 0$。第 9 讲会将「既生成全部列，又彼此独立」归纳为列空间的基；本讲从算出的零空间即可看到这两点。

■ **例 7.10　原课习题 7.2：非零乘积也能为零**　给定 $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$，找 2×2 矩阵 A_1, A_2，使 $\operatorname{rank}(A_1B) = 1$、$\operatorname{rank}(A_2B) = 0$。第二个矩阵能否也非零？

解　先看 B 的两列，都是向量 $b = (1, 1)$。乘积 A_iB 的两列因此都是 A_ib。若 $A_ib \neq 0$，两列相同且非零，秩为 1；若 $A_ib = 0$，乘积为零矩阵，秩为零。

保留 b 的最直接选择是 $A_1 = I_2$，此时 $A_1B = B$，一个主元。消灭 b 可选

$$A_2 = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \qquad A_2B = \begin{bmatrix} 1-1 & 1-1 \\ 1-1 & 1-1 \end{bmatrix} = 0.$$

A_2 自身非零，但其零空间包含 b，于是它把 B 的每一列都送到零。这个例子不需要把某个因子取成零矩阵；矩阵乘积为零只要求右边矩阵的每列都在左边矩阵的零空间中。 ♦

这一结论也可从列空间表述为 $A_2B = 0$ 当且仅当 $C(B) \subseteq N(A_2)$。一个方向是每列被消灭后任意列组合仍被消灭；另一个方向是每列本身就在列空间中。注意这里的零空间属于左因子，列空间属于右因子，二者处在同一个中间坐标空间，维数才相容。若交换乘法顺序，需要重新检查另一个包含关系。

用三种方式复核零空间　计算结束后，先把特殊解代回原矩阵检查；再看自由坐标是否组成单位阵；最后说明任意解的自由坐标能恢复组合系数。三项分别检查「确实是解」「彼此独立」「没有遗漏」。

本讲小结

- **行变换保持零空间**：初等行变换可逆，因此 $N(A)=N(U)=N(R)$；列重排必须同步重排变量。
- **主元决定角色**：r 个主元给出 r 个主元变量，其余 $n-r$ 个变量自由。零行数为 $m-r$，不要与自由变量数混淆。
- **特殊解**：每次令一个自由变量为 1、其他自由变量为 0，回代得到一个特殊解；共得 $n-r$ 个。
- **全部解**：$x=c_1s_1+\cdots+c_{n-r}s_{n-r}$，系数就是对应的自由坐标。自由坐标单位阵同时证明独立性与表示唯一。
- **分块公式**：按主元、自由变量同步重排后，$\widetilde{R}=[I_r\ F;0\ 0]$，全部解为 $y=[-F;I_{n-r}]z$；使用前须核对变量次序。
- **非零解判据**：$N(A)=\{0\}$ 当且仅当 $r=n$；若 $n>m$，必有自由变量和非零齐次解。
- **乘积为零**：$AB=0$ 当且仅当 B 的每列属于 $N(A)$，等价于 $C(B)\subseteq N(A)$。

练习

7.1 ★　对阶梯形 $U=\begin{bmatrix}1&2&2&4&6\\0&0&1&2&3\\0&0&0&0&0\end{bmatrix}$，求主元列、自由列、秩及全部特殊解。

7.2 ★　分别求下列矩阵的秩和零空间：
(a) $\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}$；(b) $\begin{bmatrix}1&1&1\\1&1&1\end{bmatrix}$；(c) $\begin{bmatrix}0&1\\0&0\end{bmatrix}$。

7.3 ★　判断真假，假的给反例：(a) 每个非零矩阵都只有零齐次解；(b) 零行个数等于自由变量个数；(c) 5×7 矩阵必有非零齐次解；(d) 行变换保持零空间。

7.4 ★★　求 $M=\begin{bmatrix}1&1&1&1\\1&2&3&4\\2&3&4&5\end{bmatrix}$ 的简化行阶梯形和全部齐次解，并将特殊解翻译成列关系。

7.5 ★★　设 $M(c)=\begin{bmatrix}1&2&1&c\\2&4&c&8\end{bmatrix}$。分 $c=2$ 与 $c\neq2$ 求主元列、特殊解。这个参数变化是否改变自由变量的个数？

7.6 ★★　对 $E=\begin{bmatrix}1&0&-2&1\\0&1&3&0\\1&1&1&1\end{bmatrix}$，求零空间矩阵 N，验证 $EN=0$。只验证 $EN=0$ 是否足以证明 N 的列给出全部解？

7.7 ★★　取 $A=\begin{bmatrix}2&-1\\4&-2\end{bmatrix}$、$B=\begin{bmatrix}1&2\\2&4\end{bmatrix}$。求 AB，并用列空间与零空间的包含解释结果。

7.8 ★★　已知 $F=\begin{bmatrix}1&0&2&1\\2&1&7&1\\3&1&9&2\end{bmatrix}$。求 $N(F)$，说明任意齐次解的组合系数从哪些坐标直接读取。

7.9 ★★★　证明：若 $N(A)=N(B)$，且二者都有 n 列，则对任意 n 维向量 x,y，$Ax=Ay$ 当且仅当 $Bx=By$。说明结论如何体现「相同的输入差被消灭」。

7.10 ★★★ 构造非零 3×3 矩阵 T，使 $N(T)=\operatorname{span}\{(1,1,1)\}$。说明仅使三列之和为零为什么还不够，并给出一个零空间更大的反例。

练习答案与提示

7.1 主元列为第 1、3 列，自由列为第 2、4、5 列，秩为 2。由第二行得 $x_3=-2x_4-3x_5$，代入第一行得 $x_1=-2x_2$。特殊解为 $(-2,1,0,0,0)$、$(0,0,-2,1,0)$、$(0,0,-3,0,1)$；全部解为它们的任意线性组合。

7.2 (a) 第二、三行分别减去第一行的 4、7 倍，得到 $[0\ -3\ -6]$、$[0\ -6\ -12]$。两条非零约束化为 $x_1-x_3=0$、$x_2+2x_3=0$，秩 2，零空间为 $\operatorname{span}\{(1,-2,1)\}$。(b) 只有一个非零独立行，秩 1，方程 $x_1+x_2+x_3=0$ 给出 $N=\operatorname{span}\{(-1,1,0),(-1,0,1)\}$。(c) 主元在第 2 列，秩 1，方程要求 $x_2=0$，故零空间为 $\operatorname{span}\{(1,0)\}$。

7.3 (a) 假，$[1\ 1]$ 的齐次解包括 $(-1,1)$。(b) 假，主例零行 1 个，自由变量 2 个；两者分别是 $m-r$ 与 $n-r$。(c) 真，因为 $r\le 5<7$，至少有 2 个自由变量。(d) 真，三种初等行变换均可逆，两个方向都保持齐次解。

7.4 第三行等于前两行之和。第二行减去第一行得 $[0\ 1\ 2\ 3]$，再清去第一行第二列，简化阶梯形为

$$\begin{bmatrix}1&0&-1&-2\\0&1&2&3\\0&0&0&0\end{bmatrix}.$$

全部解为 $s(1,-2,1,0)+t(2,-3,0,1)$。若列记为 m_j，关系分别为 $m_1-2m_2+m_3=0$、$2m_1-3m_2+m_4=0$。

7.5 第二行减去第一行的 2 倍，得 $[0\ 0\ c-2\ 8-2c]$。当 $c=2$，第 4 列出现主元，方程要求 $x_4=0$，自由变量为 x_2,x_3；特殊解为 $(-2,1,0,0)$、$(-1,0,1,0)$。当 $c\neq 2$，主元列为第 1、3 列，自由变量为 x_2,x_4，令 $x_4=t$ 得 $x_3=(2c-8)t/(c-2)$、$x_1=-2x_2+(8-c^2)t/(c-2)$。特殊解为 $(-2,1,0,0)$ 与 $((8-c^2)/(c-2),0,(2c-8)/(c-2),1)$。两种情形都有 2 个主元、2 个自由变量，只是自由位置不同。

7.6 第三行等于前两行之和，前两行已在简化阶梯形。自由变量为 $x_3=s,x_4=t$，全部解为 $s(2,-3,1,0)+t(-1,0,0,1)$，因此 $N=[2\ -1;-3\ 0;1\ 0;0\ 1]$。直接相乘得到零矩阵。仅验证乘积为零只证明其列生成的空间包含在零空间中，还要检查这些列确实能生成任意解。这里回代和自由坐标提供了所需的反向包含。

7.7 $AB=0$。B 的两列为 $(1,2)$、$2(1,2)$；$A(1,2)^T=(2-2,4-4)^T=0$，所以 $C(B)\subseteq N(A)$。两因子均非零，但右因子的每列都被左因子消灭。

7.8 第二行减去第一行的 2 倍得到 $[0\ 1\ 3\ -1]$；第三行减去第一行的 3 倍也得到同一行。故简化阶梯形为 $[1\ 0\ 2\ 1;0\ 1\ 3\ -1;0\ 0\ 0\ 0]$，零空间为 $\operatorname{span}\{(-2,-3,1,0),(-1,1,0,1)\}$。全部解写为 $s(-2,-3,1,0)+t(-1,1,0,1)$，系数从第三、第四坐标读取。

7.9 由线性性，$Ax=Ay$ 等价于 $A(x-y)=0$，等价于 $x-y\in N(A)$。由零空间相等，又等价于 $x-y\in N(B)$，即 $Bx=By$。因此两个矩阵未必有相同的输出，但它们把同一对输入识别为相同输出的条件一致：这些输入的差属于共同的零空间。

7.10 可取 $T=[1\ -1\ 0;0\ 1\ -1;1\ 0\ -1]$。前两行要求 $x_1=x_2=x_3$，第三行是它们的和，不增加限制，所以零空间恰为指定直线。仅使列和为零只保证 $(1,1,1)$ 在零空间中。例如 $[1\ -1\ 0;0\ 0\ 0;0\ 0\ 0]$ 的三列和也为零，但方程仅要求 $x_1=x_2$，x_3 自由，零空间为 $\operatorname{span}\{(1,1,0),(0,0,1)\}$，比目标更大。

第 8 讲

非齐次方程：相容条件、特解与全部解

Solving $Ax = b$: Row Reduced Form R

对应材料：MIT 18.06SC Session 1.8 课时摘要 Solving Ax = b: row reduced form R；习题 8.1、8.2、8.3 及解答。

本讲目标

(1) 会对增广矩阵同步消元，从零行读出右端向量的相容条件。

(2) 会找一个特解，并证明全部解恰好是特解加零空间。

(3) 能从简化行阶梯形读出主元变量与自由变量，写出没有遗漏的参数式。

(4) 能区分满列秩保证的唯一性与满行秩保证的存在性。

(5) 会依据秩判断解的个数，并用图解释解集的平移。

第 7 讲求出了齐次方程 $Ax = 0$ 的全部解。现在把右端换成一个指定的非零向量 b，原来的方法会多出两件事：先检查右端是否与方程之间的依赖关系相容，再决定解集相对零空间移动了多少。齐次方程总有零解，非齐次方程却可能一个解也没有；若有解，解集中的任意两点之差仍由第 7 讲的齐次方程控制。

本讲默认实数矩阵，向量均为列向量。文字中的坐标元组只为节省横向空间。始终把两个空间分清：x 属于输入空间 $\mathbb{R}^n$，而右端 b 属于输出空间 $\mathbb{R}^m$。相容条件约束的是 b，自由变量描述的是 x；它们出现在增广矩阵的不同位置。

8.1 零行怎样限制右端?

沿用上一讲的矩阵

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

第三行等于第一行与第二行之和。因此，若前三个方程都成立，把前两个等式相加，左边必然等于第三个方程的左边，右边也必须相等：$b_3 = b_1 + b_2$。例如 $b = (1, 5, 6)$ 有机会相容；换成 $(1, 5, 7)$，第三个方程要求同一个数既等于 6 又等于 7，所以无解。这里先得到的是必要条件，仍需说明条件成立时一定找得到解。

定义 8.1（相容方程组）　给定 $A \in \mathbb{R}^{m\times n}$ 和 $b \in \mathbb{R}^m$，若 $Ax = b$ 至少有一个解，就称方程组**相容**（consistent）；若没有解，就称**不相容**（inconsistent）。决定哪些 b 使方程组有解的等式称为相容条件。

方程 $x_1 + 2x_2 = 2$ 相容，因为 $(2, 0)$ 代入成立；方程组 $x_1 = 1$、$x_1 = 2$ 不相容，因为同一个实数不能同时取两个不同值。齐次方程 $Ax = 0$ 永远相容，因为 $x = 0$ 总满足它。相容只表示至少有一个解，尚未断言解是否唯一。

把右端和系数一起写成增广矩阵，按第 7 讲的路线作行变换。先作 $R_2 \leftarrow R_2 - 2R_1$、$R_3 \leftarrow R_3 - 3R_1$，再作 $R_3 \leftarrow R_3 - R_2$，得到

$$[A\ b] \longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & 2 & 2 & b_1 \\ 0 & 0 & 2 & 4 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 - b_2 - b_1 \end{array}\right].$$

最后一行对应的方程是 $0 = b_3 - b_2 - b_1$。左侧四个零并不表示整行可以直接删去；右端也必须是零，整行才是恒等式。

定理 8.2（增广消元的相容判据）　对增广矩阵 $[A\ b]$ 作初等行变换，使系数部分成为行阶梯形 U，右端相应成为 c。方程 $Ax = b$ 相容，当且仅当 U 的每个零行所对应的 c 分量也为零。

直观理解　零行表示一个已经被其他方程完全抵消的左端。相同的抵消必须发生在右端，否则留下的是 $0 = c_i \neq 0$。若没有这种矛盾，就可以从最后一个主元开始向上解出未知数；没有主元的坐标可以任选。

思路　先说明同步行变换保持解集，再分别证明两个方向。必要性来自零行方程；充分性来自回代，要说明每个主元非零，所以相应变量确实能解出。

证明　**行变换保持解集。**交换方程只改变顺序；把等式两侧同乘非零数可以除回去；给一个等式加上另一个等式的倍数，可以减回去。这三种操作都可逆，因此 $Ax = b$ 与 $Ux = c$ 的完整解集相同。

必要性。若 $Ux = c$ 有解，第 i 行系数全零时，对任何 x 都有左端为零，因此必须 $c_i = 0$。若 $c_i \neq 0$，这一个方程已经无解。

充分性。设所有零行对应的右端均为零。任意指定自由变量，例如全置零。从最后一个非零行开始，该行主元右侧的变量或者自由，或者已经在较低的主元行中解出。用右端减去这些已知项，再除以非零主元，唯一解出当前主元变量。逐行向上便构造出一个解。零行只要求 $0 = 0$，所以不会破坏构造。最后由行变换的等价性，这也是原方程的解。 ■

本定理要求检查的是**系数部分的零行**。若直接把右端也当成一个待选主元列，可能在最后一列找到一个主元；对原来的未知数而言，这正是 $0 = c_i \neq 0$ 的无解信号。计算中先标出增广竖线，可以避免把它误当成新的未知数。

例 8.1　必要条件什么时候也是充分条件？　对主例 A，求全部相容右端，并对一般相容右端给出一个解。

解　直接从阶梯形回代，既能证明充分性，又能构造解。零行要求 $b_3 = b_1 + b_2$。令 $x_2 = x_4 = 0$，第二行给出 $2x_3 = b_2 - 2b_1$，因此

$$x_3 = \frac{b_2}{2} - b_1, \qquad x_1 = b_1 - 2x_3 = 3b_1 - b_2.$$

把 $x = (3b_1 - b_2, 0, b_2/2 - b_1, 0)$ 代回，第一行得到 b_1，第二行得到 b_2，第三行作为前两行之和得到 $b_1 + b_2 = b_3$。所以所有且仅有 $b = (b_1, b_2, b_1 + b_2)$ 相容。原先的必要条件由此成为充要条件。 ♦

第 6 讲的列空间给了同一个判据的另一种写法：$Ax = b$ 有解，当且仅当 $b \in C(A)$。因为 Ax 就是 A 各列以 x_j 为系数的线性组合，这句话直接来自列空间定义。本例因此还算出了

$$C(A) = \{\, (u, v, u + v) \mid u, v \in \mathbb{R} \,\}.$$

相容条件和列空间没有给出两套互不相干的答案：前者把右端必须满足的方程写出，后者把全部可达右端当成一个集合。一个从消元读取，一个从列组合读取。

想一想　如果 A 原本没有零行，能否省略相容性检查？

答　不能。主例的三行原本全非零，消元后仍产生一个零行，限制 $b_3 = b_1 + b_2$。要检查消元后系数部分的零行。反过来，若阶梯形每行都有主元，就不会出现零行矛盾，任何右端都相容。这个结论将在满行秩情形中再次使用。

8.2　一个特解怎样给出全部解？

主例取 $b = (1, 5, 6)$，相容条件成立。把自由变量置零，上一节的公式给出 $x_p = (-2, 0, 3/2, 0)$。但这只是一个点：第 7 讲已经找到非零向量 $s_1 = (-2, 1, 0, 0)$ 和 $s_2 = (2, 0, -2, 1)$ 满足 $As_1 = As_2 = 0$，所以给 x_p 加上它们，不会改变乘积 Ax_p。要说明没有遗漏，还要反过来分析任意两个解的差。

定义 8.3（特解）　对相容方程 $Ax = b$，任意一个确定的解称为它的**特解**（particular solution），常记作 x_p。记 $S_b = \{\, x \in \mathbb{R}^n \mid Ax = b \,\}$ 为完整解集。

主例中的 $(-2, 0, 3/2, 0)$ 是特解；$(-4, 1, 3/2, 0)$ 也是特解，因为它是前者加 s_1。零向量不属于这两个解之间的选择：$A0 = 0 \neq (1, 5, 6)$。若 $b = 0$，零向量反而总是一个特解，此时 $S_0 = N(A)$。本讲的「特解」与第 7 讲按自由变量逐一取单位坐标得到的「特殊解」用途不同，后者专指描述零空间的一组方向。

定理 8.4（非齐次方程的全部解）　设 $A \in \mathbb{R}^{m \times n}$，且 $Ax = b$ 有一个特解 x_p。则

$$S_b = x_p + N(A) = \{\, x_p + z \mid z \in N(A) \,\}.$$

直观理解　从特解出发，沿任何被 A 消掉的方向移动，输出仍是 b。反过来，若两个输入产生同一个输出，输入之差就被 A 消掉。因此整个解集具有与零空间相同的方向，只是整体移动到经过 x_p 的位置。

思路　集合相等必须证明两个包含关系。先把 x_p+z 乘回去，确认写出的都是解；再对任意解 x 作差 $x-x_p$，确认所有解都在所写集合内。

证明　**构造的每个向量都是解。**若 $z\in N(A)$，则 $Az=0$，所以 $A(x_p+z)=Ax_p+Az=b+0=b$。因此 $x_p+N(A)\subseteq S_b$。

任意解都能这样写。若 $x\in S_b$，则 $Ax=b$，并且 $Ax_p=b$。两式相减得 $A(x-x_p)=b-b=0$，故 $z=x-x_p\in N(A)$。于是 $x=x_p+z$，从而 $S_b\subseteq x_p+N(A)$。两个方向合起来给出结论。 ■

「有一个特解」这个条件必须先成立。主例取 $b=(1,5,7)$ 时，零行已经给出矛盾；这时不存在可以平移的起点，解集是空集，不能任意挑一个向量放进公式充当 x_p。另一方面，定理没有要求特解唯一，也没有要求 A 是方阵。

■ **例 8.2　把上一讲的齐次解全部加回来**　写出主例在 $b=(1,5,6)$ 时的全部解，并检查参数 $s=3,t=-1$ 对应的点。

解　先取已经验证的特解，再加上第 7 讲的完整零空间。由定理 8.4，

$$x=\begin{bmatrix}-2\\0\\3/2\\0\end{bmatrix}+s\begin{bmatrix}-2\\1\\0\\0\end{bmatrix}+t\begin{bmatrix}2\\0\\-2\\1\end{bmatrix},\qquad s,t\in\mathbb{R}.$$

逐项写开是 $x=(-2-2s+2t,s,3/2-2t,t)$。自由坐标 $x_2=s,x_4=t$ 可以直接恢复参数，所以不同参数不会重复表示同一个向量。取 $s=3,t=-1$ 得 $(-10,3,7/2,-1)$。代回原三行分别为 $-10+6+7-2=1$、$-20+12+21-8=5$、$-30+18+28-10=6$。整个参数式的正确性由 $Ax_p=b$ 与 $As_1=As_2=0$ 保证，而不只靠这一个点的验算。 ♦

主例的解集在 $\mathbb{R}^4$ 中，是有两个独立参数的平移平面；画在纸上时不能把四个坐标假装成普通三维坐标。图 8.1 改用真正位于 $\mathbb{R}^2$ 的方程 $x_1+2x_2=2$ 展示同一个平移机制。蓝色齐次直线经过原点，橙色解直线经过特解 $(2,0)$。

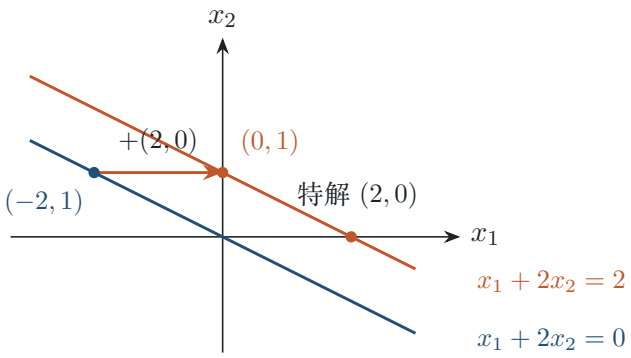


图 8.1　把齐次直线平移向量 $(2,0)$，恰好得到非齐次方程的整条解直线。

图中的水平箭头把齐次解 $(-2,1)$ 送到非齐次解 $(0,1)$。同一支平移向量也把原点送到特解 $(2,0)$。两条直线的方向相同，却只有蓝色直线包含原点；这个区别解释了为什么非零右端的解集通常不能叫子空间。

易错提醒　「全部解是特解与齐次解的任意线性组合」会把特解的系数也放开，因而产生错误输出。若 $Ax_p=b\neq 0$，则 $A(2x_p)=2b\neq b$。正确形式是 x_p+z，其中 z 遍历 $N(A)$；特解前的系数固定为 1。

想一想　换一个特解，得到的完整解集会不会变？

答　不会。若 y_p 也是特解，$A(y_p-x_p)=0$，所以 $y_p=x_p+z_0$，其中 $z_0\in N(A)$。对于每个 $z\in N(A)$，z_0+z 仍在零空间；反过来，每个 $w\in N(A)$ 都可写成 $z_0+(w-z_0)$。因此 $y_p+N(A)=x_p+N(A)$。改变特解只改变参数式的起点。

8.3　右端必须跟着系数一起变换

考虑原课习题 8.2 中的方程 $Ux=c$：

$$U=\begin{bmatrix}1&2&3\\0&0&4\end{bmatrix},\qquad c=\begin{bmatrix}5\\8\end{bmatrix}.$$

第二行给出 $4x_3=8$，所以 $x_3=2$。若只把系数中的 4 除成 1，却把右端 8 留下，就会错误地得到 $x_3=8$。增广矩阵把同一个等式的左右两侧放在一行，行操作必须作用于整行。

例 8.3　同步作 Gauss–Jordan 消元　分别把 $[U\ 0]$、$[U\ c]$ 化成简化行阶梯形，并求齐次与非齐次方程的全部解。

解　先把第二行除以 4，再把第一行减去新第二行的 3 倍。对非齐次增广矩阵，计算为

$$\left[\begin{array}{ccc|c}1&2&3&5\\0&0&4&8\end{array}\right]\longrightarrow\left[\begin{array}{ccc|c}1&2&3&5\\0&0&1&2\end{array}\right]\longrightarrow\left[\begin{array}{ccc|c}1&2&0&-1\\0&0&1&2\end{array}\right].$$

因此 $R=\left[\begin{smallmatrix}1&2&0\\0&0&1\end{smallmatrix}\right]$，变换后的右端为 $d=(-1,2)$，而非原来的 $(5,8)$。主元列为第 1、3 列，自由变量是 x_2。对齐次右端，同样两步始终把零右端变成零，故 $Rx=0$ 给出 $x_3=0$、$x_1+2x_2=0$。令

$x_2 = t$，得到

$$N(U) = \{\, t(-2,1,0) \mid t \in \mathbb{R} \,\}.$$

非齐次方程则给出 $x_3 = 2$、$x_1 + 2x_2 = -1$。令 $x_2 = t$，全部解为

$$x = (-1-2t, t, 2) = (-1,0,2) + t(-2,1,0), \qquad t \in \mathbb{R}.$$

代回原方程，第一行是 $(-1-2t) + 2t + 3\cdot 2 = 5$，第二行是 $4 \cdot 2 = 8$，所以任意 t 都可行；反过来，每个解的第二坐标就是 t，其余坐标由两条主元方程唯一确定，所以没有遗漏。原课解答取 $t = 1$ 得 $(-3,1,2)$，本讲取 $t = 0$ 得 $(-1,0,2)$；两者相差齐次方向 $(-2,1,0)$。 ♦

这个例子显示，R 与 U 一般不把同一个 x 映到同一个输出。它们有相同的齐次解集，因为零右端在行操作下仍为零；对于非齐次方程，应比较的是 $Ux = c$ 和 $Rx = d$。把系数化简后再机械地抄回原右端，会改变原问题。

把刚才的两步按右端单独跟踪，可以看清相容条件如何产生。第一步把第二行除以 4，所以 8 同时变成 2；第二步从第一行减去新第二行的 3 倍，所以第一分量从 5 变成 $5 - 3\cdot 2 = -1$。对一般右端 (c_1, c_2)，最终右端是 $(c_1 - 3c_2/4, c_2/4)$，并没有多出约束，因为系数矩阵没有零行。若系数矩阵出现零行，相同的线性组合也会作用在右端上，并给出必须为零的条件。

可以把每一步写成左乘一个可逆矩阵 E：U 变为 EU，右端 c 也必须变成 Ec。这不表示原方程的输出数值不变，而表示两边同时施加同一个可逆坐标变换后，方程真假和输入解集保持不变。消元记录表若只记系数部分，会丢失恢复原方程所需的信息；记录增广矩阵则把这两部分绑定在一起。

求非齐次全解的次序　先对增广矩阵同步消元，检查零行是否矛盾；相容时，把自由变量置零得到一个特解；再把全部自由变量放开，或者加上完整零空间。最后代回原矩阵与原右端检查。

为什么可以把自由变量全部置零？相容判据的充分性证明已经说明，任意指定自由变量，都能逐行解出主元变量。零只是一组方便的选择，并非要求所有解的自由坐标必须为零。在这个选择下，特解通常稀疏，便于检查；但「标准」来自计算约定，不表示它在长度、距离等意义下具有最优性。最近的解属于后面最小二乘与正交投影的内容，本讲不使用这些性质。

若把所有相容右端放在一起看，主例的输出不是整个 $\mathbb{R}^3$，而是平面 $b_3 = b_1 + b_2$。对应每个点 b，输入解集则是 $\mathbb{R}^4$ 中一个与二维零空间平行的仿射平面。一个右端条件和两个自由参数分别发生在不同空间：右端空间有一个线性约束，输入空间有两个可选坐标。只说「有平面」容易混淆这两个平面，判断所属空间时应先读向量长度：b 有 m 个分量，x 有 n 个分量。

若主元列已排到前面，简化行阶梯形的相容方程可写成

$$\begin{bmatrix} I_r & F \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{\mathrm{p}} \\ x_{\mathrm{f}} \end{bmatrix} = \begin{bmatrix} d_{\mathrm{p}} \\ 0 \end{bmatrix}.$$

这里 x_{p} 收集主元坐标，x_{f} 收集自由坐标，F 有 r 行、$n-r$ 列。逐行读出 $x_{\mathrm{p}} = d_{\mathrm{p}} - Fx_{\mathrm{f}}$，因此在这个重排后的坐标次序中，

$$x = \begin{bmatrix} d_{\mathrm{p}} \\ 0 \end{bmatrix} + \begin{bmatrix} -F \\ I_{n-r} \end{bmatrix} x_{\mathrm{f}}.$$

第一项是自由变量置零的特解，第二项是齐次解。如果实际主元列交错分布，必须先说明重排，再把结果还原到原变量次序。例如主例主元列为第 1、3 列，块形式的前两个坐标对应 x_1, x_3，不能直接当成原向量的前两个坐标。

■ **例 8.4　同一个解集的两个起点**　对 $Ux = c$，以 $(-3, 1, 2)$ 为特解写全解，并与以 $(-1, 0, 2)$ 为特解的写法比较。

解　先加上同一个零空间，再比较参数。第二种写法为 $x = (-3, 1, 2) + u(-2, 1, 0) = (-3 - 2u, 1 + u, 2)$。令 $t = 1 + u$，就得到 $(-1 - 2t, t, 2)$。由于 u 遍历实数时 $1 + u$ 也遍历实数，两种写法给出完全相同的集合。齐次方向没有改变，只是自由坐标的参数原点从 0 改成了 1。 ♦

8.4　有解与唯一：两件事分别由什么控制？

矩阵 $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ 有两个变量、两个主元，而矩阵 $C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ 有两行、每行一个主元。二者的秩都是 2，但对 $Bx = b$ 与 $Cx = b$ 的含义不同：前者若有解，两个变量都被确定；后者不管右端是什么，都有两个主元方程可以解出相应坐标。

> **定义 8.5（满列秩与满行秩）**　设 A 是 $m \times n$ 矩阵，秩为 r。若 $r = n$，称 A 为**满列秩**（full column rank）；若 $r = m$，称 A 为**满行秩**（full row rank）。

B 满列秩而不满行秩，因为 $r = n = 2 < m = 3$；C 满行秩而不满列秩，因为 $r = m = 2 < n = 3$。单位矩阵两者都满足。非零的 2×2 矩阵 $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ 只有一个主元，所以两者都不满足。定义中的「满」对应主元数达到某个方向的上限，与每个元素是否非零无关。

> **命题 8.6（分别判断存在性与唯一性）**　对实数矩阵 $A \in \mathbb{R}^{m \times n}$：若 $r = n$，则对每个 b，$Ax = b$ 至多有一个解；若 $r = m$，则对每个 b，$Ax = b$ 至少有一个解。若 $r = m = n$，则对每个 b 都有唯一解。

证明　满列秩时没有自由变量，齐次方程只有零解。若 x, y 都解 $Ax = b$，则 $A(x - y) = 0$，所以 $x - y = 0$，即 $x = y$。这证明有解时唯一，但没有保证任意右端都能出现。

满行秩时阶梯形没有零行。对任意 b 同步变换得到右端 c，由定理 8.2 都相容。自由变量置零后可以回代得到特解，所以每个右端至少有一个解。若两种满秩同时成立，上述存在性与唯一性合在一起，对每个右端都给出唯一解。 ■

■ **例 8.5　满列秩：有解时唯一**　求 $Bx = b$ 的相容条件，并求 $b = (2, 3, 5)$ 时的解。

解　先把三行直接写成方程：$x_1 = b_1$、$x_2 = b_2$、$x_1 + x_2 = b_3$。前两行已经确定全部变量，第三行因此要求 $b_3 = b_1 + b_2$。条件满足时唯一解为 (b_1, b_2)，不满足时无解。取 $(2, 3, 5)$ 得 $x = (2, 3)$，代回得到 $Bx = (2, 3, 2 + 3) = (2, 3, 5)$。这里的第三个方程负责检查右端，没有再增加一个可自由指定的变量。 ♦

例 8.6　满行秩：每个右端都能达到　求 $Cx=b$ 的全部解，并说明为何不需要相容条件。

解　直接从主元方程读出 $x_1+x_3=b_1$、$x_2+x_3=b_2$。令自由变量 $x_3=t$，得到 $x=(b_1-t,b_2-t,t)$，其中 $t\in\mathbb{R}$。代回两行分别为 $b_1-t+t=b_1$、$b_2-t+t=b_2$，所以对任意 b_1,b_2 都有解。每个解的第三坐标恢复参数 t，故这就是全部解。取 $b=(2,3)$ 时，$t=0$ 给 $(2,3,0)$，$t=1$ 给 $(1,2,1)$；两点之差为 $(-1,-1,1)$，恰好属于零空间。 ♦

满列秩条件不足以保证任意 b 可解：B 要求 $b_3=b_1+b_2$，所以 $(2,3,6)$ 不可达。满行秩条件不足以保证唯一：$C(2-t,3-t,t)^T=(2,3)^T$ 对所有实数 t 成立。两个反例对应两个不同问题，做题时应分别提问「有多少右端可以解」和「一个相容右端对应多少输入」。

图 8.2 用两个更小的矩阵区分输出限制与输入不唯一。左图画 $D=\begin{bmatrix}1\\2\end{bmatrix}$ 的输出空间：标量 t 只能产生直线上的 $(t,2t)$。右图画 $E=\begin{bmatrix}1&2\end{bmatrix}$ 对右端 2 的输入解集：整条直线都被送到同一个数 2。

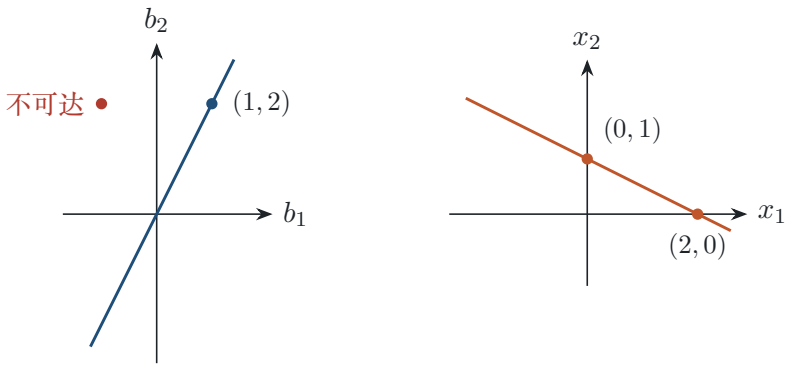


图 8.2　满列秩可以限制可达右端，满行秩可以让同一右端对应多个输入。

左图中的蓝线属于右端所在的 $\mathbb{R}^2$；红点 $(-1,2)$ 不在蓝线上，所以方程不相容。右图中的橙线属于未知向量所在的 $\mathbb{R}^2$，它给出固定右端的全部输入。虽然两幅图都画了直线，横纵轴代表的量不同，读图时必须先看坐标名。用同一张没有标轴的平面图同时代替输入解集和输出列空间，会掩盖这两个概念的区别。

8.5　四种秩情形各有多少解？

已知矩阵大小 $m\times n$ 和秩 r，只能推出 $r\le m$、$r\le n$；还必须知道它实际有多少主元。将主元列暂时排在前面，简化行阶梯形具有块结构 $\begin{bmatrix}I_r&F\\0&0\end{bmatrix}$。增广右端的零行约束决定有解还是无解，自由变量个数 $n-r$ 决定相容后有几个解。

秩关系	简化系数形	相容性	每个相容右端的解数
$r=m=n$	I	任意 b 相容	唯一解
$r=n<m$	$\begin{bmatrix}I\\0\end{bmatrix}$	需满足行约束	唯一解
$r=m<n$	$[I\ F]$	任意 b 相容	无穷多解
$r<m,\ r<n$	$\begin{bmatrix}I&F\\0&0\end{bmatrix}$	需满足行约束	无穷多解

表中「唯一」与「无穷」都以右端相容为前提。若有零行，违反任一右端约束就得到零个解；若没有自由变量，每个相容右端只有一个解；若至少有一个自由变量，任意改变它就得到另一个解，所以解无穷多。这里没有「恰好两个解」等中间情况，因为变量取实数，自由坐标可以取任意实数。

这四格并非四个互斥的口号，而是把两个独立问题交叉起来：零行数 $m-r$ 控制右端有没有约束，自由变量数 $n-r$ 控制相容后的解有多少参数。满方阵满秩时两种数量都为零；满列秩长方阵只有右端约束；满行秩宽矩阵只有自由变量；双缺秩时两者都出现。若右端恰好满足全部约束，解的参数个数仍只由 $n-r$ 决定，不会因某些右端坐标恰为零而减少。

图 8.3 用真正的二维输入、输出空间细看双缺秩情形。取 $A_0=\left[\begin{smallmatrix}1&2\\2&4\end{smallmatrix}\right]$，则 $A_0x=(x_1+2x_2)(1,2)$。输出只能沿 $(1,2)$ 方向变化；固定可达右端 $b=(3,6)$ 后，输入仍可沿零空间方向 $(-2,1)$ 移动。这两条直线分别位于不同空间，不能因都画成直线就把它们当成同一个集合。

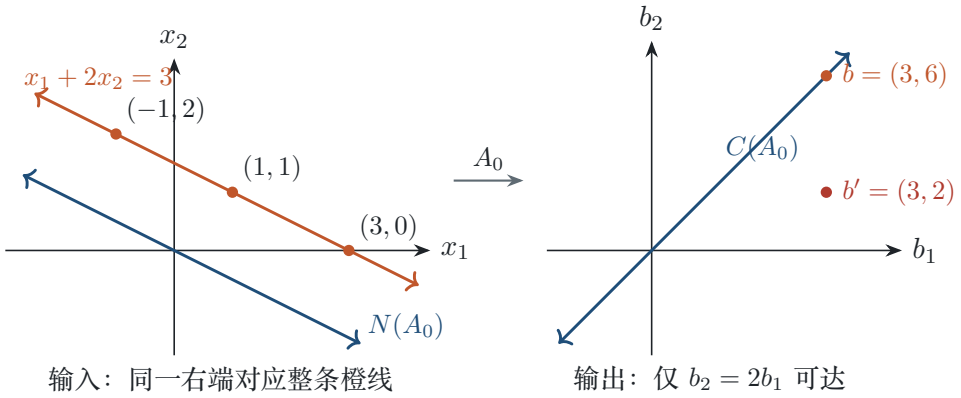


图 8.3　秩为 1 的 A_0 把左图整条橙线送到右图的同一个点 $(3,6)$，而红点 $(3,2)$ 没有原像。

左图的三个输入分别给出 $x_1+2x_2=3$，所以输出全是 $3(1,2)=(3,6)$。右图的红点违反 $b_2=2b_1$，因而没有任何输入能达到它。两条蓝线都可无限延伸，但所在空间、方向与意义不同：左图是被送到零的输入，右图是可以达到的输出。这里 $m-r=n-r=1$，前者给出一个右端约束，后者给出一个自由参数。图只画了显示窗口内的部分直线，箭头表示它们在两端都没有界限。

矩阵有 $m<n$ 仍不保证满行秩。例如 $G=\begin{bmatrix}1&2&3\\2&4&6\end{bmatrix}$ 的第二行是第一行的两倍，秩为 1；右端 $(1,0)$ 不相容。每个 b 是否可解取决于实际秩是否达到 m，不能只比较矩阵形状。

■ **例 8.7　行列都缺主元**　对下列矩阵 J，求 $Jx=b$ 的相容条件、全部解和秩情形：

$$J=\begin{bmatrix}1&0&1&2\\0&1&1&-1\\0&0&0&0\end{bmatrix}$$

解　主元列为第 $1,2$ 列，第三行要求 $b_3=0$。相容时令两个自由变量 $x_3=s,x_4=t$，前两行给出

$$x=(b_1-s-2t,\ b_2-s+t,\ s,\ t),\qquad s,t\in\mathbb{R}.$$

因此秩 $r=2$ 同时小于 $m=3$ 和 $n=4$；每个相容右端对应无穷多解，不相容右端没有解。取

$b=(1,2,0)$、$s=1,t=0$ 得 $(0,1,1,0)$，代回三行分别为 $1,2,0$。取 $b=(1,2,1)$ 时第三行变成 $0=1$，没有解。 ♦

想一想　看到 $m<n$，能否立即断言 $Ax=b$ 对每个 b 都可解？

答　不能。只有实际秩 $r=m$ 时，消元后才没有零行，因而每个右端都相容。上面的 G 虽有 $2<3$，秩却是 1；消元后留下一个零行，右端必须满足相应关系。列数给出的是秩的上界，不是秩一定达到的保证。

零矩阵是分类表的边界情形。若 $A=0_{m\times n}$，秩 $r=0$；当 $b=0$ 时全部 $x\in\mathbb{R}^n$ 都是解，当 $b\neq 0$ 时无解。若 $m,n>0$，它属于双缺秩情形；若 $m=0$ 或 $n=0$，应按空矩阵的维数单独解释，不能用「至少有一个自由变量」这句话代替维数检查。

8.6　特解、特殊解与同一解集

原习题 8.1 把「特解」与「齐次解」放在一起辨析。若 x_p 满足 $Ax_p=b$，而 x_n 满足 $Ax_n=0$，正确的组合是 x_p+x_n，因为 $A(x_p+x_n)=b$。把它们任意线性组合会把 x_p 的系数也改变，右端随之改变。

■ **例 8.8**　**原题 8.1：三个说法都错在哪里？**　设 $Ax=b$ 相容。分别解释以下说法为何错误：完整解是 x_p,x_n 的任意线性组合；特解至多一个；若 A 可逆则零空间没有解。

解　第一句错误，因为 $2x_p+0x_n$ 的像是 $2b$，通常不等于 b；正确形式是固定系数 1 的 x_p+x_n。第二句错误，因为每个 $z\in N(A)$ 都给出另一个特解 x_p+z；若零空间只有零向量，才只有一个特解。第三句把「非零零空间解」误当成「零空间中的解」：零向量总满足 $A0=0$，可逆矩阵的零空间恰为 $\{0\}$。三个纠正都由 A 的线性性直接验证。 ♦

特解确实不唯一时，改变特解只会改变参数原点。主例中的 $x_p=(-2,0,3/2,0)$ 和 $x_p+s_1=(-4,1,3/2,0)$ 都可以作为起点。改用第二个向量为起点时，原参数 (s,t) 对应新参数 $(s-1,t)$：只有 s_1 方向的参数平移，s_2 方向的参数不变，所覆盖的几何集合仍相同；这一点在图 8.1 中用一条直线展示。

8.7　同一个右端下，矩阵是否相同？

原题 8.3 的量词很强：对每个右端 b，两个方程的完整解集相同。固定任意向量 y，令 $b=Ay$，则 y 肯定是 $Ax=b$ 的解；由假设它也必须是 $Cx=b$ 的解。因此 $Cy=Ay$。由于 y 任意，两矩阵作用相同，特别地分别取标准基向量，就能逐列得到 $C=A$。

命题 8.7（完整解集确定线性映射）　设 $A,C\in\mathbb{R}^{m\times n}$。若对每个 $b\in\mathbb{R}^m$，方程 $Ax=b$ 与 $Cx=b$ 的解集相同，则 $A=C$。

证明　任取 $y\in\mathbb{R}^n$，令 $b=Ay$。因 $Ay=b$，向量 y 属于第一个方程的解集；按假设也属于第二

个解集，所以 $Cy = b = Ay$。这对每个 y 成立。取 $y = e_j$，便有 $Ce_j = Ae_j$，即 C 与 A 的第 j 列相同；对全部 j 都成立，故 $A = C$。 ■

■ **例 8.9　同一个零空间，不同的非齐次解**　求 $A = I_2, C = 2I_2$ 在 $b = 0, (2,4)$ 时的解。

解　分别解出两个方程，就能看清量词的区别。右端为零时，$Ax = 0$ 与 $Cx = 0$ 都只有 $x = (0,0)$。右端为 $(2,4)$ 时，$Ax = b$ 的唯一解为 $(2,4)$，而 $Cx = b$ 的唯一解为 $(1,2)$；代入分别得到 $I_2(2,4) = (2,4)$ 与 $2I_2(1,2) = (2,4)$。因此，同一个齐次解集只说明两矩阵把哪些输入送到零，不确定非零输入的像。比较全部右端才足以逐列确定矩阵。 ♦

想一想　为什么证明 $A = C$ 时只需检查 $e_1, \dots, e_n$？

答　矩阵的第 j 列正是 Ae_j，因为 e_j 只有第 j 个坐标等于 1，其余为 0。所以若两个矩阵对所有向量作用相同，它们在每个标准基向量上的结果相同，也就是每一列逐项相同。检查基向量不是额外假设，而是把「作用相同」翻译成矩阵元素相同。

本讲小结

(1) **相容条件**：消元后系数零行对应的右端必须为 0；否则 $Ax = b$ 不相容。

(2) **全部解**：若 x_p 是特解，则解集为 $x_p + N(A)$；任意两解之差属于 $N(A)$。

(3) **参数个数**：r 个主元留下 $n - r$ 个自由变量；每个相容右端的解集有 $n - r$ 个参数。

(4) **存在与唯一**：满列秩 $r = n$ 保证至多一个解；满行秩 $r = m$ 保证每个右端至少一个解。

(5) **四种秩情形**：$r = m = n$ 时每个右端恰有唯一解；$r < m$ 时要检查零行对应的相容条件。

(6) **量词不能省略**：对所有 b 的完整解集相同可推出 $A = C$；只比较 $b = 0$ 不能确定矩阵。

练习

8.1 ★　对 $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$，求 $Ax = b$ 的相容条件。给出一个相容右端、一个不相容右端，并求相容时的全部解。

8.2 ★　设 $U = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix}$、$c = (5,8)^T$。写出一个特解、零空间方向和全部解。

8.3 ★　满列秩的 3×2 矩阵对每个右端是否有解？若有解，最多有几个？说明依据。

8.4 ★★　对 $Jx = b$，其中 $J = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$，求相容条件与全部解。

8.5 ★★　取 $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$。求 $Ax = (u,v)^T$ 的全部解，并说明解集有几个自由参数。

8.6 ★★　设 $A = 0_{2\times 3}$。分别求 $Ax = 0$ 和 $Ax = (1,0)^T$ 的解集，并说明秩与相容性。

8.7 ★★ 证明：若 x_p, x_q 都满足 $Ax = b$，则 $x_q - x_p \in N(A)$；反之，对任意 $z \in N(A)$，$x_p + z$ 仍是解。

8.8 ★★★ 判断并修正说法：「若 $Ax = b$ 与 $Cx = b$ 在 $b = 0$ 时有相同解集，则 $A = C$。」给出反例。

8.9 ★★★ 假设 $A, C \in \mathbb{R}^{m\times n}$，且对每个 $b \in \mathbb{R}^m$ 两方程完整解集相同。证明 $A = C$，并指出证明中如何使用「每个 b」。

练习答案与提示

8.1 设 $b = (b_1, b_2)^T$。第二行是第一行的 2 倍，所以相容条件为 $b_2 = 2b_1$。取 $b = (1, 2)^T$ 相容，$b = (1, 0)^T$ 不相容。相容时方程化为 $x_1 + 2x_2 = b_1$，令 $x_2 = t$，全部解是 $(b_1 - 2t, t)$，$t \in \mathbb{R}$。代入可得第二分量 $2b_1$。

8.2 消元得 $x_3 = 2$、$x_1 + 2x_2 = -1$。取 $x_2 = 0$ 得特解 $(-1, 0, 2)$；齐次方向为 $(-2, 1, 0)$。故全部解 $(-1 - 2t, t, 2)$。代入 U 得 $(5, 8)^T$。

8.3 不保证每个右端有解；有解时至多一个。秩为 2 等于列数，故无自由变量，齐次零空间只有零向量；但秩 $2 < 3$，消元有零行，存在不相容右端。

8.4 第三行给出相容条件 $b_3 = 0$。令 $x_3 = s, x_4 = t$，解为 $(b_1 - s - 2t, b_2 - s + t, s, t)$。代入前两行分别为 b_1, b_2，第三行恒为 0；两个自由坐标可从解中直接读回。

8.5 第三列无主元，令 $x_3 = t$。前两行给出全部解 $(u - t, v - t, t)$，其中 $t \in \mathbb{R}$。对任意 u, v 都相容，且有一个自由参数。

8.6 秩为 0。当右端为零时，所有 $(x_1, x_2, x_3) \in \mathbb{R}^3$ 都是解；当右端为 $(1, 0)^T$ 时，第一行成为 $0 = 1$，无解。因 $r < m$ 且 $r < n$，它是双缺秩情形。

8.7 两解相减：$A(x_q - x_p) = Ax_q - Ax_p = b - b = 0$，故差属于 $N(A)$。反过来若 $Az = 0$，则 $A(x_p + z) = b + 0 = b$，所以仍是解。

8.8 该说法错误。取 $A = I_2, C = 2I_2$，两者的齐次解集都为 $\{0\}$，但矩阵不同。相同零空间只说明齐次解相同，不确定矩阵对非零向量的作用。

8.9 任取 $y \in \mathbb{R}^n$，用假设中的「每个 b」选 $b = Ay$。则 y 解 $Ax = b$，也解 $Cx = b$，所以 $Cy = Ay$。依次取 $y = e_j$，两矩阵第 j 列相同；对每个 j 成立，故 $A = C$。

第 9 讲

线性无关、基与维数

Independence, Basis, and Dimension

对应材料：MIT 18.06SC Session 1.9 课时摘要 Independence, basis, and dimension；习题 9.1、9.2 及官方解答；ZoomNotes Part 3 的向量空间、基与维数。

本讲目标

(1) 会把线性无关转化为齐次方程只有零解，并通过主元检验向量之间的依赖。

(2) 能分别检查「张成」和「无关」，说清一组向量为什么构成给定空间的基。

(3) 理解基坐标存在且唯一，会在一组非标准基下求坐标并用图解释。

(4) 会证明有限维空间中所有基的向量数相同，区分空间的维数、向量的分量数与矩阵的秩。

(5) 会选原矩阵的主元列作为列空间基，用特殊解作为零空间基，并算出 r 与 $n-r$。

第 8 讲已经能把方程的全部解写成一个特解加若干自由方向。可是方向向量能否随便挑？若某个方向可以由其他方向组合得到，把它也列进去会怎样？本讲要给出一种没有遗漏、没有冗余的描述：先确定哪些方向不可替代，再用这些方向记录每个向量的坐标。

本讲默认实向量空间。每组向量都保留给定顺序；把向量拼成矩阵时，它们依次是列，组合系数因此组成列向量。行文中的 $(1,2,3)$ 表示列向量的坐标，横排只是节省空间。

9.1 哪些方向可以被其他方向替代？

取 $u=(1,0)$、$v=(1,1)$ 与 $w=(2,1)$。因为 $w=u+v$，用三者来表示向量时会有重复。例如 $(3,1)=2u+v=u+w$。差别在系数，结果却相同；两种表示相减得到 $u+v-w=0$。一个不全为零的系数组合就暴露了冗余。

定义 9.1（线性无关与线性相关） 设 $v_1,\dots,v_k$ 在同一个实向量空间中。若 $c_1v_1+\cdots+c_kv_k=0$ 只能在 $c_1=\cdots=c_k=0$ 时成立，称这组向量**线性无关**（linearly independent）。若存在一组不全为零的系数使组合为零，称它们**线性相关**（linearly dependent）；这样的等式称为非平凡线性关系。

例如 $(1,0),(0,1)$ 无关，因为两个坐标分别迫使两个系数为零；$(1,2),(2,4)$ 相关，因为 $2(1,2)-(2,4)=0$。单独一个非零向量无关：$cv=0$ 时，选 v 中一个非零分量便得到 $c=0$。只含零向量的组相关，因为 $1\cdot 0=0$ 已是非平凡关系。任何包含零向量的非空向量组都相关，其他向量的系

数取零即可。

命题 9.2（相关就是可以删去一个向量）　非空有限向量组线性相关，当且仅当其中至少有一个向量可写成其余向量的线性组合。

证明　若 $\sum_i c_i v_i = 0$ 且某个 $c_j \neq 0$，就可移项并除以这个非零系数，得到 $v_j = -\sum_{i\neq j}(c_i/c_j)v_i$。反过来，若 $v_j = \sum_{i\neq j} a_i v_i$，移项得 $v_j - \sum_{i\neq j} a_i v_i = 0$，其中 v_j 的系数为 1，所以关系非平凡。两步都不需要假定向量自身非零；它们也覆盖包含零向量的情形。 ■

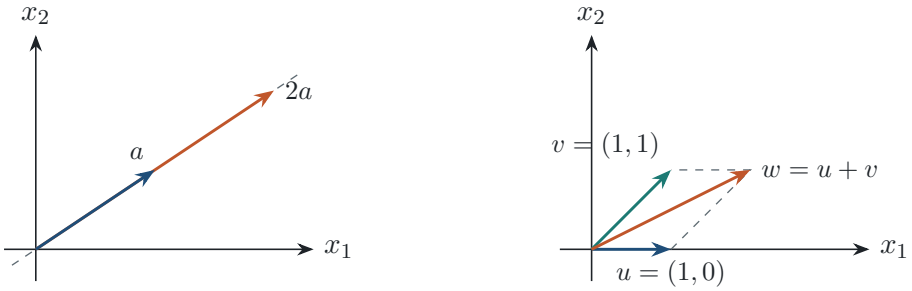
这里说「至少一个」，并不要求每个向量都能删。比如 $(1,0),(2,0),(0,1)$ 相关，前两个中任一个可由另一个代替，但删去 $(0,1)$ 后就只能得到横轴上的向量。判断能否删某个指定向量，要单独检查它是否在其余向量的张成空间里。

例 9.1　把关系写成齐次方程　对 $u=(1,0)$、$v=(1,1)$、$w=(2,1)$，求 $c_1u+c_2v+c_3w=0$ 的所有系数，并判断前两个向量是否无关。

解　先按坐标列方程：$c_1+c_2+2c_3=0$、$c_2+c_3=0$。令 $c_3=t$，得到 $c_2=-t$、$c_1=-t$，所以全部系数为 $t(-1,-1,1)$。取 $t=1$ 就是 $-u-v+w=0$，代回为 $(-1,0)+(-1,-1)+(2,1)=(0,0)$，因此三者相关。

只取前两者时，第二个坐标迫使 $c_2=0$，第一个坐标再迫使 $c_1=0$，因此它们无关。加上 w 增加了一个生成元，却没有增加新的方向：它已经是 $u+v$。这个例子也说明，只找到一个非平凡关系就足以证明相关；若要证明无关，则必须排除全部非平凡关系。 ♦

图 9.1 对照两种二维情形。左边的两个箭头在一条过原点直线上，右边的两个箭头不共线，虚线画出它们的和。箭头末端的位置和方向一起决定向量，不能只看两个箭头的长度是否不同。



两个非零向量共线：相关　　u,v 无关；u,v,w 相关

图 9.1　两个非零向量共线时一个可替代另一个；两个不共线方向已足以生成平面，加上它们的和会带来冗余。

右图的平行四边形说明「两个方向够用」和「三个向量相关」可以同时成立。线性无关描述一组向量内部是否冗余，不是说它们在图上相互垂直。右图 u 与 v 的夹角为 $45°$，仍然无关。

命题 9.3（用主元判断无关）　将 n 个向量排成 $m \times n$ 矩阵 A。这些列线性无关，当且仅当 $N(A) = \{0\}$，也当且仅当每列都是主元列，即 $\operatorname{rank} A = n$。若 $n > m$，这些列必线性相关。

证明　关系 $\sum_j c_j a_j = 0$ 就是 $Ac = 0$，故无关等价于只有零解。第 7 讲的消元给出 $n - r$ 个自由变量：没有自由变量时只能取零解，有自由变量时令其中一个等于 1、其余自由变量等于 0，便得到非零特殊解。所以无关等价于 $r = n$。最后，$r \le m < n$ 时必有自由变量，故相关。 ■

想一想　「任意两个都不共线」能否保证三个向量无关？

答　不能。例 9.1 的 u, v, w 任意两者都不共线，但 $u + v - w = 0$。两两无关只检查两项关系，三项关系还可能存在。三个实向量的完整判断仍要检查三系数方程，或把它们拼成矩阵找主元。

9.2　张成回答的是哪些向量能够得到

仍取 $u = (1, 0)$ 与 $v = (1, 1)$。给定任意目标 (p, q)，解 $cu + dv = (p, q)$ 得到 $d = q$、$c = p - q$。这说明平面里没有它们组合不到的点。若只留 u，则所有组合的第二坐标都为零，目标 $(0, 1)$ 就无法达到。这里比较的是整个可达集合，尚未问表示是否唯一。

定义 9.4（张成空间与生成组）　向量 $v_1, \dots, v_k$ 的所有线性组合构成的集合记为

$$\operatorname{span}\{v_1, \dots, v_k\} = \{\, c_1 v_1 + \cdots + c_k v_k \mid c_i \in \mathbb{R} \,\}.$$

若这个集合等于子空间 S，称这些向量**张成**（span）S，也称它们为 S 的一组生成元。

零系数给出零向量，两个组合相加仍是组合，一个组合乘常数仍是组合，所以张成集合确实是子空间。只含零向量的组张成 $\{0\}$，不会张成一条直线；一个非零向量的张成集合是整条过原点直线，允许正、负和任意大小的实系数。按空和为零的约定，空向量组也张成 $\{0\}$。

命题 9.5（最小包含空间）　$\operatorname{span}\{v_1, \dots, v_k\}$ 是包含这些向量的最小子空间：任何包含每个 v_i 的子空间 W 都包含其张成空间。

证明　张成集合已经是子空间，且把第 i 个系数取 1、其余取 0 就得到 v_i。另一方面，若 W 包含每个 v_i，由数乘封闭性它包含每个 $c_i v_i$，由加法封闭性它也包含这些向量的和。故所有线性组合都在 W 中，即张成空间包含于 W。 ■

「最小」按集合包含关系理解。它没有说选到的生成元最少；给一个生成组加上重复向量，张成集合仍旧不变。用矩阵记号，第 6 讲的列空间恰好是 $C(A) = \operatorname{span}\{a_1, \dots, a_n\}$，所以「$Ax = b$ 是否有解」就是「b 是否在这些列的张成空间中」。

■ **例 9.2　三个三维向量能否张成整个空间？**　原摘要中的三个向量是 $a = (1, 1, 2)$、$b = (2, 2, 5)$、$c = (3, 3, 8)$。判断它们是否张成 $\mathbb{R}^3$；若不能，写出恰好张成的集合，并判断前两个是否无关。

解　先找每个组合必须满足的坐标约束。三者第一、第二坐标相同，所以它们的组合都满足 $x_1 = x_2$，不可能得到 $(1,0,0)$，因而不能张成 $\mathbb{R}^3$。又 $c = 2b - a$，所以第三个向量可以删除而不改变张成集合。

现在证明前两个恰好张成平面 $P = \{(p,p,q) : p,q \in \mathbb{R}\}$，不能只说它们都在平面内。方程 $\alpha a + \beta b = (p,p,q)$ 化为 $\alpha + 2\beta = p$、$2\alpha + 5\beta = q$。第二式减第一式两倍得到 $\beta = q - 2p$，再得 $\alpha = 5p - 2q$。直接代回：

$$(5p - 2q)(1,1,2) + (q - 2p)(2,2,5) = (p,p,q).$$

所以平面中的每个向量都能得到。令 $p = q = 0$，上面的系数只能为零，说明 a, b 无关；也可在齐次方程里直接消元验证。于是两个无关向量已经给出了整个平面，第三个向量仅增加重复表示。 ♦

这个平面在 $\mathbb{R}^3$ 中，向量各有三个分量，但可自由指定的坐标只有 p, q 两个。第三个坐标没有给空间增加第三个独立方向，因为前两个坐标之间有约束。这会在后面的维数定义里变成精确的数字，而不只是一句几何直觉。

向量个数不等于独立方向数　「三个三维向量一定张成 $\mathbb{R}^3$」是错的：例 9.2 中三个向量都被限制在 $x_1 = x_2$。应当检查主元数或实际可达集合。向量很多只能提供更多候选方向，不能保证它们没有重复。

■ **例 9.3　只添一个向量，空间何时变大？**　从 $\operatorname{span}\{e_1, e_2\} = \{(p,q,0)\}$ 出发，分别加入 $e_1 + e_2$ 与 e_3，所得张成空间有什么变化？

解　检查新向量能否由旧向量表示。$e_1 + e_2 = (1,1,0)$ 已在旧平面内，所以每个 $ae_1 + be_2 + c(e_1 + e_2) = (a + c, b + c, 0)$ 仍在原平面，且旧组合仍被保留，张成空间不变。$e_3 = (0,0,1)$ 不在该平面，加入它后任意 $(p,q,r) = pe_1 + qe_2 + re_3$ 都能得到，张成空间变为 $\mathbb{R}^3$。一个新向量是否有用，取决于它是否在旧张成空间中。 ♦

9.3　基让每个向量有且只有一套坐标

前两节的问题可以放在一起：既要组合得到所有目标，又希望每个目标的系数只有一套。前一个要求排除遗漏，后一个要求排除重复。标准向量 e_1, e_2, e_3 满足这两点，因为 $(x_1, x_2, x_3) = x_1e_1 + x_2e_2 + x_3e_3$，而比较三个坐标就能确定三个系数。

定义 9.6（基与坐标）　子空间 S 的一组**基**（basis）是既线性无关又张成 S 的有序向量组 $\mathcal{B} = (v_1, \dots, v_d)$。若 $x = \sum_i c_i v_i$，系数列 $(c_1, \dots, c_d)^T$ 称为 x 关于 $\mathcal{B}$ 的**坐标**（coordinates），记为 $[x]_{\mathcal{B}}$。

按空和约定，空向量组线性无关且张成零子空间，因此它是 $\{0\}$ 的基，$\dim\{0\} = 0$。

例如 e_1, e_2, e_3 是 $\mathbb{R}^3$ 的标准基。例 9.2 的 a, b 是平面 P 的基，却不是 $\mathbb{R}^3$ 的基：它们无关，但是不能张成整个 $\mathbb{R}^3$。同一组向量是否构成基，必须说明「哪个空间的基」。u, v, w 可以张成 $\mathbb{R}^2$，

但相关，因此也不能称为它的基。

命题 9.7（方阵的列构成基当且仅当矩阵可逆）　设 $V=[v_1\ \cdots\ v_n]\in\mathbb{R}^{n\times n}$。则 $(v_1,\dots,v_n)$ 是 $\mathbb{R}^n$ 的一组基，当且仅当 V 可逆。

证明　若这些列构成基，那么每个 $x\in\mathbb{R}^n$ 都能唯一写成 $x=Vc$，所以线性映射 $c\mapsto Vc$ 是双射，V 可逆。反过来，若 V 可逆，则对每个 x，$c=V^{-1}x$ 是唯一满足 $Vc=x$ 的系数，故这些列张成 $\mathbb{R}^n$ 且线性无关。 ■

定理 9.8（基坐标存在且唯一）　有限向量组 $\mathcal{B}=(v_1,\dots,v_d)$ 是 S 的基，当且仅当每个 $x\in S$ 都能唯一地写成 $\sum_i c_iv_i$。

思路　先把两项要求分开：张成保证表示存在，无关保证两种表示的差只能是零系数组合。反过来，唯一表示零向量的要求便能检验无关，不必逐个比较所有非零向量。

证明　若 $\mathcal{B}$ 是基，张成性保证表示存在。设 $x=\sum_i c_iv_i=\sum_i d_iv_i$，相减得到 $\sum_i(c_i-d_i)v_i=0$。由无关性，所有 $c_i-d_i=0$，所以系数逐项相等。

反过来，若每个 $x\in S$ 都有唯一表示，存在性说明向量组张成 S。零向量已经有全零系数表示，若另有 $\sum_i c_iv_i=0$，唯一性就强迫 $c_i=0$，所以向量组无关，因而是基。 ■

■ **例 9.4　在斜的基下读坐标**　取 $\mathcal{B}=(u,v)$，其中 $u=(1,0)$、$v=(1,1)$。求 $x=(3,1)$ 的基坐标；若把基的顺序换成 (v,u)，结果如何？

解　解组合方程 $cu+dv=x$，而不是直接把 x 的标准分量抄作答案。第二坐标给 $d=1$，第一坐标给 $c+d=3$，故 $c=2$。所以 $[x]_{\mathcal{B}}=(2,1)^T$，代回 $2(1,0)+(1,1)=(3,1)$。基倒序后，$x=1v+2u$，坐标变为 $(1,2)^T$。向量 x 没有移动，只是记录其组合系数的顺序改变。 ♦

图 9.2 用两套网格解释这一计算。灰格对应标准基，橙色斜格对应 u,v。从原点先走两个 u，再走一个 v，就到达同一点 $(3,1)$；斜格的每个基本平行四边形对应两个系数各增加一个单位。

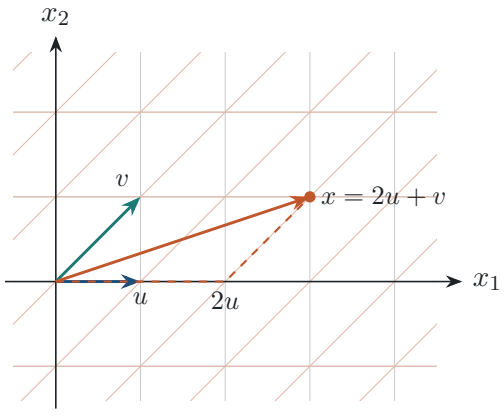


图 9.2　向量 $x=(3,1)$ 在基 $u=(1,0),v=(1,1)$ 下的坐标为 $(2,1)$；两套坐标描述同一个端点。

想一想　零向量的坐标是什么？如果某个非零向量也有全零坐标，会发生什么？

答　零向量的每个坐标都为零，因为全零组合给出零向量；基坐标唯一，所以没有另一种坐标表示。非零向量不能有全零坐标，否则组合结果就是零，和该向量非零矛盾。零向量本身若放进候选基，会让无关性立即失败。

注 9.1　坐标依赖基的顺序。把基向量重排，坐标按同一置换重排；换成不同基则坐标一般由一个可逆矩阵变换。第 29 讲会系统推导换基矩阵。本讲先使用最直接的办法：把基向量作列，解 $Vc = x$。

9.4　维数是每一组基都有的向量数

基同时无关、张成。但如果同一个空间有两组不同的基，为什么它们必须包含相同个数的向量？这一事实不依赖图形，也不依赖选用哪一种基。其核心是：若 k 个向量能张成空间，那么其中任何无关组至多有 k 个向量。

定理 9.9（无关组的大小不超过生成组）　若 $u_1, \dots, u_\ell$ 线性无关，而 $v_1, \dots, v_k$ 张成同一个有限维空间 S，则 $\ell \le k$。

思路　把每个 u_i 用生成组表示，系数排成 $k \times \ell$ 矩阵 C。如果 $\ell > k$，方程 $Cc = 0$ 有非零解；这组系数会把某些 u_i 组合成零，与无关性冲突。这里用到的正是「未知数比方程多必有自由变量」。

证明　因为 $v_1, \dots, v_k$ 张成 S，对每个 j 存在系数 c_{ij} 使 $u_j = \sum_{i=1}^{k} c_{ij} v_i$。令 $C = (c_{ij})$，则 $[u_1 \ \cdots \ u_\ell] = [v_1 \ \cdots \ v_k] C$。若 $\ell > k$，齐次系统 $Ct = 0$ 有非零解 $t \in \mathbb{R}^\ell$。于是

$$\sum_{j=1}^{\ell} t_j u_j = [v_1 \ \cdots \ v_k] C t = 0.$$

这是一条非平凡关系，因为 $t \neq 0$，与 u_j 无关矛盾。因此 $\ell \le k$。 ■

定理 9.10（同一空间的所有基等长）　若 $\mathcal{B}$ 与 $\mathcal{C}$ 都是有限维空间 S 的基，则它们含有相同个数的向量。这个共同的数称为 S 的**维数**（dimension），记作 $\dim S$。

证明　设 $\mathcal{B}$ 有 k 个向量，$\mathcal{C}$ 有 ℓ 个。$\mathcal{B}$ 张成 S，而 $\mathcal{C}$ 无关，由定理 9.9 得 $\ell \le k$。交换两组基的角色，同理得到 $k \le \ell$，因此 $k = \ell$。 ■

这个证明同时给出可用的比较原则：无关组不能比生成组更长；若两组各自既无关又张成，两个不等式逼出相同数量。平面的一组基无论选 (e_1, e_2) 还是 (u, v) 都含两个向量，所以平面维数为 2。这数字不等于平面向量在环境空间里的分量数；$\mathbb{R}^3$ 中的平面维数是 2，每个向量仍有三个分量。

■ **例 9.5　删减与扩充**　设 $S = \operatorname{span}\{(1,0,0), (0,1,0)\} \subset \mathbb{R}^3$。从这组生成元中删去一个会发生什么？加入 $(1,1,0)$ 或 $(0,0,1)$ 又分别发生什么？

解　原来的两个向量无关并张成 xy 平面，所以它们已经是 S 的一组基。删去任意一个只剩一条直线，不能张成 S。向量 $(1,1,0)$ 是原来两个向量的和，加入它只增加重复表示；向量 $(0,0,1)$ 不在 S 中，加入后得到三个无关向量，张成整个 $\mathbb{R}^3$。因此删减要检查是否仍张成，扩充要检查新向量是否落在原空间中。 ♦

维数不是坐标个数　「$\mathbb{R}^3$ 中的子空间都三维」是错的。$\operatorname{span}\{(1,0,0)\}$ 的维数为 1，平面 $x_1=x_2$ 的维数为 2；它们的向量虽然写成三个分量，独立方向却分别只有一个和两个。判断维数要数一组基的向量，而不是数环境空间的坐标分量。

9.5　列空间的基：主元列要回到原矩阵

上一讲的消元告诉我们如何解方程，本节把同一结果翻译成列空间的基。关键细节是：主元位置在行最简形中确定，但基向量必须从原矩阵中取出。

定理 9.11（原矩阵主元列构成列空间的基）　设 $A\in\mathbb{R}^{m\times n}$，对 A 作可逆的初等行变换得到 R。若 R 的主元列编号为 $j_1,\dots,j_r$，则原矩阵的列

$$a_{j_1},\dots,a_{j_r}$$

构成 $C(A)$ 的一组基；特别地，$\dim C(A)=r=\operatorname{rank}A$。

思路　行变换左乘一个可逆矩阵 E，所以 $R=EA$。这会改变列向量的坐标，却不会改变列之间的线性关系：$Ac=0$ 当且仅当 $Rc=0$。先在 R 中看出主元列无关且生成所有列，再把同一组系数关系拉回 A。

证明　设 R 的主元列为 $r_{j_1},\dots,r_{j_r}$。主元列无关：若它们有非平凡关系，把系数放在对应的 n 个位置、其余位置置零，就得到 $Rc=0$；查看对应的主元行，主元变量系数只能全为零，矛盾。每个非主元列也都在主元列的张成中：令对应自由变量为 1，其余自由变量为 0，解 $Rx=0$ 后把主元变量项移到等式另一侧，就得到该列是主元列的线性组合。

由于 $R=EA$ 且 E 可逆，$Ac=0\iff Rc=0$。因此原矩阵中对应的主元列也无关；对于任意列关系 $a_j=\sum_i\alpha_i a_{j_i}$，左乘 E 得到 $r_j=\sum_i\alpha_i r_{j_i}$，故原矩阵的每列也在这些主元列的张成中。它们既无关又张成 $C(A)$，所以是基。 ■

■ **例 9.6　主例的列空间基**　对

$$A=\begin{bmatrix}1&2&3&1\\1&1&2&1\\1&2&3&1\end{bmatrix}\quad\longrightarrow\quad R=\begin{bmatrix}1&0&1&1\\0&1&1&0\\0&0&0&0\end{bmatrix},$$

求 $C(A)$ 的一组基及其维数。

解　R 的主元列是第 $1,2$ 列，所以要回到 A 取

$$a_1=(1,1,1)^T,\qquad a_2=(2,1,2)^T.$$

它们无关，且 $a_3 = a_1 + a_2$、$a_4 = a_1$，所以四列的张成恰好等于 $\operatorname{span}\{a_1, a_2\}$。因此 $\{a_1, a_2\}$ 是 $C(A)$ 的基，$\dim C(A) = \operatorname{rank} A = 2$。注意 R 的前两列不能直接当作 $C(A)$ 的向量；它们属于行变换后的坐标表示。 ♦

图 9.3 把「主元位置」和「原列向量」分成两栏。左栏只负责标出主元列编号，右栏才列出实际作为基向量的 a_1, a_2；非主元列由它们组合得到。

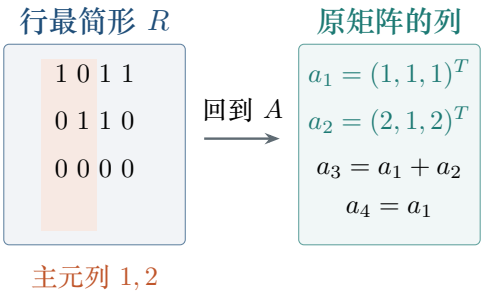


图 9.3　主元编号由行最简形决定，列空间基向量必须从原矩阵取出。

想一想　为什么行变换可以帮我们找主元，却不能把行最简形的列直接当作原列空间的基？

答　行变换左乘可逆矩阵，保留的是列之间的线性关系和秩，不保留每个列向量在环境空间中的具体位置。比如主例中原列 $a_1 = (1,1,1)^T$，而 R 的第一列是 $(1,0,0)^T$；它们都能作为「第 1 列是主元」的证据，却不是同一个向量。找完编号后必须回到 A。

9.6　零空间的基：特殊解记录自由方向

列空间描述输出能到哪里，零空间描述哪些输入会被送到零。对 $A \in \mathbb{R}^{m\times n}$，若秩为 r，就有 $n - r$ 个自由变量；每个自由变量取一次 1、其余自由变量取 0，得到一个特殊解。

定理 9.12（特殊解构成零空间的基）　设 R 是 A 的行最简形，主元列数为 r，自由变量为 $x_{j_1}, \dots, x_{j_{n-r}}$。对每个自由变量依次取 1，其余自由变量取 0，回代得到特殊解 $s_1, \dots, s_{n-r}$。则

$$N(A) = \operatorname{span}\{s_1, \dots, s_{n-r}\}, \qquad \dim N(A) = n - r.$$

证明　行变换不改变齐次方程的解集，所以只需解 $Rx = 0$。每个主元变量都是自由变量的线性组合，因而任意解都等于各自由变量系数乘对应特殊解之和，得到张成性。若 $\sum_i \alpha_i s_i = 0$，观察第 j_i 个自由坐标：第 i 个特殊解在那里为 1，其他特殊解在那里为 0，所以 $\alpha_i = 0$。因此特殊解无关，构成零空间的基。基中向量数正好是自由变量数 $n - r$。 ■

■ **例 9.7　主例的零空间基**　对主例的 R，求 $N(A)$ 的一组基，并验证其中两个方向确实被 A 消掉。

解　由 $Rx = 0$ 得 $x_1 + x_3 + x_4 = 0$、$x_2 + x_3 = 0$。令 $(x_3, x_4) = (1,0)$ 和 $(0,1)$，得到

$$s_1 = (-1,-1,1,0)^T, \qquad s_2 = (-1,0,0,1)^T.$$

任意零空间向量都写成 $x_3s_1+x_4s_2$，所以这两个向量构成基。直接代入原矩阵：$As_1=0$、$As_2=0$；而它们在第 3、4 个坐标上分别是 $(1,0)$ 与 $(0,1)$，因此无关。这里 $n=4, r=2$，维数公式也给出 $\dim N(A)=2$。 ♦

■ **例 9.8**　**错误的坐标符号会破坏特殊解**　有人把主例的第二个特殊解写成 $(-1,0,0,-1)^T$。判断它是否属于 $N(A)$，并说明如何快速发现错误。

解　代入 A 得

$$A(-1,0,0,-1)^T=(-2,-2,-2)^T\neq 0,$$

所以它不是零空间向量。错误来自 $x_1+x_3+x_4=0$：当 $x_4=1$ 时应有 $x_1=-1$，而不是 $x_4=-1$。最快的检查是把候选向量乘回原矩阵；若结果不是零，就不能称为特殊解。 ♦

注 9.2　列空间基和零空间基来自同一次消元，但取法相反：列空间基取原矩阵的主元列，零空间基取行最简形的自由变量特殊解。前者生活在 $\mathbb{R}^m$，后者生活在 $\mathbb{R}^n$；当 $m\neq n$ 时，不能把两者放在同一个坐标空间里比较。

9.7　两道综合题：从关系到几何

前面把独立、张成、基、维数和主元方法分别建立起来，现在用两道原题把它们串起来。第一题只问一组向量中最多能保留多少个独立方向；第二题把平面的基、与坐标平面的交线和法向方向放在同一幅图中。

■ **例 9.9**　**原题 9.1：六个四维向量中的最大独立组**　设

$$v_1=(1,-1,0,0)^T,\quad v_2=(1,0,-1,0)^T,\quad v_3=(1,0,0,-1)^T,$$
$$v_4=(0,1,-1,0)^T,\quad v_5=(0,1,0,-1)^T,\quad v_6=(0,0,1,-1)^T.$$

求这六个向量中最多有多少个线性无关，并给出一组达到最大值的向量。

解　先找明显关系：$v_4=v_2-v_1$、$v_5=v_3-v_1$、$v_6=v_3-v_2$，所以后三个向量不会增加前面三者的张成空间。把 v_1,v_2,v_3 作列，得到

$$M=\begin{bmatrix}1&1&1\\-1&0&0\\0&-1&0\\0&0&-1\end{bmatrix}.$$

前三行的三个主元已经说明三列无关；更直接地，$c_1v_1+c_2v_2+c_3v_3=0$ 的第 2、3、4 坐标依次给出 $c_1=c_2=c_3=0$。因此 $\{v_1,v_2,v_3\}$ 是一组含三个向量的无关组，而全部六个向量都在它们张成的空间中，最大值就是 3。这也说明「向量个数为六」不等于「独立方向有六个」。 ♦

■ **例 9.10**　**原题 9.2：平面、交线与法向量**　在 $\mathbb{R}^3$ 中考虑平面

$$P=\{(x,y,z):x-2y+3z=0\}.$$

求 P 的一组基、P 与 xy 平面的交线的一组基，以及所有垂直于 P 的向量空间的一组基。

解　把平面方程写成齐次方程 $\begin{bmatrix}1 & -2 & 3\end{bmatrix}x = 0$。令 $y = s, z = t$，则 $x = 2s - 3t$，所以

$$(x, y, z) = s(2, 1, 0) + t(-3, 0, 1).$$

故 P 的一组基为

$$p_1 = (2, 1, 0)^T, \qquad p_2 = (-3, 0, 1)^T.$$

它们无关，因为第二坐标和第三坐标分别锁定两个系数。与 xy 平面的交集还要求 $z = 0$，于是只剩 $s(2, 1, 0)$，交线的一组基为 $\{p_1\}$。方程的系数向量

$$n = (1, -2, 3)^T$$

分别与 p_1, p_2 点乘为 0，所以它垂直于平面。若 $q = (a, b, c)^T$ 也垂直于 P，则 $q \cdot p_1 = 2a + b = 0$、$q \cdot p_2 = -3a + c = 0$，从而 $q = a(1, -2, 3) = an$。因此法向空间恰为 $\operatorname{span}\{n\}$，一组基就是 $\{n\}$。 ♦

图 9.4 用一块半透明平面、平面内的两条基向量和 xy 平面中的交线显示三种对象。图中的投影是三维坐标到纸面的表示，代数判断仍由方程和点积完成。

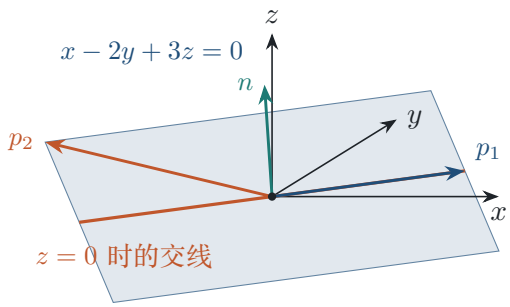


图 9.4　平面 $x - 2y + 3z = 0$ 的两条基向量在平面内，橙线是它与 xy 平面的交线，法向量 $n = (1, -2, 3)$ 垂直于二者。

平面基有两个向量，交线基有一个向量，法向空间也有一个向量；它们分别是维数 $2, 1, 1$ 的三个不同子空间。法向量不是平面中的第三个基向量，因为它不在平面内。

最后把秩与维数的关系再写成一张可计算的表。对 $m \times n$ 矩阵 A，列空间位于 $\mathbb{R}^m$，所以 $\dim C(A) = r \le m$；零空间位于 $\mathbb{R}^n$，所以 $\dim N(A) = n - r \le n$。列空间等于 $\mathbb{R}^m$ 当且仅当 $r = m$；零空间等于整个输入空间 $\mathbb{R}^n$ 当且仅当 $r = 0$。例如满列秩的 3×2 矩阵有 $r = 2$，零空间是 $\{0\}$，列空间却只是 $\mathbb{R}^3$ 中的二维子空间。两个空间的维数相加为 m 或 n 并不成立，正确的恒等式是 $\dim N(A) + \operatorname{rank} A = n$，因为它们都由未知数个数 n 计数。

想一想　为什么平面方程的系数向量天然给出法向量？

答　平面中的任意向量 p 都满足 $n^T p = 0$，其中 $n = (1, -2, 3)^T$ 是方程系数。若 $q = (a, b, c)^T$ 也垂直于平面，前面的基向量给出 $2a + b = 0$、$-3a + c = 0$，所以 $q = a(1, -2, 3)^T$。因此所有法向量都是 n 的倍数。

交集条件不能漏写　求平面与 xy 平面的交线时，不能只使用 $x-2y+3z=0$。还必须加入 $z=0$；否则得到的是整个二维平面 P，不是它在 xy 平面中的交线。

本讲小结

- 线性无关等价于齐次方程 $Ac=0$ 只有零解，也等价于矩阵每一列都是主元列。
- 张成描述所有线性组合；生成组不要求无关，加入重复向量不会改变张成空间。
- 基同时满足张成和无关，因此每个向量都有存在且唯一的基坐标。
- 同一有限维空间的所有基含有相同数量的向量，这个数量是空间的维数。
- 列空间基取原矩阵的主元列，维数等于秩；不能把行最简形的列直接当作原列空间基。
- 零空间基由自由变量的特殊解组成，维数为 $n-r$；它们生活在输入空间 $\mathbb{R}^n$。
- 平面、交线和法向空间的维数可以不同；方程系数向量给出法向方向，但不属于平面本身。

练习

9.1 ★　判断 $(1,2)$ 与 $(2,4)$ 是否线性无关；写出一条非平凡关系。

9.2 ★　判断 e_1, e_2, e_1+e_2 是否构成 $\mathbb{R}^2$ 的基，并说明理由。

9.3 ★　求 $A=\begin{bmatrix}1&2&3\\2&4&7\\1&2&4\end{bmatrix}$ 的列空间基和零空间基。

9.4 ★　在基 $\mathcal{B}=((1,1),(1,-1))$ 下求 $(4,2)$ 的坐标。

9.5 ★★　证明：若一组向量无关，则删去其中任意一个后仍无关；反过来由「删去任意一个后仍无关」能否推出原组无关？

9.6 ★★　求平面 $x+y+z=0$ 与 xz 平面的交线基。

9.7 ★★　说明向量 $(1,0,0),(0,1,0),(1,1,0)$ 张成什么空间，并给出一组基。

9.8 ★★　设 A 是 3×5 矩阵且秩为 2。求 $\dim C(A)$ 和 $\dim N(A)$，并说明五列是否可能线性无关。

9.9 ★★★　设 $u_1,\dots,u_\ell$ 无关而 $v_1,\dots,v_k$ 张成同一空间。证明 $\ell\le k$，并指出证明中使用了哪一个矩阵事实。

9.10 ★★★　求 $P=\{(x,y,z):2x-y+z=0\}$ 的一组基、与 yz 平面的交线基和法向空间基。

练习答案与提示

9.1　相关，因为 $2(1,2)-(2,4)=0$。

9.2　不构成基。它们张成 $\mathbb{R}^2$，但三列的系数 $(1,1,-1)$ 给出非平凡关系 $e_1+e_2-(e_1+e_2)=0$，故相关；删去第三个后 e_1,e_2 才是基。

9.3　行最简形主元列为第 1, 3 列，故列空间基取原列 $(1,2,1)^T,(3,7,4)^T$。齐次方程给 $x_1+2x_2+3x_3=0$、

$x_2 + x_3 = 0$，令 $x_3 = t$ 得 $(-2t, t, 0)$，所以零空间基为 $(-2, 1, 0)^T$。

9.4 解 $c(1,1) + d(1,-1) = (4,2)$，两式相加得 $2c = 6$，所以 $c = 3, d = 1$；坐标为 $(3,1)^T$。

9.5 正向：若删去后的组有非平凡关系，把被删向量系数取 0，便得到原组的非平凡关系，矛盾。反向不成立；例如 $(1,0), (0,1), (1,1)$ 中删去任意一个后剩余两向量无关，但原组三者相关。

9.6 交集还要求 $y = 0$，方程变为 $x + z = 0$，所以交线为 $t(1, 0, -1)$，基为 $(1, 0, -1)^T$。

9.7 三个向量都在 xy 平面，且前两个已经张成整个 xy 平面；一组基是 e_1, e_2，第三个向量是冗余的和。

9.8 $\dim C(A) = r = 2$，$\dim N(A) = n - r = 5 - 2 = 3$；五列不可能无关，因为秩只有 2。

9.9 把每个 u_j 写成生成组的组合，得到 $k \times \ell$ 系数矩阵 C。若 $\ell > k$，未知数多于方程，齐次系统 $Ct = 0$ 有非零解，于是 $\sum t_j u_j = 0$，违背无关性。因此 $\ell \le k$；使用的矩阵事实是「未知数多于方程的齐次系统有非零解」。

9.10 令 $y = s, z = t$，则 $x = (s - t)/2$，故平面基可取 $(1, 2, 0)^T, (-1, 0, 2)^T$。与 yz 平面交集要求 $x = 0$，得到 $s = t$，交线基为 $(0, 1, 1)^T$。若 $q = (a, b, c)^T$ 垂直于平面基向量，则 $a + 2b = 0$、$-a + 2c = 0$，所以 $q = -b(2, -1, 1)^T$。法向空间基为 $(2, -1, 1)^T$。

第 10 讲

四个基本子空间

The Four Fundamental Subspaces

对应材料：MIT 18.06SC Session 1.10 摘要；ZoomNotes 第 3.5 节；习题集 10.1–10.2。

本讲目标

(1) 会从 $m \times n$ 矩阵写出四个基本子空间，并标明它们各自所在的 $\mathbb{R}^m$ 或 $\mathbb{R}^n$。

(2) 会用行最简形求行空间、列空间与零空间的基，并用消元矩阵求左零空间的基。

(3) 能证明两对正交关系，并解释为什么它们分别互为正交补。

(4) 会从秩 r 计算四个空间的维数，判断 $Ax = b$ 是否有解。

(5) 能判断 $A^{\mathsf{T}}y = d$ 的可解性与解的唯一性，并说明全部解的结构。

10.1 四个子空间分别在哪里？

对一个 3×4 矩阵 A，乘积 Ax 把输入 $x \in \mathbb{R}^4$ 送到输出 $b \in \mathbb{R}^3$。有些输入会被送到零，有些输出根本到不了；行向量和行之间的关系也各自形成子空间。先把这四种集合分别放回正确的环境空间。

定义 10.1（矩阵的四个基本子空间） 设 $A \in \mathbb{R}^{m\times n}$。$A$ 的**列空间**（column space）、**行空间**（row space）、**零空间**（nullspace）和**左零空间**（left nullspace）分别定义为

$$C(A) = \{Ax : x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m, \qquad C(A^{\mathsf{T}}) = \{A^{\mathsf{T}}y : y \in \mathbb{R}^m\} \subseteq \mathbb{R}^n,$$

$$N(A) = \{x \in \mathbb{R}^n : Ax = 0\} \subseteq \mathbb{R}^n, \qquad N(A^{\mathsf{T}}) = \{y \in \mathbb{R}^m : A^{\mathsf{T}}y = 0\} \subseteq \mathbb{R}^m.$$

命题 10.2（四个集合确实是子空间） 定义 10.1 中的四个集合都是各自环境空间的子空间。

证明 以列空间为例，$0 = A0$ 属于 $C(A)$；若 $u = Ax_1$、$v = Ax_2$，则 $\alpha u + \beta v = A(\alpha x_1 + \beta x_2)$，仍属于 $C(A)$。把 A 换成 A^{T}，同样证明行空间是子空间。

若 $x_1, x_2 \in N(A)$，则 $A(\alpha x_1 + \beta x_2) = \alpha Ax_1 + \beta Ax_2 = 0$，所以零空间对线性组合封闭且含零向量。将 A 换成 A^{T} 得左零空间也是子空间。 ■

例如取

$$A=\begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}.$$

它的四个空间为

$$C(A)=\mathbb{R}^2,\qquad C(A^{\mathsf T})=\operatorname{span}\{(1,0,0)^{\mathsf T},(0,1,0)^{\mathsf T}\},$$

$$N(A)=\operatorname{span}\{(0,0,1)^{\mathsf T}\},\qquad N(A^{\mathsf T})=\{0\}.$$

若 A 是 $m\times n$ 零矩阵，则 $C(A)=C(A^{\mathsf T})=\{0\}$，而 $N(A)=\mathbb{R}^n$、$N(A^{\mathsf T})=\mathbb{R}^m$。相反，例中矩阵 $D^{\mathsf T}$ 满列秩，$N(D^{\mathsf T})=\{0\}$；所以零空间也可能只有零向量。

这组结论可以逐项从矩阵读出。A 的前两列是 $\mathbb{R}^2$ 的标准基，故列空间为整个输出平面；$A^{\mathsf T}y=0$ 的前两行分别给出 $y_1=0,y_2=0$，所以左零空间为零。齐次方程 $Ax=0$ 则要求 $x_1=x_2=0$，留下第三个输入方向；转置的两列正是原来的两行，故行空间由前两个标准方向张成。

行空间写成 $C(A^{\mathsf T})$ 是为了继续使用列向量记号：原矩阵的每一行转置后成为 $A^{\mathsf T}$ 的一列。类似地，$A^{\mathsf T}y$ 是 A 各行的线性组合，而 Ax 是 A 各列的线性组合。两种乘法把「输入系数」和「输出向量」区分开来。

四个空间不在同一个坐标空间　把四个空间都看成 $\mathbb{R}^n$ 中的集合是错误的。上例有 $C(A),N(A^{\mathsf T})\subseteq\mathbb{R}^2$，而 $C(A^{\mathsf T}),N(A)\subseteq\mathbb{R}^3$；矩阵的列决定输出坐标数，行决定输入坐标数。

图 10.1 将四个空间按环境空间分成左右两组。对秩为 r 的矩阵，左右两组的维数分别是 $r,n-r$ 与 $r,m-r$；这些计数将在后面由消元证明。

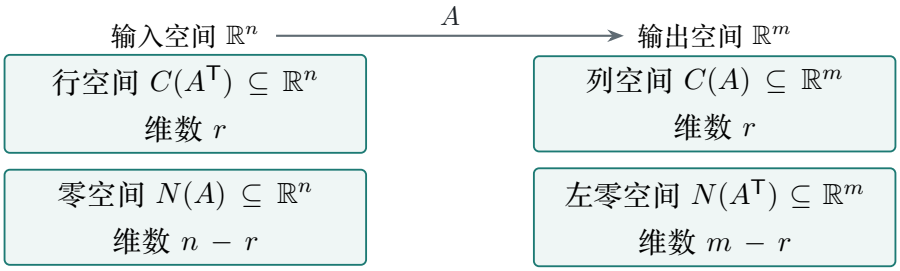


图 10.1　四个子空间按环境空间分成两组；每组维数之和等于该环境空间的维数。

图中左侧两空间由输入向量 x 描述，右侧两空间由输出或输出约束 y 描述。行空间记录矩阵实际使用的输入方向，零空间记录被 A 消去的输入方向；列空间记录能够产生的输出，左零空间记录行方程之间的依赖。

想一想　列空间是否总等于整个输出空间 $\mathbb{R}^m$？

答　不总是。若 A 的秩小于 m，列空间的维数也小于 m，所以它只是 $\mathbb{R}^m$ 的真子空间；零矩阵的列空间就是 $\{0\}$。满行秩时列空间才等于 $\mathbb{R}^m$。

10.2　行最简形怎样给出两种基?

取本讲贯穿使用的矩阵

$$A=\begin{bmatrix}1&2&3&1\\1&1&2&1\\1&2&3&1\end{bmatrix}.$$

消元得到

$$R=\begin{bmatrix}1&0&1&1\\0&1&1&0\\0&0&0&0\end{bmatrix}.$$

具体地，先作 $R_2\leftarrow R_1-R_2$、$R_3\leftarrow R_3-R_1$，再作 $R_1\leftarrow R_1-2R_2$：

$$\begin{bmatrix}1&2&3&1\\1&1&2&1\\1&2&3&1\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&3&1\\0&1&1&0\\0&0&0&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&1&1\\0&1&1&0\\0&0&0&0\end{bmatrix}.$$

主元在第 $1,2$ 列，非零行恰有两行。这里能直接从 R 读出行空间的基；列空间的基却必须回到原矩阵 A 取列。

定理 10.3（原矩阵主元列与行最简形非零行）　设 $A\in\mathbb{R}^{m\times n}$，其行最简形为 R，秩为 r，主元列编号为 $j_1,\dots,j_r$。则 A 的原矩阵主元列 $a_{j_1},\dots,a_{j_r}$ 构成 $C(A)$ 的一组基；R 的前 r 个非零行转置后构成 $C(A^{\mathsf T})$ 的一组基。因此

$$\dim C(A)=\dim C(A^{\mathsf T})=r.$$

思路　行变换由可逆矩阵 E 左乘完成，即 $R=EA$。这保留了行之间的线性组合关系，却通常改变列向量本身。证明分两部分：用 E 可逆把 R 的列关系搬回 A，再用行空间相等和主元位置读出行空间的基。

证明　**列空间。**行最简形每个非主元列都是主元列的线性组合。若 $R_j=\sum_i c_iR_{j_i}$，由 $R=EA$ 得

$$E\left(a_j-\sum_i c_ia_{j_i}\right)=0.$$

因 E 可逆，$a_j=\sum_i c_ia_{j_i}$，所以原矩阵的主元列张成 $C(A)$。这些列线性无关：若 $\sum_i d_ia_{j_i}=0$，左乘 E 得 $\sum_i d_iR_{j_i}=0$；主元列在主元行的坐标是标准基向量，故每个 $d_i=0$。

行空间。每个行变换都是可逆的行组合，所以 R 的行由 A 的行组合得到；反过来 $A=E^{-1}R$，故两者行空间相等。R 只有前 r 行非零，而且这些行在各自主元列上的坐标构成 $r\times r$ 单位矩阵，因而线性无关并张成全部行空间；将它们转置后，就得到 $C(A^{\mathsf T})$ 的一组基。 ■

■ **例 10.1　同一矩阵的行空间和列空间**　对上面的 3×4 矩阵 A，分别写出 $C(A)$ 和 $C(A^{\mathsf T})$ 的一组基，并给出它们的维数。

解　思路：主元位置从 R 读，列向量则回到 A 取。主元列是第 $1,2$ 列，所以

$$C(A)=\operatorname{span}\{(1,1,1)^{\mathsf T},(2,1,2)^{\mathsf T}\}.$$

第 3 列是前两列之和，第 4 列等于第 1 列，故这两列已张成全部列空间。行空间由 R 的前两行张成：

$$C(A^{\mathsf T})=\operatorname{span}\{(1,0,1,1)^{\mathsf T},(0,1,1,0)^{\mathsf T}\}.$$

两组基各有两个向量，因此两个空间维数都是 2。这题说明主元列编号可从 R 读取，但列空间基向量必须取自 A。 ♦

行变换不会保留列空间　把 R 的主元列直接当成 $C(A)$ 的基是错误的。此例中 R 的第一列是 $(1,0,0)^{\mathsf T}$，而 A 的第一列是 $(1,1,1)^{\mathsf T}$；前者属于 $C(R)\subseteq\mathbb{R}^3$，后者才是题目所求的原矩阵列向量。

■ **例 10.2　秩为一时的两个空间**　令 $B=\begin{bmatrix}1&2&3\\2&4&6\end{bmatrix}$。求 $C(B)$ 与 $C(B^{\mathsf T})$ 的基和维数。

解　第二行是第一行的 2 倍，第三列也是第一列的 3 倍、第二列的 $3/2$ 倍。行最简形只有一行非零行，主元列为第 1 列，所以

$$C(B)=\operatorname{span}\{(1,2)^{\mathsf T}\},\qquad C(B^{\mathsf T})=\operatorname{span}\{(1,2,3)^{\mathsf T}\}.$$

每个空间都是一条过原点的直线，维数都为 1。虽然一个空间在 $\mathbb{R}^2$、另一个在 $\mathbb{R}^3$，它们的基向量个数仍相同。 ♦

10.3　两个零空间怎样给出基?

对主矩阵的行最简形，方程 $Rx=0$ 化为

$$x_1+x_3+x_4=0,\qquad x_2+x_3=0.$$

主元变量由自由变量决定。每个自由变量单独取 1，其余取 0，会得到一个基本的零方向；左零空间则记录行方程中的线性依赖。

定理 10.4（特殊解构成零空间的基）　设 $A\in\mathbb{R}^{m\times n}$ 的秩为 r，主元变量为 $x_{j_1},\dots,x_{j_r}$，自由变量为 $x_{f_1},\dots,x_{f_{n-r}}$。令 s_k 为 $x_{f_k}=1$ 且其余自由变量为 0 时 $Ax=0$ 的解，则

$$N(A)=\operatorname{span}\{s_1,\dots,s_{n-r}\},\qquad \dim N(A)=n-r.$$

思路　行变换可逆，因此 $Ax=0$ 与 $Rx=0$ 有相同的解。任意解由其自由变量值唯一确定；把这些值拆成标准坐标方向，就得到特殊解的张成式。自由变量坐标也会直接证明这些特殊解线性无关。

证明　设自由变量值为 $t_1,\dots,t_{n-r}$。方程 $Rx=0$ 的主元变量是这些 t_k 的线性组合，所以解可写成 $x=\sum_{k=1}^{n-r} t_k s_k$。这证明特殊解张成 $N(A)$。

若 $\sum_k c_k s_k=0$，看自由变量 f_j 对应的坐标：该坐标在 s_j 中为 1，在其余特殊解中为 0，故 $c_j=0$。这对每个 j 都成立，所以特殊解线性无关，构成一组基。 ■

■ **例 10.3　主矩阵的零空间**　求贯穿本讲的矩阵 A 的零空间基，并写出全部齐次解。

解　令两个自由变量 $x_3=s$、$x_4=t$。由两条方程得到 $x_1=-s-t$、$x_2=-s$，所以

$$x=s(-1,-1,1,0)^{\mathsf T}+t(-1,0,0,1)^{\mathsf T}.$$

两向量的第 $3,4$ 坐标分别为 $(1,0)$ 和 $(0,1)$，故线性无关；直接相乘也得 $A(-1,-1,1,0)^{\mathsf T}=A(-1,0,0,1)^{\mathsf T}=0$。因此它们构成零空间的一组基，$\dim N(A)=2$。 ♦

若没有自由变量，即 $r=n$，上述基为空组，零空间只有零向量，维数为 0。这与「每个自由变量贡献一个独立零方向」一致。

定理 10.5（消元矩阵的底部各行给出左零空间基）　设可逆矩阵 $E\in\mathbb{R}^{m\times m}$ 满足 $EA=R$，其中 R 是 A 的行最简形，前 r 行非零。则 E 的最后 $m-r$ 行转置后构成 $N(A^{\mathsf T})$ 的一组基，因而

$$\dim N(A^{\mathsf T})=m-r.$$

思路　若 R 的某一行是零行，对应的 E 的行就是原矩阵各行的一个线性组合，且该组合为零。要证明这些行找全了全部依赖，取任意左零向量，用 E^{-1} 写出它在 E 的行中的系数，再利用 R 的非零行线性无关。

证明　记 E 的第 i 行为 $e_i^{\mathsf T}$。当 $i>r$ 时，$e_i^{\mathsf T}A$ 是 R 的第 i 行，等于零，所以 $A^{\mathsf T}e_i=0$。这些 e_i 线性无关，因为 E 可逆、其行线性无关。

反过来，设 $A^{\mathsf T}y=0$，即 $y^{\mathsf T}A=0$。令 $c^{\mathsf T}=y^{\mathsf T}E^{-1}$，则

$$0=y^{\mathsf T}A=c^{\mathsf T}EA=c^{\mathsf T}R.$$

R 的前 r 行线性无关，故 $c_1=\cdots=c_r=0$。于是 $y^{\mathsf T}=c^{\mathsf T}E$ 只是 E 的最后 $m-r$ 行的线性组合。两方向合起来，这些行转置后张成且线性无关，构成 $N(A^{\mathsf T})$ 的基。 ■

■ **例 10.4　从行依赖求左零空间**　对主矩阵 A，求一组左零空间基，并将它解释成 A 的行关系。

解　对增广矩阵 $[A\ I]$ 同步消元，可得

$$EA=R,\qquad E=\begin{bmatrix}-1&2&0\\1&-1&0\\-1&0&1\end{bmatrix}.$$

第三行给出 $(-1,0,1)A=0$，也就是 $-A_{1,:}+A_{3,:}=0$；这反映了 A 的第 1、第 3 行相同。因为 $m-r=3-2=1$，这一行转置 $(-1,0,1)^{\mathsf T}$ 已构成左零空间的基。 ♦

10.4　行空间与零空间为何正交?

对 A 的每一行 a_i^{T}，方程 $Ax=0$ 都要求 $a_i^{\mathsf{T}}x=0$。因此零空间里的输入方向垂直于每一行，也垂直于行空间；输出一侧也有相同关系。这里「垂直」都按标准点积理解。

定义 10.6（正交补）　设 S 是 $\mathbb{R}^q$ 的子空间。所有与 S 中每个向量都正交的向量组成的空间称为 S 的**正交补**（orthogonal complement），记为

$$S^{\perp}=\{v\in\mathbb{R}^q : v^{\mathsf{T}}s=0 \text{ 对每个} s\in S\}.$$

例如 $S=\operatorname{span}\{(1,1)^{\mathsf{T}}\}\subseteq\mathbb{R}^2$ 时，$S^{\perp}=\operatorname{span}\{(1,-1)^{\mathsf{T}}\}$，因为 $(1,1)\cdot(1,-1)=0$。零子空间的正交补是整个环境空间；整个 $\mathbb{R}^q$ 的正交补则只有零向量。

在矩阵问题里，$x\in N(A)$ 的意思正是 A 的每一行与 x 点积为零；所有行的线性组合也因此与 x 正交。反过来，只要 x 垂直于每一行，逐行相乘便得到 $Ax=0$。输出侧同理：$y\in N(A^{\mathsf{T}})$ 当且仅当 y 与 A 的每一列以及它们的所有组合都正交。

引理 10.7（正交补的维数）　若 $S\subseteq\mathbb{R}^q$ 且 $\dim S=k$，则 $\dim S^{\perp}=q-k$。

证明　把 S 的一组基作为列组成 $Q\in\mathbb{R}^{q\times k}$。向量 v 属于 $S^{\perp}$ 当且仅当 $Q^{\mathsf{T}}v=0$。Q 的列无关，故 $\operatorname{rank}Q=k$；由前述行空间与列空间维数相等，$\operatorname{rank}Q^{\mathsf{T}}=k$。零空间维数定理应用于 Q^{T}，给出 $\dim N(Q^{\mathsf{T}})=q-k$，这就是 $\dim S^{\perp}$。 ■

定理 10.8（四个基本子空间的正交关系）　对任意 $A\in\mathbb{R}^{m\times n}$，行空间与零空间互为正交补，列空间与左零空间也互为正交补：

$$C(A^{\mathsf{T}})=N(A)^{\perp}\subseteq\mathbb{R}^n,\qquad C(A)=N(A^{\mathsf{T}})^{\perp}\subseteq\mathbb{R}^m.$$

证明　若 $x\in N(A)$，$u\in C(A^{\mathsf{T}})$，则存在 y 使 $u=A^{\mathsf{T}}y$，于是 $u^{\mathsf{T}}x=y^{\mathsf{T}}Ax=0$。所以 $C(A^{\mathsf{T}})\subseteq N(A)^{\perp}$。两边维数均为 r：左边为秩定理，右边由引理 10.7 和 $\dim N(A)=n-r$ 得到。包含且维数相同，故两者相等。

若 $y\in N(A^{\mathsf{T}}),b\in C(A)$，则存在 x 使 $b=Ax$，于是 $y^{\mathsf{T}}b=(A^{\mathsf{T}}y)^{\mathsf{T}}x=0$。故 $C(A)\subseteq N(A^{\mathsf{T}})^{\perp}$。两边维数也都是 r，因为 $\dim C(A)=r$、$\dim N(A^{\mathsf{T}})=m-r$；引理给出另一边维数为 r，所以相等。 ■

取 $H=\begin{bmatrix}1&1\\1&1\end{bmatrix}$。它的每一列和每一行都沿着 $(1,1)^{\mathsf{T}}$，而 $(1,-1)^{\mathsf{T}}$ 被 H 和 H^{T} 都送到零。图 10.2 把输入侧与输出侧的两对直线并排画出。

例 10.5　用一组直线检查两对正交空间　求 H 的四个基本子空间，并逐对验证它们正交。

解　H 的两列相同且非零，所以 $C(H)=\operatorname{span}\{(1,1)^{\mathsf{T}}\}$。行也相同，故 $C(H^{\mathsf{T}})=\operatorname{span}\{(1,1)^{\mathsf{T}}\}$。

解 $Hx=0$ 得 $x_1+x_2=0$，于是 $N(H)=\operatorname{span}\{(1,-1)^{\mathsf T}\}$；同理 $N(H^{\mathsf T})=\operatorname{span}\{(1,-1)^{\mathsf T}\}$。点积

$$(1,1)\cdot(1,-1)=1-1=0$$

验证了两对空间都正交。它们各自维数为 1，在 $\mathbb{R}^2$ 中互为正交补。 ♦

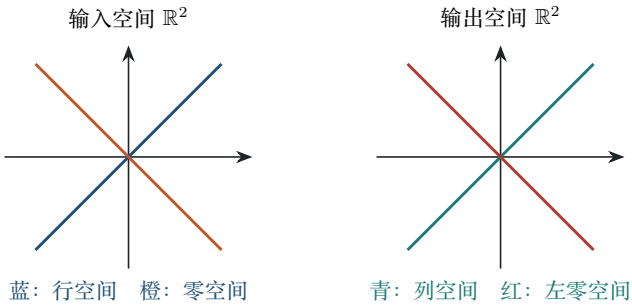


图 10.2　矩阵 H 的行空间与零空间、列空间与左零空间分别是两条互相垂直的直线。

图中的每个平面都是一个 $\mathbb{R}^2$。两条方向线斜率分别为 1 与 -1，对应向量 $(1,1)^{\mathsf T}$ 和 $(1,-1)^{\mathsf T}$；它们既互相垂直，也恰好张成整个环境平面。

想一想　两条子空间互相正交，就一定互为正交补吗?

答　不一定。$\operatorname{span}\{(1,0,0)^{\mathsf T}\}$ 与 $\operatorname{span}\{(0,1,0)^{\mathsf T}\}$ 在 $\mathbb{R}^3$ 中正交，但它们只张成 xy 平面，第三坐标方向仍在正交补里。前述两对基本子空间之所以互为正交补，还用到了维数之和等于环境空间维数。

10.5　秩如何分配四个空间的维数?

一个 $m\times n$ 矩阵有 n 个输入变量、m 个方程。若消元留下 r 个主元，输入端就有 $n-r$ 个自由方向；输出端则有 $m-r$ 条独立的行关系。行空间和列空间都恰有 r 个独立方向。

定理 10.9（四个基本子空间的维数）　设 $A\in\mathbb{R}^{m\times n}$ 的秩为 r。则

$$\dim C(A)=r,\quad \dim C(A^{\mathsf T})=r,\quad \dim N(A)=n-r,\quad \dim N(A^{\mathsf T})=m-r.$$

证明　定理 10.3 给出列空间和行空间各有 r 个基向量。行最简形有 r 个主元、$n-r$ 个自由变量，所以定理 10.4 给出 $\dim N(A)=n-r$。消元矩阵 E 有 $m-r$ 个对应零行，定理 10.5 给出 $\dim N(A^{\mathsf T})=m-r$。 ■

于是

$$r\le m,\qquad r\le n,\qquad \dim C(A)+\dim N(A)=n,\qquad \dim C(A^{\mathsf T})+\dim N(A^{\mathsf T})=m.$$

前两个不等式表示主元数不能超过行数或列数；后两个等式分别把输入维数和输出维数分成正交的两部分。特别地，$n>m$ 时 $N(A)$ 一定含非零向量；$r<m$ 时 $N(A^{\mathsf T})$ 一定含非零向量。

若 $r=n$，矩阵满列秩，$N(A)=\{0\}$，所以每个可解的 $Ax=b$ 都有唯一解；这并不保证每个 $b\in\mathbb{R}^m$ 都可达到。若 $r=m$，矩阵满行秩，$C(A)=\mathbb{R}^m$ 且 $N(A^{\mathsf{T}})=\{0\}$，所以每个 b 都可达到，而每个可解的 $A^{\mathsf{T}}y=d$ 都有唯一解。只有在 $m=n=r$ 时，两种满秩条件同时成立。

例如前面的 G 是 3×2 满列秩矩阵，$Gx=0$ 只有零解，但仍有右端项无法达到；K 是 2×3 满行秩矩阵，任意 $\mathbb{R}^2$ 右端项都能达到，但齐次系统有一个自由输入方向。矩阵是否方阵，不能替代对秩的检查。

■ **例 10.6**　**有不可解右端项意味着什么？**　设 A 是 $m\times n$ 矩阵，且至少有一个 $b\in\mathbb{R}^m$ 使 $Ax=b$ 无解。求 m,n,r 之间必然成立的不等式，并说明 $A^{\mathsf{T}}y=0$ 有非零解。

解　可解的右端项恰好是 $C(A)$ 中的向量。存在不可解的 b 意味着 $C(A)\neq\mathbb{R}^m$，所以 $\dim C(A)=r<m$。秩不超过列数，故 $r\le n$；综上，

$$0\le r\le n,\qquad r<m.$$

由维数定理，$\dim N(A^{\mathsf{T}})=m-r>0$，因此齐次方程 $A^{\mathsf{T}}y=0$ 有非零解。这正是左零空间维数为正的含义。 ♦

■ **例 10.7**　**一个三行两列的不可解系统**　令 $G=\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}$，考察 $Gx=(0,0,1)^{\mathsf{T}}$。求 r、四个空间的维数，并找出一个非零左零向量。

解　两列线性无关，故 $r=2$。因此 $C(G)$、$C(G^{\mathsf{T}})$、$N(G)$、$N(G^{\mathsf{T}})$ 的维数依次为 $2,2,0,1$。若 $Gx=(0,0,1)^{\mathsf{T}}$，前两行迫使 $x_1=x_2=0$，第三行便给出 $0=1$，所以无解。

取 $y=(-1,-1,1)^{\mathsf{T}}$，则 $G^{\mathsf{T}}y=(0,0)^{\mathsf{T}}$，但 $y\neq 0$。而且 $y^{\mathsf{T}}(0,0,1)^{\mathsf{T}}=1$，所以这个右端项没有与所有左零向量正交，也就不在 $C(G)=N(G^{\mathsf{T}})^{\perp}$ 中。 ♦

10.6　$Ax=b$ 何时有解？

Ax 是 A 的列以 x 为系数的线性组合，所以右端项 b 能否达到，先看它是否落在列空间。消元提供了逐行检查的方法；左零空间则把所有相容条件一次写全。

定理 10.10（相容条件与全部解）　设 $A\in\mathbb{R}^{m\times n}$、$b\in\mathbb{R}^m$。方程 $Ax=b$ 有解，当且仅当

$$b\in C(A),$$

也当且仅当 $y^{\mathsf{T}}b=0$ 对每个 $y\in N(A^{\mathsf{T}})$ 成立。若 x_p 是一个特解，则全部解为 $x_p+N(A)$；可解时解唯一，当且仅当 $N(A)=\{0\}$。

思路　列空间条件直接来自 Ax 是列组合。左零条件由四空间正交补等式改写；全部解则比较任意解和一个特解，两者之差满足齐次方程。唯一性只需检查齐次解是否只有零解。

证明　$Ax=b$ 有解的定义就是存在 x 使 $b=Ax$，等价于 $b\in C(A)$。又由定理 10.8，$C(A)=N(A^{\mathsf T})^{\perp}$，所以它等价于 b 与每个左零向量正交。

若 x_p 是特解，则 $Ax=b$ 等价于 $A(x-x_p)=0$，也就是 $x-x_p\in N(A)$。所以全部解恰为 $x_p+N(A)$。若零空间仅含零向量，这个集合只有 x_p；若零空间含非零 z，则 x_p 和 x_p+z 是两个不同解，故不唯一。 ■

更具体地，若 $EA=R$ 且 R 的底部 $m-r$ 行为零，则增广矩阵消元给出 $Rx=Eb$。底部零行要求 $(Eb)_{r+1},\dots,(Eb)_m$ 全为零；任何一个非零值都会化成矛盾 $0=c$。这些正是 $N(A^{\mathsf T})$ 基向量与 b 的点积条件。

对主矩阵，第三行是第一行的重复，因此每个列组合都满足 $b_3=b_1$。反过来，若 $b=(u,v,u)^{\mathsf T}$，取第一列系数 $2v-u$、第二列系数 $u-v$，就得到 $Ax=b$。所以本例的列空间恰为 $\{(u,v,u)^{\mathsf T}:u,v\in\mathbb{R}\}$，唯一的独立相容条件是 $b_3-b_1=0$；这也正是消元矩阵底行 $(-1,0,1)$ 与 b 的点积。

■ **例 10.8**　**同一系数矩阵的相容与不相容右端项**　对主矩阵 A，分别求解 $Ax=b_1$ 与 $Ax=b_2$，其中 $b_1=(1,5,1)^{\mathsf T}$、$b_2=(1,5,2)^{\mathsf T}$；若有解，写出全部解。

解　继续使用消元矩阵 E。对 b_1，

$$Eb_1=(9,-4,0)^{\mathsf T}.$$

方程 $Rx=Eb_1$ 化为 $x_1+x_3+x_4=9$、$x_2+x_3=-4$。令 $x_3=s$、$x_4=t$，得到全部解

$$x=(9,-4,0,0)^{\mathsf T}+s(-1,-1,1,0)^{\mathsf T}+t(-1,0,0,1)^{\mathsf T}.$$

取 $s=t=0$ 得特解 $(9,-4,0,0)^{\mathsf T}$，直接相乘为 $A(9,-4,0,0)^{\mathsf T}=(1,5,1)^{\mathsf T}$。

对 b_2，$Eb_2=(9,-4,1)^{\mathsf T}$。第三行原本是零行，却要求 $0=1$，所以无解。对应左零向量 $y=(-1,0,1)^{\mathsf T}$ 满足 $y^{\mathsf T}b_2=1\neq0$，也从正交条件排除了这个右端项。 ♦

> **命题 10.11（转置方程的可解性与唯一性）**　方程 $A^{\mathsf T}y=d$ 有解，当且仅当 $d\in C(A^{\mathsf T})$，也就是 d 在 A 的行空间中。若 y_p 是一个特解，则全部解为 $y_p+N(A^{\mathsf T})$；可解时解唯一，当且仅当 $N(A^{\mathsf T})=\{0\}$。

证明　按定义，$C(A^{\mathsf T})$ 正是所有 $A^{\mathsf T}y$ 的集合。若 y_p 是一个特解，则 $A^{\mathsf T}y=d$ 等价于 $A^{\mathsf T}(y-y_p)=0$，故 $y-y_p\in N(A^{\mathsf T})$。因此解集为 $y_p+N(A^{\mathsf T})$，其只有一个元素恰好当且仅当左零空间为零空间。 ■

■ **例 10.9**　**行空间中的向量有唯一原像吗？**　令 $K=\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}$，解 $K^{\mathsf T}y=d$，其中 $d=(2,-1,1)^{\mathsf T}$，并判断解是否唯一。

解　逐坐标写出 $K^{\mathsf T}y=(y_1,y_2,y_1+y_2)^{\mathsf T}$。前两坐标要求 $y_1=2$、$y_2=-1$，第三坐标随即为 1，所以唯一解是 $y=(2,-1)^{\mathsf T}$。从四空间看，K 的秩为 $2=m$，故 $\dim N(K^{\mathsf T})=m-r=0$，与唯一性判据一致。 ♦

唯一性条件不表示每个 $d \in \mathbb{R}^3$ 都有原像。对同一个 K，$K^{\mathsf{T}}y = (0,0,1)^{\mathsf{T}}$ 的前两坐标迫使 $y = (0,0)^{\mathsf{T}}$，第三坐标却应为 0 而不是 1，因此这个 d 不在行空间中，方程无解。只有先确认 $d \in C(K^{\mathsf{T}})$，再用 $N(K^{\mathsf{T}}) = \{0\}$，才能得出唯一解。

■ **例 10.10　同一个行向量可以有多种系数**　对主矩阵 A，求 $A^{\mathsf{T}}y = (1,2,3,1)^{\mathsf{T}}$ 的全部解。

解　右端项正是 A 的第一行转置，所以 $y_p = (1,0,0)^{\mathsf{T}}$ 是一个解。左零空间基为 $(-1,0,1)^{\mathsf{T}}$，故全部解为

$$y = (1,0,0)^{\mathsf{T}} + t(-1,0,1)^{\mathsf{T}}, \qquad t \in \mathbb{R}.$$

由于左零空间非零，解不唯一；例如 $t = 0$ 和 $t = 1$ 给出两个不同系数向量，却产生同一个 d。 ♦

两种方程的唯一性看不同零空间　$Ax = b$ 的唯一性由 $N(A)$ 决定，不能拿 $N(A^{\mathsf{T}})$ 代替。上例矩阵 G 的左零空间有非零向量 $(-1,-1,1)^{\mathsf{T}}$，但 $Gx = 0$ 仍只有唯一解 $x = 0$，因为 G 的两列线性无关。转置方程 $A^{\mathsf{T}}y = d$ 的唯一性才由 $N(A^{\mathsf{T}})$ 决定。

想一想　检查 b 是否与左零空间中每个向量正交，需要检验无穷多个向量吗?
答　不需要。若 $y_1, \dots, y_{m-r}$ 是左零空间的一组基，只要逐个检查 $y_i^{\mathsf{T}}b = 0$；任意左零向量都是它们的线性组合，点积也随之为零。

方法：先看列空间，再看零空间　判断 $Ax = b$ 是否可解，先检查 b 是否满足左零空间给出的相容条件；若可解，一个特解加上 $N(A)$ 就是全部解。唯一性只需再问 $N(A)$ 是否为零空间。

10.7　矩阵集合也有基和维数

把 3×3 矩阵的九个位置当成九个坐标，矩阵加法和数乘逐项进行。全体实 3×3 矩阵构成一个矩阵空间（matrix space）$\mathcal{M}_{3\times 3}$，其维数为 9。记 E_{ij} 为第 (i,j) 个位置是 1、其他位置全为 0 的矩阵；九个 E_{ij} 张成所有矩阵，并由九个互不相同的位置保证线性无关，因而构成一组基。

具体地，任意 $X = (x_{ij})$ 都有唯一展开 $X = \sum_{i=1}^{3}\sum_{j=1}^{3} x_{ij}E_{ij}$。这与普通坐标向量完全相同，只是每个「坐标方向」现在是一张矩阵。上三角和对称条件分别删去或绑定某些坐标，因此可以直接数出相应基的个数。

上三角矩阵由条件 $a_{21} = a_{31} = a_{32} = 0$ 定义，其基可取

$$E_{11}, E_{12}, E_{13}, E_{22}, E_{23}, E_{33},$$

所以维数为 6。对称矩阵满足 $a_{ij} = a_{ji}$；三个对角位置可独立取值，三个非对角位置成对相等，故一组基为

$$E_{11}, E_{22}, E_{33}, E_{12}+E_{21}, E_{13}+E_{31}, E_{23}+E_{32},$$

维数也为 6。这些集合都是子空间：零矩阵满足相应的零约束；若 U, V 满足约束，则任意 $\alpha U + \beta V$ 仍满足。对称性也保持，因为 $(\alpha U + \beta V)^{\mathsf{T}} = \alpha U^{\mathsf{T}} + \beta V^{\mathsf{T}} = \alpha U + \beta V$。各组基的矩阵由不同位置或不同成对位置支撑，故系数逐项比较可知线性无关。

■ **例 10.11　两个矩阵子空间的交集**　求上三角矩阵空间与对称矩阵空间的交集，并判断

$$D=\begin{bmatrix}1&0&0\\0&3&0\\0&0&7\end{bmatrix}$$

是否属于交集。

解　交集中的矩阵既满足 $a_{ij}=0\ (i>j)$，又满足 $a_{ij}=a_{ji}$。若 $i<j$，对称性给出 $a_{ij}=a_{ji}$，而上三角条件使右边为 0，所以所有对角线外的位置都为 0。因此交集恰好是对角矩阵空间

$$\mathcal{D}=\operatorname{span}\{E_{11},E_{22},E_{33}\},\qquad \dim\mathcal{D}=3.$$

矩阵 D 满足 $D^{\mathsf T}=D$，且下三角三个非对角位置为零，所以属于交集。展开式 $D=E_{11}+3E_{22}+7E_{33}$ 的系数就是它的三个对角元素；三个基矩阵分别占据不同位置，故展开唯一。交集的维数要通过交集本身的条件计数，不能把两个子空间的维数直接相加。 ♦

两个子空间的交集仍是子空间：零矩阵属于两者，而两者中任意两个矩阵的线性组合仍同时满足两边的条件。对角矩阵空间正是这一一般事实在上三角与对称矩阵上的具体结果。

想一想　上三角空间与对称空间各有维数 6，为什么交集不是 6 维?
答　对称性把每个上三角非对角元与其下三角位置绑在一起；上三角条件又把下三角位置置零，于是相应的上三角元也必须为零。交集中只剩三个独立对角元，所以维数是 3。

本讲小结

- 对 $A\in\mathbb{R}^{m\times n}$，$C(A),N(A^{\mathsf T})\subseteq\mathbb{R}^m$；$C(A^{\mathsf T}),N(A)\subseteq\mathbb{R}^n$。
- 行最简形的主元列编号定位 $C(A)$ 的基，但基向量取原矩阵列；行最简形前 r 个非零行转置后给出 $C(A^{\mathsf T})$ 的基。
- 自由变量的特殊解构成 $N(A)$ 的基；消元矩阵中对应零行的行向量转置后构成 $N(A^{\mathsf T})$ 的基。
- 若 $\operatorname{rank}A=r$，四个维数依次为 $r,r,n-r,m-r$；输入端与输出端的两对正交补分别分解 $\mathbb{R}^n$ 与 $\mathbb{R}^m$。
- $Ax=b$ 可解当且仅当 $b\in C(A)=N(A^{\mathsf T})^{\perp}$；若 x_p 是特解，全部解为 $x_p+N(A)$。
- 可解的 $Ax=b$ 唯一当且仅当 $N(A)=\{0\}$；$A^{\mathsf T}y=d$ 可解当且仅当 $d\in C(A^{\mathsf T})$，唯一当且仅当 $N(A^{\mathsf T})=\{0\}$。

练习

10.1 ★　判断并说明：(a) 四个基本子空间都在 $\mathbb{R}^n$ 中；(b) $\dim C(A)=\dim C(A^{\mathsf T})$；(c) $N(A)$ 与 $C(A)$ 总能直接作点积；(d) $Ax=b$ 可解当且仅当 $b\in C(A)$。

10.2 ★　A 是 5×7 矩阵且 $\operatorname{rank}A=3$。求四个基本子空间的维数。

10.3 ★ 对 $H=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$，求行空间与零空间的一组基，并验证它们正交。

10.4 ★★ 对主矩阵 A，求 $C(A)$ 与 $C(A^{\mathsf T})$ 的一组基，并解释为什么列空间基不能直接取自行最简形。

10.5 ★★ 对主矩阵 A，求 $N(A)$ 的全部解及一组基。

10.6 ★★ 对主矩阵 A，求 $N(A^{\mathsf T})$ 的一组基，并将其写成行向量依赖关系。

10.7 ★★ 设 A 是 $m\times n$ 矩阵且某个 $Ax=b$ 无解。证明必须有 $r<m$、$r\le n$，并说明 $A^{\mathsf T}y=0$ 有非零解。

10.8 ★★ 对主矩阵 A，判断 $Ax=(1,5,1)^{\mathsf T}$ 与 $Ax=(1,5,2)^{\mathsf T}$ 是否可解；给出可解系统的全部解。

10.9 ★★ 填空并证明：$A^{\mathsf T}y=d$ 可解当且仅当 d 属于 _____；若可解，解唯一当且仅当 _____ 只含零向量。

练习答案与提示

10.1 (a) 假，$C(A),N(A^{\mathsf T})$ 在 $\mathbb{R}^m$，而 $C(A^{\mathsf T}),N(A)$ 在 $\mathbb{R}^n$。(b) 真，两者维数均为 r。(c) 假；非方阵时两个空间可能处于不同环境空间，正确的正交关系是 $C(A^{\mathsf T})\perp N(A)$ 与 $C(A)\perp N(A^{\mathsf T})$。(d) 真，因为 Ax 正是 A 的列以 x 为系数的线性组合。

10.2 维数依次为 $\dim C(A)=3$、$\dim C(A^{\mathsf T})=3$、$\dim N(A)=7-3=4$、$\dim N(A^{\mathsf T})=5-3=2$。

10.3 $C(H^{\mathsf T})=\operatorname{span}\{(1,1)^{\mathsf T}\}$，$N(H)=\operatorname{span}\{(1,-1)^{\mathsf T}\}$。点积为 $1\cdot1+1\cdot(-1)=0$，故行空间与零空间正交。

10.4 主元列为第 $1,2$ 列，两空间的基可并列写为

$$C(A)=\operatorname{span}\{(1,1,1)^{\mathsf T},(2,1,2)^{\mathsf T}\},$$
$$C(A^{\mathsf T})=\operatorname{span}\{(1,0,1,1)^{\mathsf T},(0,1,1,0)^{\mathsf T}\}.$$

第 3 列等于前两列之和，第 4 列等于第 1 列。行变换左乘可逆矩阵会改变列向量本身，所以只能从 R 读主元列编号，基向量要回到 A 取。

10.5 令 $x_3=s,x_4=t$，则 $x_1=-s-t,x_2=-s$。全部解为 $x=s(-1,-1,1,0)^{\mathsf T}+t(-1,0,0,1)^{\mathsf T}$，基为这两个向量。代入 A 得两者都映到零；它们在第 $3,4$ 个坐标分别为 $(1,0)$、$(0,1)$，故线性无关。

10.6 基为 $(-1,0,1)^{\mathsf T}$，因为 A 的第 3 行减第 1 行为零，即 $(-1,0,1)A=0$。左零空间维数 $m-r=1$，所以这一个非零向量已张成全部左零空间。

10.7 由不可解右端项知 $C(A)\neq\mathbb{R}^m$，故 $r=\dim C(A)<m$；又 r 不超过列数 n，故 $0\le r\le n$ 且 $r<m$。维数定理给 $\dim N(A^{\mathsf T})=m-r>0$，所以 $A^{\mathsf T}y=0$ 有非零解。

10.8 令 $E=\begin{bmatrix}-1 & 2 & 0\\ 1 & -1 & 0\\ -1 & 0 & 1\end{bmatrix}$，则 $EA=R$。对 $b_1=(1,5,1)^{\mathsf T}$，$Eb_1=(9,-4,0)^{\mathsf T}$，故 $x=(9,-4,0,0)^{\mathsf T}+s(-1,-1,1,0)^{\mathsf T}+t(-1,0,0,1)^{\mathsf T}$ 是全部解；取 $s=t=0$ 可验 $Ax=b_1$。对 $b_2=(1,5,2)^{\mathsf T}$，$Eb_2=(9,-4,1)^{\mathsf T}$，底部零行要求 $0=1$，因此无解。

10.9 $A^{\mathsf T}y=d$ 可解当且仅当 $d\in C(A^{\mathsf T})$，即 d 在行空间中。若 y_p 是特解，则全部解为 $y_p+N(A^{\mathsf T})$，所以唯一当且仅当 $N(A^{\mathsf T})=\{0\}$。

第 11 讲

矩阵空间、秩 1 矩阵与小世界图

Matrix Spaces; Rank 1; Small World Graphs

对应材料：MIT 18.06SC Session 1.11 课时摘要（对应 Strang 第 11 讲视频 Matrix Spaces; Rank 1; Small World Graphs）；Session 1.11 习题 11.1–11.2 及解答；Strang《Introduction to Linear Algebra》第 3.5–3.6 节。

本讲目标

(1) 把「向量」推广到矩阵和函数：会判断一族矩阵或函数是否构成向量空间，并求出基与维数。

(2) 掌握 3×3 矩阵空间 $\mathbf{M}$ 及其子空间 $\mathbf{S}$（对称）、$\mathbf{U}$（上三角）、$\mathbf{D}$（对角）的基与维数。

(3) 分清子空间的并与和，会用维数公式 $\dim V+\dim W=\dim(V+W)+\dim(V\cap W)$ 做计算。

(4) 理解 $y''+y=0$ 的解构成以 $\cos x,\sin x$ 为基的二维空间。

(5) 掌握秩 1 矩阵 $A=uv^{\mathsf{T}}$ 的结构：四个子空间、$A^2=(v^{\mathsf{T}}u)A$，以及「秩 r 矩阵是 r 个秩 1 矩阵之和」。

(6) 认识图、距离与「小世界」现象，会用邻接矩阵的幂求两点间的距离。

第 9、10 讲建立了线性无关、基、维数和四个基本子空间的理论，舞台一直是 $\mathbb{R}^n$。这些概念真正用到的只有两种运算：加法和数乘。只要一族对象能相加、能乘以数，并且遵守通常的运算律，整套理论就可以原样搬过去。

本讲先看两个新的「向量空间」：一个由矩阵组成，一个由函数组成。随后转向一类特别简单的矩阵，秩为 1 的矩阵，它们是搭建所有矩阵的积木。最后一节讲图和「小世界」，为第 12 讲用矩阵研究图做准备。和第 6 讲一样，列向量常写成 $(1,2,3)$ 的横排形式以节省篇幅。

11.1 矩阵空间

11.1.1 矩阵也是向量

记 $\mathbf{M}$ 为全体 3×3 实矩阵组成的集合。两个 3×3 矩阵可以逐个元素相加，一个矩阵可以逐个元素乘以实数，结果仍是 3×3 矩阵：

$$A,B\in\mathbf{M},\ c\in\mathbb{R}\quad\Longrightarrow\quad A+B\in\mathbf{M},\quad cA\in\mathbf{M}.$$

加法交换律、结合律、数乘分配律等八条运算律逐个元素成立，零矩阵充当零向量。所以 $\mathbf{M}$ 是一个向量空间，它的「向量」是矩阵。

定义 11.1（矩阵空间）　全体 $m\times n$ 实矩阵在矩阵加法和数乘下构成向量空间，称为**矩阵空间**（matrix space），记作 $\mathbf{M}_{m\times n}$。本讲 $\mathbf{M}$ 默认指 $\mathbf{M}_{3\times 3}$。

在这个空间里，矩阵乘法 AB 暂时用不上。向量空间只关心线性组合 $cA+dB$。两个矩阵相乘是矩阵这种对象额外具有的结构，与「$\mathbf{M}$ 是向量空间」无关。

11.1.2 标准基与维数

记 E_{ij} 为第 (i,j) 个元素是 1、其余元素全是 0 的矩阵。任何 3×3 矩阵都能写成

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix}=\sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}E_{ij},$$

组合系数就是矩阵的元素。这 9 个矩阵还线性无关：若 $\sum c_{ij}E_{ij}=0$，左边是第 (i,j) 个元素为 c_{ij} 的矩阵，它等于零矩阵就迫使所有 $c_{ij}=0$。所以 $\{E_{ij}\}$ 是 $\mathbf{M}$ 的一组基，称为**标准基**（standard basis）。

$$\boxed{\dim \mathbf{M}_{3\times 3}=9,\qquad \dim \mathbf{M}_{m\times n}=mn.}$$

直观理解　把 3×3 矩阵的 9 个元素逐行抄成一列，A 就变成 $\mathbb{R}^9$ 中的向量 $(a_{11},a_{12},a_{13},a_{21},\dots,a_{33})$。矩阵加法和数乘在这种「拉直」下就是 $\mathbb{R}^9$ 里的加法和数乘，E_{ij} 对应 $\mathbb{R}^9$ 的标准基向量。所以关于 $\mathbf{M}$ 的线性无关、张成、坐标问题，都可以拉直成 $\mathbb{R}^9$ 里的消元问题来解。

例 11.1　**判断一组 2×2 矩阵是不是基**　设

$$B_1=\begin{bmatrix}1&1\\0&0\end{bmatrix},\quad B_2=\begin{bmatrix}0&1\\1&0\end{bmatrix},\quad B_3=\begin{bmatrix}0&0\\1&1\end{bmatrix},\quad B_4=\begin{bmatrix}1&0\\0&0\end{bmatrix}.$$

(1) 证明 B_1,B_2,B_3,B_4 是 $\mathbf{M}_{2\times 2}$ 的一组基，并求 $A=\begin{bmatrix}3&1\\2&5\end{bmatrix}$ 在这组基下的坐标。(2) 若把 B_4 换成单位矩阵 I，还是基吗?

解　(1) 先写出一般的组合：

$$c_1B_1+c_2B_2+c_3B_3+c_4B_4=\begin{bmatrix}c_1+c_4 & c_1+c_2\\ c_2+c_3 & c_3\end{bmatrix}.$$

令它等于零矩阵。从右下角得 $c_3=0$，左下角得 $c_2=0$，右上角得 $c_1=0$，左上角得 $c_4=0$。所以四个矩阵线性无关。$\dim \mathbf{M}_{2\times 2}=4$，4 个线性无关的向量自动张成整个空间（第 9 讲），因此它们构成一组基。

求坐标就是令组合等于 A，按同样的顺序回代：

$$c_3 = 5, \qquad c_2 = 2 - c_3 = -3, \qquad c_1 = 1 - c_2 = 4, \qquad c_4 = 3 - c_1 = -1.$$

所以 $A = 4B_1 - 3B_2 + 5B_3 - B_4$，坐标是 $(4, -3, 5, -1)$。验算左上角：$4 - 1 = 3$，与 A 一致。

(2) 换成 I 后，组合变成

$$c_1B_1 + c_2B_2 + c_3B_3 + c_4I = \begin{bmatrix} c_1 + c_4 & c_1 + c_2 \\ c_2 + c_3 & c_3 + c_4 \end{bmatrix}.$$

令它为零：$c_1 = -c_4$，$c_2 = -c_1 = c_4$，$c_3 = -c_2 = -c_4$，最后一个方程 $c_3 + c_4 = 0$ 自动成立。c_4 可以任取，取 $c_4 = -1$ 得非零解 $(1, -1, 1, -1)$，即

$$B_1 - B_2 + B_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

四个矩阵线性相关，不是基。拉直成 $\mathbb{R}^4$ 的向量后，这四个向量组成的矩阵秩为 3。 ♦

例 11.1 的两问只差一个矩阵，结论却相反。判断依据始终是同一件事：齐次方程组「组合 $= 0$」有没有非零解。这和第 9 讲在 $\mathbb{R}^n$ 中的做法完全相同。

11.1.3 对称、上三角与对角矩阵

$\mathbf{M}$ 里有三个重要的子集：

$$\mathbf{S} = \{\, A \in \mathbf{M} \mid A^{\mathsf{T}} = A \,\}, \qquad \mathbf{U} = \{\, A \in \mathbf{M} \mid a_{ij} = 0\ (i > j) \,\}, \qquad \mathbf{D} = \{\, A \in \mathbf{M} \mid a_{ij} = 0\ (i \neq j) \,\},$$

分别是对称矩阵、上三角矩阵和对角矩阵。它们都是子空间。以 $\mathbf{S}$ 为例：若 $A^{\mathsf{T}} = A$，$B^{\mathsf{T}} = B$，则 $(cA + dB)^{\mathsf{T}} = cA^{\mathsf{T}} + dB^{\mathsf{T}} = cA + dB$，组合仍对称。$\mathbf{U}$ 和 $\mathbf{D}$ 的条件是「某些位置为 0」，两个矩阵在这些位置为 0，它们的任意组合在这些位置也为 0。

求维数的办法是数「自由元素」的个数，再把每个自由元素对应成一个基矩阵。

上三角矩阵 $\mathbf{U}$。 对角线及其上方的 6 个元素可以任取，下方 3 个元素必须为 0。基就取标准基中的 6 个：

$$E_{11},\ E_{12},\ E_{13},\ E_{22},\ E_{23},\ E_{33}, \qquad \dim \mathbf{U} = 6.$$

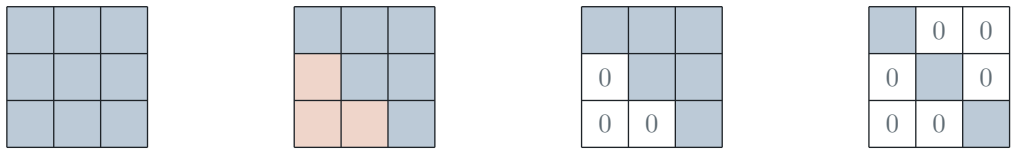
对称矩阵 $\mathbf{S}$。 对称矩阵由对角线及其上方的 6 个元素决定，下方元素是上方的镜像：$a_{21} = a_{12}$，$a_{31} = a_{13}$，$a_{32} = a_{23}$。一个自由的非对角元素 a_{12} 同时出现在两个位置，对应的基矩阵是 $E_{12} + E_{21}$。于是

$$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} = aE_{11} + dE_{22} + fE_{33} + b(E_{12} + E_{21}) + c(E_{13} + E_{31}) + e(E_{23} + E_{32}),$$

这 6 个矩阵张成 $\mathbf{S}$。若组合为零矩阵，则 $a, b, \ldots, f$ 作为元素全为 0，所以它们线性无关。$\dim \mathbf{S} = 6$。

对角矩阵 **D**。　只有对角线上 3 个元素自由，基为 E_{11}, E_{22}, E_{33}，$\dim \mathbf{D} = 3$。

图 11.1 把四类矩阵的自由位置并排标出：蓝格可任取，橙格由对称条件复制，标 0 的格必须为零。



M：$\dim = 9$　　**S**：$\dim = 6$　　**U**：$\dim = 6$　　**D**：$\dim = 3$

图 11.1　$\mathbf{M}, \mathbf{S}, \mathbf{U}, \mathbf{D}$ 中的自由元素。蓝色格子可任取；橙色格子由对称性决定（等于对角线另一侧的元素）；写 0 的格子必须为零。维数就是蓝色格子的个数。

同样的计数对 $n \times n$ 矩阵成立。对角线及其上方共有 $1 + 2 + \cdots + n$ 个位置，所以

$$\dim \mathbf{M}_{n\times n} = n^2, \qquad \dim \mathbf{S}_n = \dim \mathbf{U}_n = \frac{n(n+1)}{2}, \qquad \dim \mathbf{D}_n = n.$$

$n = 3$ 时依次是 $9, 6, 6, 3$。

注 11.1　**U** 和 **D** 的基都可以直接从标准基 $\{E_{ij}\}$ 中挑出来，**S** 却不行。9 个 E_{ij} 中只有 E_{11}, E_{22}, E_{33} 是对称矩阵，它们只张成 3 维的 **D**，凑不出 6 维的 **S**。所以**大空间的一组基，一般不能删减成子空间的基**。反方向总是可行的：子空间的任一组基都能扩充成大空间的基（第 9 讲）。例如在 **S** 的 6 个基矩阵之外添上 E_{12}, E_{13}, E_{23}，就得到 **M** 的一组基，因为 $E_{21} = (E_{12} + E_{21}) - E_{12}$ 等也都能组合出来。

想一想　全体 3×3 可逆矩阵组成的集合是 **M** 的子空间吗？全体迹为 0 的 3×3 矩阵呢？
答　可逆矩阵不是子空间：零矩阵不可逆，不在集合里；也可以看 $I + (-I) = 0$，两个可逆矩阵之和不可逆。迹为 0 的矩阵是子空间，因为 $\operatorname{tr}(cA + dB) = c \operatorname{tr} A + d \operatorname{tr} B$，条件 $\operatorname{tr} A = 0$ 是一个线性方程。它是 $\mathbb{R}^9$ 中一个方程的解集，维数为 $9 - 1 = 8$。

■ 例 11.2　**固定秩集合与线性约束**　说明全体秩为 4 的 5×17 矩阵不是子空间（即使把零矩阵也添进去）。再求 $A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ 的四个基本子空间的维数，并给出 $N(A)$ 的一组基。

解　先看两个矩阵

$$X = \begin{bmatrix} \operatorname{diag}(1,1,1,1,0) & 0_{5\times 12} \end{bmatrix}, \qquad Y = \begin{bmatrix} \operatorname{diag}(1,1,1,-1,0) & 0_{5\times 12} \end{bmatrix}.$$

它们的秩都是 4，而 $X + Y = \begin{bmatrix} \operatorname{diag}(2,2,2,0,0) & 0_{5\times 12} \end{bmatrix}$ 的秩是 3。所以这个集合对加法不封闭；添入零矩阵也不能改变这一点。

其次，$N(A)$ 是 $\mathbb{R}^4$ 中满足 $x_1 + x_2 + x_3 + x_4 = 0$ 的超平面。令自由变量 x_2, x_3, x_4 依次取 1，得到基

$$(-1,1,0,0), \qquad (-1,0,1,0), \qquad (-1,0,0,1),$$

故 $\dim N(A) = 3$。$C(A) = \mathbb{R}$，$C(A^{\mathsf{T}}) = \operatorname{span}\{(1,1,1,1)\}$，$N(A^{\mathsf{T}}) = \{0\}$，四个维数依次为 $1, 1, 3, 0$。 ♦

11.2　子空间的交、并、和与维数公式

11.2.1　交与并

$\mathbf{S}\cap\mathbf{U}$ 是既对称又上三角的矩阵。上三角要求对角线下方为 0，对称又要求上方等于下方的镜像，所以上方也只能是 0，剩下的只有对角矩阵：

$$\mathbf{S}\cap\mathbf{U}=\mathbf{D},\qquad \dim(\mathbf{S}\cap\mathbf{U})=3.$$

两个子空间的交总是子空间（第 6 讲）。并就不同了。取

$$X=E_{12}=\begin{bmatrix}0&1&0\\0&0&0\\0&0&0\end{bmatrix}\in\mathbf{U},\qquad Y=E_{12}+E_{21}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&0\end{bmatrix}\in\mathbf{S},$$

则 $X+Y=2E_{12}+E_{21}$ 既不对称（$(1,2)$ 元是 2，$(2,1)$ 元是 1），也不是上三角（$(2,1)$ 元非零）。所以 $X+Y\notin\mathbf{S}\cup\mathbf{U}$，并集 $\mathbf{S}\cup\mathbf{U}$ 对加法不封闭，**不是子空间**。第 6 讲在 $\mathbb{R}^n$ 里得到过同样的结论：两个子空间的并是子空间，当且仅当其中一个包含另一个。

11.2.2　和

要得到包含 $\mathbf{S}$ 与 $\mathbf{U}$ 的最小子空间，应取**和**（sum）

$$\mathbf{S}+\mathbf{U}=\{\,S+U\mid S\in\mathbf{S},\ U\in\mathbf{U}\,\},$$

即一个对称矩阵加一个上三角矩阵所能得到的全部矩阵。答案是整个 $\mathbf{M}$。

■ **例 11.3　把矩阵拆成对称加上三角**　把 $A=\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}$ 写成一个对称矩阵与一个上三角矩阵之和，并说明这样的拆分有多少种。

解　上三角矩阵对角线下方是 0，所以 A 下方的元素 $4,7,8$ 只能由对称部分提供。对称部分在下方放 $4,7,8$，在上方镜像位置也必须放 $4,7,8$，对角线暂时放 0：

$$S_0=\begin{bmatrix}0&4&7\\4&0&8\\7&8&0\end{bmatrix},\qquad U_0=A-S_0=\begin{bmatrix}1&-2&-4\\0&5&-2\\0&0&9\end{bmatrix}.$$

S_0 对称，U_0 上三角，$S_0+U_0=A$。这个构造对任意 A 都能做：把 A 的严格下三角部分记作 L，取 $S_0=L+L^{\mathsf T}$，$U_0=A-S_0$。所以 $\mathbf{S}+\mathbf{U}=\mathbf{M}$。

拆分不唯一。对任意对角矩阵 D'，S_0+D' 仍对称，U_0-D' 仍上三角，两者之和仍是 A。反过来，若 $A=S_1+U_1=S_0+U_0$，则 $S_1-S_0=U_0-U_1$ 既对称又上三角，属于 $\mathbf{S}\cap\mathbf{U}=\mathbf{D}$。所以全部拆分恰好是 (S_0+D',U_0-D')，D' 跑遍 3 维的 $\mathbf{D}$。交 $\mathbf{S}\cap\mathbf{U}$ 越大，拆分的自由度就越大。 ♦

11.2.3 维数公式

现在把四个维数放在一起：

$$\dim \mathbf{S} = 6, \qquad \dim \mathbf{U} = 6, \qquad \dim(\mathbf{S}+\mathbf{U}) = 9, \qquad \dim(\mathbf{S}\cap\mathbf{U}) = 3.$$

有 $6+6=9+3$。$\mathbf{S}$ 和 $\mathbf{U}$ 的基加起来有 12 个矩阵，可 $\mathbf{S}+\mathbf{U}$ 只有 9 维，多出的 3 正是两者共有的 $\mathbf{D}$ 被数了两次。这对任意两个子空间成立。

定理 11.2（维数公式）　设 V, W 是有限维向量空间的两个子空间，则

$$\dim V + \dim W = \dim(V+W) + \dim(V\cap W).$$

思路　从「重复计数」的那部分 $V\cap W$ 出发。先取它的一组基，再分别扩充成 V 的基和 W 的基。把三组向量合在一起，证明它们恰好是 $V+W$ 的一组基，然后数个数。

证明　设 $\dim(V\cap W)=k$，取 $V\cap W$ 的一组基 $z_1,\dots,z_k$。它是 V 中的线性无关组，可以扩充成 V 的基 $z_1,\dots,z_k,v_1,\dots,v_p$，于是 $\dim V = k+p$；同理扩充成 W 的基 $z_1,\dots,z_k,w_1,\dots,w_q$，$\dim W = k+q$。下面证明 $k+p+q$ 个向量

$$z_1,\dots,z_k,\ v_1,\dots,v_p,\ w_1,\dots,w_q$$

是 $V+W$ 的一组基。

张成：$V+W$ 中的元素形如 $v+w$，$v\in V$ 是 z 和 v_i 的组合，$w\in W$ 是 z 和 w_j 的组合，相加后是这 $k+p+q$ 个向量的组合。

无关：设 $\sum a_i z_i + \sum b_i v_i + \sum c_j w_j = 0$。把含 w_j 的项移到右边：

$$x := \sum a_i z_i + \sum b_i v_i = -\sum c_j w_j.$$

左边在 V 中，右边在 W 中，所以 $x\in V\cap W$，可写成 $x=\sum d_i z_i$。于是 $\sum d_i z_i + \sum c_j w_j = 0$，而 $z_1,\dots,z_k,w_1,\dots,w_q$ 是 W 的基，所以所有 $c_j=0$，从而 $x=0$。再由 z, v 是 V 的基得所有 $a_i=b_i=0$。

所以 $\dim(V+W) = k+p+q = (k+p)+(k+q)-k = \dim V + \dim W - \dim(V\cap W)$。 ■

易错提醒　维数公式里出现的是和 $V+W$，不是并 $V\cup W$。并集一般不是子空间，谈不上维数。Session 1.11 的课时摘要把这个式子印成了 $\dim\mathbf{S}+\dim\mathbf{U}=\dim(\mathbf{S}\cup\mathbf{U})+\dim(\mathbf{S}\cap\mathbf{U})$，这是笔误，正确写法是 $\dim(\mathbf{S}+\mathbf{U})$。

图 11.2 同时展示包含关系和两组空间的交、和；连线方向是从子空间指向包含它的空间。

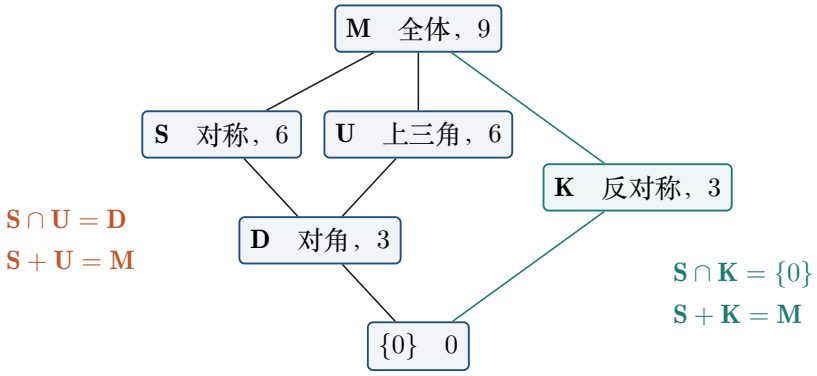


图 11.2　M 的几个子空间及其维数。连线表示下方的空间包含于上方的空间。S 与 U 的交 D 位于两者下方，和 M 位于两者上方。

■ 例 11.4　对称部分与反对称部分　记 $\mathbf{K}=\{\,A\in\mathbf{M}\mid A^{\mathsf T}=-A\,\}$ 为反对称矩阵空间。求 $\dim\mathbf{K}$，用维数公式求 $\dim(\mathbf{S}+\mathbf{K})$，并把例 11.3 中的 A 写成对称矩阵与反对称矩阵之和。

解　$A^{\mathsf T}=-A$ 在对角线上给出 $a_{ii}=-a_{ii}$，所以对角元全为 0；下方元素等于上方元素的相反数。自由元素只有上方 3 个，基为 $E_{12}-E_{21}$，$E_{13}-E_{31}$，$E_{23}-E_{32}$，$\dim\mathbf{K}=3$。

若 A 既对称又反对称，则 $A=A^{\mathsf T}=-A$，$A=0$。所以 $\mathbf{S}\cap\mathbf{K}=\{0\}$。由维数公式

$$\dim(\mathbf{S}+\mathbf{K})=6+3-0=9,$$

$\mathbf{S}+\mathbf{K}$ 是 M 中的 9 维子空间，只能是 M 本身。

具体的拆分是 $A=\frac12(A+A^{\mathsf T})+\frac12(A-A^{\mathsf T})$，前一项对称，后一项反对称：

$$\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}=\begin{bmatrix}1&3&5\\3&5&7\\5&7&9\end{bmatrix}+\begin{bmatrix}0&-1&-2\\1&0&-1\\2&1&0\end{bmatrix}.$$

这次拆分是唯一的：两种拆分之差同时属于 S 和 K，只能是零矩阵。与例 11.3 对比，交为 $\{0\}$ 时没有任何自由度。 ♦

想一想　M 中能否找到两个 6 维子空间 V,W，使它们只在零矩阵处相交？

答　不能。由维数公式，$\dim(V\cap W)=6+6-\dim(V+W)\ge 12-9=3$，因为 $V+W$ 是 M 的子空间，维数不超过 9。S 与 U 的交恰好是 3 维，已经是最小可能。$\mathbb{R}^3$ 中两个过原点的平面至少交于一条直线，是同一个道理：$2+2-3=1$。

11.3　函数空间：微分方程的解

注 11.2　[阅读准备] 本节用到一元导数、三角函数求导和「导数恒为零则函数为常数」。尚未学过微积分时，可先略过本节，接着读下一节的秩 1 矩阵；前面建立的基、维数和子空间结论不依赖这个函数空间例子。

向量空间的元素也可以是函数。两个函数相加、函数乘以常数仍是函数，零函数（处处取 0

的函数）充当零向量。Strang 在课上举的例子是微分方程

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+y=0.$$

直接求导得 $(\cos x)''+\cos x=-\cos x+\cos x=0$，$(\sin x)''+\sin x=-\sin x+\sin x=0$。若允许复值函数，$e^{ix}$ 也满足方程，因为 $(e^{ix})''+e^{ix}=-e^{ix}+e^{ix}=0$。更重要的是，解的线性组合仍是解：若 $y_1''+y_1=0$，$y_2''+y_2=0$，则

$$(c_1y_1+c_2y_2)''+(c_1y_1+c_2y_2)=c_1(y_1''+y_1)+c_2(y_2''+y_2)=0.$$

所以全体解构成一个向量空间。用第 6 讲的语言说，它是线性算子 $L(y)=y''+y$ 的「零空间」：满足 $L(y)=0$ 的全体 y。这里 L 作用在函数上，扮演矩阵 A 的角色，而 L 对线性组合可以逐项作用，就是 $A(cx+dz)=cAx+dAz$ 的翻版。

> **定理 11.3（$y''+y=0$ 的解空间）**　在定义于 $\mathbb{R}$ 上的二阶可导实函数中，方程 $y''+y=0$ 的全体解恰好是
>
> $$y=c_1\cos x+c_2\sin x,\qquad c_1,c_2\in\mathbb{R}.$$
>
> $\cos x,\sin x$ 是解空间的一组基，解空间的维数为 2。

证明　线性无关：若 $c_1\cos x+c_2\sin x$ 是零函数，它对每个 x 都为 0。取 $x=0$ 得 $c_1=0$，取 $x=\pi/2$ 得 $c_2=0$。

张成：设 y 是任意一个解，令 $z=y-y(0)\cos x-y'(0)\sin x$。$z$ 是解的组合，所以也是解，并且 $z(0)=0$，$z'(0)=0$。考虑 $E(x)=z^2+(z')^2$，则

$$E'(x)=2zz'+2z'z''=2z'(z+z'')=0,$$

所以 E 是常数，$E(x)=E(0)=0$。平方和为 0 迫使 $z\equiv0$，即 $y=y(0)\cos x+y'(0)\sin x$。 ■

若允许复值解，分别对实部和虚部应用上面的结论，仍有 $y=c_1\cos x+c_2\sin x$，只是 $c_1,c_2\in\mathbb{C}$。

证明的第二步顺便给出了坐标：解 y 在基 $\cos x,\sin x$ 下的坐标就是它的初值 $(y(0),y'(0))$。e^{ix} 呢？允许复数系数时，欧拉公式给出

$$e^{ix}=\cos x+i\sin x,$$

它的坐标是 $(1,i)$。复数解空间中 e^{ix},e^{-ix} 也是一组基，两组基之间的换算是 $\cos x=\frac12(e^{ix}+e^{-ix})$，$\sin x=\frac{1}{2i}(e^{ix}-e^{-ix})$。三个解 $\cos x,\sin x,e^{ix}$ 在复数意义下线性相关，因为空间只有 2 维。

■ **例 11.5　解空间中的坐标**　(1) 求 $y''+y=0$ 满足 $y(0)=1,y'(0)=\sqrt3$ 的解，并写成 $R\cos(x-\varphi)$ 的形式。(2) 求 $\cos(x+\pi/3)$ 在基 $\cos x,\sin x$ 下的坐标。(3) $\cos x$ 与 $\cos(x+\pi/3)$ 也是解空间的一组基吗？若是，求 $\sin x$ 在这组基下的坐标。

解　(1) 坐标就是初值，$y=\cos x+\sqrt3\sin x$。提出 2：

$$y=2\Big(\tfrac12\cos x+\tfrac{\sqrt3}{2}\sin x\Big)=2\Big(\cos x\cos\tfrac{\pi}{3}+\sin x\sin\tfrac{\pi}{3}\Big)=2\cos\Big(x-\frac{\pi}{3}\Big).$$

(2) 展开 $\cos(x+\pi/3)=\cos x\cos\frac{\pi}{3}-\sin x\sin\frac{\pi}{3}=\frac{1}{2}\cos x-\frac{\sqrt{3}}{2}\sin x$，坐标为 $\left(\frac{1}{2},-\frac{\sqrt{3}}{2}\right)$。$\cos(x+\pi/3)$ 本身是解，这与「平移一个解仍是解」相符。

(3) 用坐标把问题搬到 $\mathbb{R}^2$：$\cos x\leftrightarrow(1,0)$，$\cos(x+\pi/3)\leftrightarrow\left(\frac{1}{2},-\frac{\sqrt{3}}{2}\right)$。这两个向量不共线，所以是 $\mathbb{R}^2$ 的基，对应的两个函数也是解空间的基。设 $\sin x=a\cos x+b\cos(x+\pi/3)$，比较坐标：

$$a+\tfrac{1}{2}b=0,\qquad -\tfrac{\sqrt{3}}{2}b=1\quad\Longrightarrow\quad b=-\tfrac{2}{\sqrt{3}},\quad a=\tfrac{1}{\sqrt{3}}.$$

验算：$\frac{1}{\sqrt{3}}\cos x-\frac{2}{\sqrt{3}}\left(\frac{1}{2}\cos x-\frac{\sqrt{3}}{2}\sin x\right)=\sin x$。 ♦

例 11.5 的第 (3) 问显示了坐标的用处：一旦选定一组基，2 维解空间里的问题就变成 $\mathbb{R}^2$ 里的问题。表 11.1 列出几个常系数线性齐次方程的解空间；这些例子中，解空间的维数等于方程的阶数。

表 11.1　几个线性齐次常微分方程的解空间

方程	一组基	另一组基	维数
$y'-y=0$	e^{x}	$2e^{x}$	1
$y''=0$	$1,\ x$	$1+x,\ 1-x$	2
$y''-y=0$	$e^{x},\ e^{-x}$	$\cosh x,\ \sinh x$	2
$y''-3y'+2y=0$	$e^{x},\ e^{2x}$	$e^{x}+e^{2x},\ e^{2x}$	2
$y''+y=0$	$\cos x,\ \sin x$	$e^{ix},\ e^{-ix}$（复系数）	2
$y'''=0$	$1,\ x,\ x^{2}$	$1,\ 1+x,\ (1+x)^{2}$	3

易错提醒　$y''+y=1$ 的解集不是向量空间。零函数不满足它，两个解之和满足的是 $y''+y=2$。它的全部解是 $y=1+c_1\cos x+c_2\sin x$：一个特解 $y_p=1$ 加上齐次方程的全部解。这和第 6 讲中 $Ax=b$ 的解集 $x_p+N(A)$ 结构相同。

前后联系　18.03（常微分方程）中的「叠加原理」说的就是齐次方程的解对线性组合封闭。在系数连续且最高阶导数项系数处处非零的区间上，n 阶线性齐次方程的解空间是 n 维的，这是 18.03 的基本定理之一。第 22 讲会研究 $\mathrm{d}u/\mathrm{d}t=Au$；当 A 可对角化时，特征向量给出 $e^{\lambda t}v$ 形式的一组解，矩阵指数则处理一般方阵。

11.4　秩 1 矩阵

11.4.1　列乘以行

看矩阵

$$A=\begin{bmatrix}1&4&5\\2&8&10\end{bmatrix}.$$

第二行是第一行的 2 倍，消元一步就得到零行，所以 $\operatorname{rank}A=1$。行空间由 $(1,4,5)$ 张成，列空间由第一列 $(1,2)$ 张成，另两列 $(4,8)$、$(5,10)$ 分别是它的 4 倍和 5 倍。A 的每一列都是 $(1,2)$ 的

倍数，倍数依次是 $1,4,5$，所以

$$A=\begin{bmatrix}1\\2\end{bmatrix}\begin{bmatrix}1&4&5\end{bmatrix}.$$

一个列向量乘以一个行向量。这个形状对所有秩 1 矩阵都成立。

定理 11.4（秩 1 矩阵的结构）　$m\times n$ 矩阵 A 的秩为 1，当且仅当存在非零向量 $u\in\mathbb{R}^m$，$v\in\mathbb{R}^n$ 使得 $A=uv^{\mathsf T}$。此时

$$\begin{aligned}&C(A)=\operatorname{span}\{u\},&&C(A^{\mathsf T})=\operatorname{span}\{v\},\\&N(A)=\{\,x\mid v^{\mathsf T}x=0\,\},&&N(A^{\mathsf T})=\{\,y\mid u^{\mathsf T}y=0\,\}.\end{aligned}$$

证明　设 $A=uv^{\mathsf T}$，$u,v\neq 0$。A 的第 j 列是 v_ju，所以每一列都在 $\operatorname{span}\{u\}$ 中；$v\neq 0$ 时至少有一个 $v_j\neq 0$，这一列非零，所以 $C(A)=\operatorname{span}\{u\}$，$\operatorname{rank}A=1$。同理 A 的第 i 行是 $u_iv^{\mathsf T}$，行空间是 $\operatorname{span}\{v\}$。由 $Ax=u(v^{\mathsf T}x)$ 且 $u\neq 0$，$Ax=0$ 当且仅当数 $v^{\mathsf T}x=0$；对 $A^{\mathsf T}=vu^{\mathsf T}$ 同理。

反过来，设 $\operatorname{rank}A=1$。列空间是一条直线，取其中一个非零向量 u，则每一列都是 u 的倍数：第 j 列 $=v_ju$。把这些倍数排成 $v=(v_1,\dots,v_n)$，就有 $A=uv^{\mathsf T}$。$A\neq 0$，所以 $v\neq 0$。 ■

定理中的四个子空间维数是 $1,1,n-1,m-1$，与第 10 讲的 $r,\ r,\ n-r,\ m-r$（$r=1$）一致。对上面的 A（$m=2$，$n=3$），$N(A)$ 是 $\mathbb{R}^3$ 中的平面 $x_1+4x_2+5x_3=0$，$N(A^{\mathsf T})$ 是 $\mathbb{R}^2$ 中的直线 $y_1+2y_2=0$。

图 11.3 用 $u=(1,2,\dots,6)^T$ 和 $v=(1,2,\dots,6)^T$ 展开列乘行；每个格子都是一个外积元素 u_iv_j。

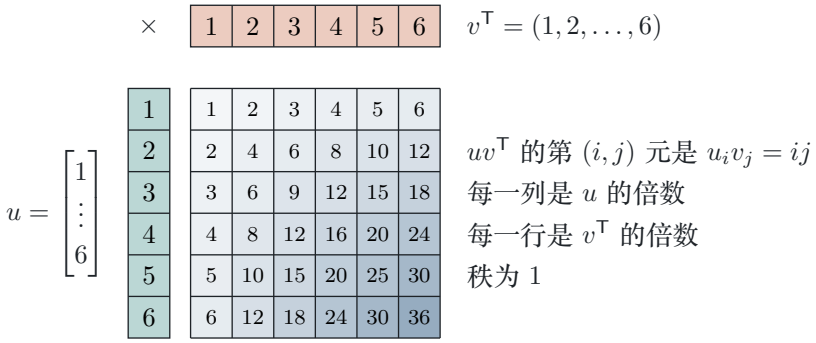


图 11.3　乘法表是秩 1 矩阵。左边一列 u 乘以上面一行 $v^{\mathsf T}$，得到 6×6 的乘法表 $uv^{\mathsf T}$。颜色越深表示数越大。

定理 11.4 给出子空间的结构；对方阵，外积还带来一个简单的平方公式。

■ **例 11.6　平方与四个子空间**　设

$$A=\begin{bmatrix}2&4\\1&2\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}\begin{bmatrix}1&2\end{bmatrix}.$$

求 A^2、$C(A)$、$C(A^{\mathsf T})$、$N(A)$ 和 $N(A^{\mathsf T})$。

解　先用外积相乘：

$$A^2 = u(v^{\mathsf T}u)v^{\mathsf T} = \big(1\cdot 2 + 2\cdot 1\big)A = 4A = \begin{bmatrix} 8 & 16 \\ 4 & 8 \end{bmatrix}.$$

A 的列都是 $(2,1)$ 的倍数，$A^{\mathsf T}$ 的列都是 $(1,2)$ 的倍数，所以

$$C(A) = \operatorname{span}\{(2,1)\}, \qquad C(A^{\mathsf T}) = \operatorname{span}\{(1,2)\}.$$

由 $Ax = 0 \iff x_1 + 2x_2 = 0$，得 $N(A) = \operatorname{span}\{(-2,1)\}$；由 $A^{\mathsf T}y = 0 \iff 2y_1 + y_2 = 0$，得 $N(A^{\mathsf T}) = \operatorname{span}\{(1,-2)\}$。四个子空间的维数依次为 $1,1,1,1$，与第 10 讲的秩—零度计数相符。 ♦

命题 11.5（秩 1 方阵的平方）　若 $A = uv^{\mathsf T}$，其中 $u, v \in \mathbb{R}^n$ 非零，则

$$A^2 = (v^{\mathsf T}u)A.$$

证明　直接结合矩阵乘法的结合律：$A^2 = (uv^{\mathsf T})(uv^{\mathsf T}) = u(v^{\mathsf T}u)v^{\mathsf T} = (v^{\mathsf T}u)A$。中间的 $v^{\mathsf T}u$ 是一个数，因此可以移到矩阵外。 ■

若 $v^{\mathsf T}u = 0$，非零的秩 1 方阵也可能满足 $A^2 = 0$。例如 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ 对应 $u = (1,0)$、$v = (0,1)$。因此「秩 1」并不意味着 A 可逆或 $A^2 \neq 0$。

定理 11.6（秩 r 矩阵是秩 1 矩阵之和）　若 $\operatorname{rank} A = r$，则 A 可以写成 r 个秩 1 矩阵之和；反过来，r 个秩 1 矩阵之和的秩至多为 r。

证明　取 $C(A)$ 的一组基，将基向量排成 $C = [c_1 \ \cdots \ c_r]$。每列 a_j 都是这些基向量的线性组合，把系数逐列排成 R，得到 $A = CR$。按列乘行展开：

$$A = \sum_{k=1}^{r} c_k R_{k,:},$$

其中 $R_{k,:}$ 是 R 的第 k 行。C 的列独立，所以映射 $x \mapsto Cx$ 不会把非零向量送到零向量；由 $A = CR$ 得 $\operatorname{rank} R = \operatorname{rank} A = r$，故 R 的每一行都非零。于是每一项 $c_k R_{k,:}$ 都是秩 1 矩阵。反方向中，每一项的列空间至多一维，r 项相加后的列空间包含于这 r 条直线张成的空间，维数不超过 r。 ■

例 11.7　把秩 2 矩阵拆成两个秩 1 矩阵　将

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 5 \end{bmatrix}$$

写成两个秩 1 矩阵之和。

解　前两列 $c_1=(1,1)$、$c_2=(2,3)$ 线性无关，第三列 $(4,5)=2c_1+c_2$。按列组合的系数写出 R：

$$A=\begin{bmatrix}1&2&4\\1&3&5\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}1&0&2\end{bmatrix}+\begin{bmatrix}2\\3\end{bmatrix}\begin{bmatrix}0&1&1\end{bmatrix}.$$

两个外积都非零，秩各为 1；A 的前两列独立，故 $\operatorname{rank}A=2$。这次分解恰好用了两项，和定理 11.6 的下界一致。 ♦

11.5　图距离与小世界

如果五个人之间有若干好友关系，最少经过几次「朋友的朋友」能把任意两个人连起来？把人看成节点，把好友关系看成边，这个问题就变成图上的距离问题。

定义 11.7（无向图）　一个无向图由节点集合 V 和边集合 E 组成，记作 $G=(V,E)$。每条边连接两个不同的节点，边没有方向；本节的图不含重边。

例如，若节点 1 与节点 2 表示两个人，边 $\{1,2\}$ 表示他们互为好友；没有这条边只表示两人未直接连接，不排除有共同好友。网站也可作节点、超链接作边；若保留链接方向，得到的是有向图，本节先讨论无向图。

固定节点次序 $1,\dots,n$，定义 $n\times n$ 矩阵 A：有边连接 i,j 时 $A_{ij}=1$，否则 $A_{ij}=0$。这个矩阵称为**邻接矩阵**（adjacency matrix）。无向图的邻接矩阵对称，对角线为零。

命题 11.8（邻接矩阵幂与游走数）　对任意 $k\ge 1$，$(A^k)_{ij}$ 等于从节点 i 到节点 j 的长度为 k 的**游走**（walk）数；游走允许重复经过节点和边。

证明　$k=1$ 时，A_{ij} 正好记录是否有一条边。若 A^k 统计长度 k 的游走，则

$$(A^{k+1})_{ij}=\sum_{\ell=1}^{n}(A^k)_{i\ell}A_{\ell j}$$

把最后一步落在 ℓ 的所有长度 k 游走逐一接上一条 ℓ 到 j 的边，恰好统计全部长度 $k+1$ 的游走。归纳即得结论。 ■

令 $A^0=I$。若 i 与 j 连通，它们的**图距离**（graph distance）定义为

$$d(i,j)=\min\{\,k\in\mathbb{Z}_{\ge 0}\mid (A^k)_{ij}>0\,\}.$$

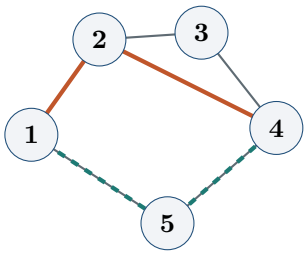
若两点不连通，距离记为 ∞。因此，A^k 某一元非零只说明存在一条长度恰为 k 的游走；要找最短距离，必须从 $k=0,1,2,\dots$ 依次找第一个非零的位置。

例 11.8　五人网络中的距离　五个人之间的好友边为

$$E = \{\{1,2\},\{1,5\},\{2,3\},\\ \{2,4\},\{3,4\},\{4,5\}\}.$$

写出邻接矩阵，计算 A^2，并求节点 1 到 4 的距离。

图 11.4 中的橙色和青色各标出一条从 1 到 4 的两步路线。接下来用矩阵核实它们，并检查图中最远的节点对。



橙色：$1 \to 2 \to 4$
青色：$1 \to 5 \to 4$

图 11.4　该网络有 5 个节点和 6 条边；节点 1 到 4 有两条长度为 2 的路线。

解　按节点 $1,2,3,4,5$ 排列行与列，邻接矩阵为

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

它对称且对角线为零；各行和 $2,3,2,3,2$ 是对应节点的度数。相乘得到

$$A^2 = \begin{bmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & 3 & 1 & 1 & 2 \\ 1 & 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 3 & 0 \\ 0 & 2 & 1 & 0 & 2 \end{bmatrix}.$$

$A_{14} = 0$，但 $(A^2)_{14} = 2$，所以 $d(1,4) = 2$；两条游走正是 $1 \to 2 \to 4$ 和 $1 \to 5 \to 4$。逐对取第一次出现的正数，可得

$$\big(d(i,j)\big)_{i,j=1}^{5} = \begin{bmatrix} 0 & 1 & 2 & 2 & 1 \\ 1 & 0 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 \end{bmatrix}.$$

任意两人最多隔两条边，故这张图的直径为 2。现实社交网络节点更多，「六度分隔」指的正是远距离节点也可能通过很短的好友链连接；它描述的是网络现象，不是对每张图都成立的定理。 ♦

想一想　简单无向图的 $(A^2)_{ii}$ 表示什么？

答　每一条从 i 出发再回到 i 的两步游走，先任选一个与 i 相邻的节点再沿原边返回。因此 $(A^2)_{ii}$ 等于节点 i 的度数；本例对角线 $2,3,2,3,2$ 与邻接矩阵的行和一致。

易错提醒　$(A^k)_{ij}>0$ 不一定表示距离恰好是 k。例如本例 $(A^2)_{11}=2$，但 $d(1,1)=0$。它只表示存在长度 k 的游走；图距离要取使矩阵元首次非零的最小 k。

本讲小结

- $\mathbf{M}_{m\times n}$ 的标准基是 E_{ij}，维数为 mn；对称、上三角、对角矩阵空间的维数依次为 $\frac{n(n+1)}{2},\frac{n(n+1)}{2},n$。
- 子空间的交仍是子空间；维数公式为 $\dim(V+W)=\dim V+\dim W-\dim(V\cap W)$，并 $V\cup W$ 通常不是子空间。
- $y''+y=0$ 的实解空间以 $\cos x,\sin x$ 为基，解 y 的坐标是 $(y(0),y'(0))$。
- $\operatorname{rank}A=1\iff A=uv^{\mathsf T}$，其中 u,v 非零；四个基本子空间的维数是 $1,1,n-1,m-1$。
- 秩 1 方阵满足 $A^2=(v^{\mathsf T}u)A$；秩为 r 的矩阵可分解为 r 个秩 1 矩阵之和。
- 邻接矩阵幂的元素 $(A^k)_{ij}$ 统计长度 k 的游走数；图距离是使它首次为正的最小 k。
- 「最小步数」不能由一个较高次幂的非零元素直接读出；要逐阶检查 $A,A^2,\dots$。

练习

11.1 ★　判断真假，假的给出反例：(a) 全体秩恰为 1 的 2×2 矩阵构成子空间；(b) 若 $A^2=0$，则 $A=0$。

11.2 ★　令 $V=\{\,A\in\mathbf{M}_{3\times3}\mid \operatorname{tr}A=0\,\}$。证明 V 是子空间，并给出一组基和维数。

11.3 ★　已知 $\dim V=7$，$\dim W=5$，$\dim(V\cap W)=3$，求 $\dim(V+W)$。

11.4 ★　求 $y''+y=0$ 满足 $y(0)=0$、$y'(0)=2$ 的解。

11.5 ★★　对 $A=\begin{bmatrix}2&4\\1&2\end{bmatrix}$，写成 $uv^{\mathsf T}$，求 A^2 与 $N(A)$、$N(A^{\mathsf T})$。

11.6 ★★　在例 11.8 的图中，求 $d(2,5)$，并列出长度为 2 的所有从 2 到 5 的游走。

11.7 ★★★　令 P_{21},P_{31},P_{32} 分别交换 3×3 单位矩阵的第 $2,1$、第 $3,1$、第 $3,2$ 行。证明 $I=P_{21}+P_{31}+P_{32}-P_{32}P_{21}-P_{21}P_{32}$，并证明右边的五个非恒等置换矩阵线性无关。

11.8 ★★★　设 $L=\begin{bmatrix}1&0&-1\\-1&1&0\\0&-1&1\end{bmatrix}$。描述所有满足 $LX=0$ 的 3×3 矩阵 X，再描述所有形如 LX 的矩阵及这两个空间的维数。

练习答案与提示

11.1 (a) 假。E_{11} 与 $-E_{11}$ 都秩为 1，但它们的和是秩 0 的零矩阵。(b) 假：$N=\begin{bmatrix}0&1\\0&0\end{bmatrix}\neq 0$，而 $N^2=0$。

11.2 由 $\operatorname{tr}(cA+dB)=c\operatorname{tr}A+d\operatorname{tr}B$，$V$ 对线性组合封闭。可取基

$$E_{12},E_{13},E_{21},E_{23},E_{31},E_{32},E_{11}-E_{33},E_{22}-E_{33}.$$

前六个系数分别指定非对角元。最后两个系数指定前两个对角元；迹为零确定第三个对角元。因此 $\dim V=8$。

11.3 维数公式给出 $\dim(V+W)=7+5-3=9$。

11.4 解的一般形式是 $y=c_1\cos x+c_2\sin x$。初值给出 $c_1=0$、$c_2=2$，所以 $y=2\sin x$。

11.5 $u=(2,1)^T$、$v=(1,2)^T$，故 $A=uv^{\mathsf T}$ 且 $v^{\mathsf T}u=4$。于是 $A^2=4A=\begin{bmatrix}8&16\\4&8\end{bmatrix}$。由 $x_1+2x_2=0$ 得 $N(A)=\operatorname{span}\{(-2,1)\}$；由 $2y_1+y_2=0$ 得 $N(A^{\mathsf T})=\operatorname{span}\{(1,-2)\}$。

11.6 邻接矩阵中 $A_{25}=0$，而 $(A^2)_{25}=2$，故最短距离为 2。两条长度为 2 的游走是 $2\to1\to5$ 和 $2\to4\to5$。

11.7 三个换行矩阵之和是全 1 矩阵 J；$P_{32}P_{21}+P_{21}P_{32}=J-I$，相减即得题中等式。若五矩阵的线性组合为零，其元素矩阵为

$$\begin{bmatrix}c_3&c_1+c_4&c_2+c_5\\c_1+c_5&c_2&c_3+c_4\\c_2+c_4&c_3+c_5&c_1\end{bmatrix}=0.$$

对角元先给出 $c_1=c_2=c_3=0$，再由位置 $(1,2)$、$(1,3)$ 得 $c_4=c_5=0$，所以它们线性无关。等行和、等列和的矩阵可由 a,b,d,e,s 五个自由数写成

$$\begin{bmatrix}a&b&s-a-b\\d&e&s-d-e\\s-a-d&s-b-e&a+b+d+e-s\end{bmatrix},$$

因而该子空间维数是 5，五个独立的置换矩阵构成其基。

11.8 L 的第三行为前两行之和的相反数，前两行独立，故 $\operatorname{rank}L=2$；解方程得 $N(L)=\operatorname{span}\{(1,1,1)\}$。因此 $LX=0$ 当且仅当 X 的每一列都是 $(1,1,1)$ 的倍数，即

$$X=\begin{bmatrix}a&b&c\\a&b&c\\a&b&c\end{bmatrix},$$

所以零空间维数为 3。$C(L)=\{\,(a,d,-a-d)\mid a,d\in\mathbb{R}\,\}$：给定此列向量，取 $x=(a,a+d,0)$ 就有 $Lx=(a,d,-a-d)$。因此像空间中的矩阵恰为

$$B=\begin{bmatrix}a&b&c\\d&e&f\\-a-d&-b-e&-c-f\end{bmatrix},\qquad a,b,c,d,e,f\in\mathbb{R}.$$

每列有两个自由数，故像空间维数为 6。对上述 B 取

$$X=\begin{bmatrix}a&b&c\\a+d&b+e&c+f\\0&0&0\end{bmatrix},$$

直接相乘得 $LX = B$，所以列和为零的条件也充分。左乘算子 $T(X) = LX$ 的零度为 3、秩为 6；秩—零度定理给出 $3 + 6 = 9 = \dim \mathbf{M}_{3\times 3}$。

第 12 讲

图、网络与关联矩阵

Graphs, Networks, Incidence Matrices

对应材料：MIT 18.06SC Session 1.12 课时摘要；Problems 12.1–12.2 及 Solutions；Strang《Introduction to Linear Algebra》第 4 版第 8.2 节 / 第 5 版第 10.1 节；ZoomNotes 第 3.6 节。

本讲目标

(1) 会从带方向的图写出关联矩阵，并解释矩阵乘以节点电势所得的边电势差。

(2) 理解连通图的关联矩阵秩为节点数减一，并描述它的零空间与四个基本子空间。

(3) 会用 $A^{\mathsf{T}}y=0$ 写出 Kirchhoff 电流定律，并用独立循环流构造其解。

(4) 会从 Ohm 定律和节点平衡推出 $A^{\mathsf{T}}CAx=f$，并解带接地点的简单网络。

(5) 能说明边方向只影响关联矩阵与电流坐标的符号，不改变网络本身。

12.1 图、网络与边的方向

第 11 讲把人或网站看成节点，用边表示好友关系或链接，并用邻接矩阵研究距离。电路也能画成同样的节点与边：节点记录电势，边表示连接两个节点的导线和元件。两种应用共享同一种图结构。

定义 12.1（图、回路与树） 有限图（graph）$G=(V,E)$ 由节点集 V 与边集 E 组成，每条边连接两个不同节点。若任意两节点可由边序列连接，称图连通（connected）；沿边出发又回到起点且中途不重复节点的路径称为回路（cycle）；连通且无回路的图称为树（tree）。若树包含 G 的全部节点、边都是 G 的边，称它为 G 的生成树（spanning tree）。

例如，节点 1 到 2、2 到 3、1 到 3 三条边组成一个回路；删去 1 到 3 后仍连通而无回路，所得三节点图是一棵树。只有一个节点且没有边的图也连通，是一棵树；两个孤立节点组成的图不连通。连通图总能找到生成树：若图中还有回路，就删去回路上的一条边；删边后回路的其余部分仍连接这条边的两个端点，所以图仍连通。重复删边直到没有回路，剩下的就是生成树。

树的边数由节点数决定。从一个节点开始逐步扩展树，每加入一条边，都必须接入一个尚未出现的新节点，否则新边会和已有路径构成回路。因此一棵含 n 个节点的树有 $n-1$ 条边。反过来，连通图有 m 条边时，生成树用掉 $n-1$ 条；每条剩余边加回去，都会与树中连接它两个端点的唯一路径构成一个回路。这些树外边将成为循环流自由度的来源。

电流要有正负号，所以还要给每条边选一个方向（orientation）。方向只是记号：边上的电流

若沿箭头流，坐标取正；若逆箭头流，坐标取负。图 12.1 展示一个有 4 个节点、5 条边的连通网络。

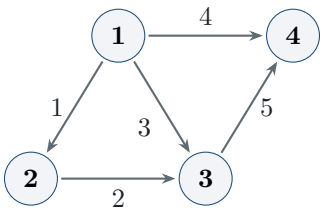


图 12.1　箭头给出边的正方向；该连通图有 4 个节点、5 条边和 2 个独立回路。

取边 $1,2,3$ 得到回路 $1\to2\to3\to1$，其中最后一步逆着边 3 的箭头。边 $3,4,5$ 也组成回路。两个回路共享边 3，所以后面数「独立回路」时，不能把不同的简单回路个数直接相加。

■ **例 12.1　树与回路**　图 12.1 有多少个独立回路？删去边 3 后，剩下的图是什么？

解　生成树可以取边 $1,2,4$。其余边 $3,5$ 各自闭合一个基本回路，因此这里有 $5-4+1=2$ 个独立回路。还可以直接找出两个三角回路：边 $1,2,3$ 构成一个，边 $3,4,5$ 构成另一个。删去边 3 后，边 $1,2,4,5$ 仍连通，却留下了由边 $1,2,4,5$ 构成的外圈；所以这四条边不是树。要得到树，还需再删一条外圈上的边，例如删边 5，就得到边 $1,2,4$。这题提醒我们：删掉一条回路边只会消除当前回路；若图中还有其他回路，必须继续检查。 ♦

想一想　若把图 12.1 的边 3 反向，节点与回路会改变吗？
答　都不改变。边方向是给电流坐标选的正方向；改箭头会使该边关联矩阵那一行和对应电流坐标同时变号，图的连通关系、回路和电路解保持不变。

12.2　关联矩阵记录端点差

图 12.1 的每条边都连接两个节点。若把节点电势记为 $x_1,\dots,x_4$，边 1 两端的电势差是 x_2-x_1，边 4 两端的电势差是 x_4-x_1。把这类差值一次算完，需要一张把边与端点对应起来的矩阵。

定义 12.2（关联矩阵）　给每条边选定方向后，图的**关联矩阵**（incidence matrix）A 有 m 行、n 列：行按边编号，列按节点编号。若第 k 条边从节点 i 指向节点 j，则第 k 行在第 i 列为 -1、第 j 列为 1，其余位置为 0。

对图 12.1，节点次序和边次序都取图中的编号，得到

$$A=\begin{bmatrix}-1&1&0&0\\0&-1&1&0\\-1&0&1&0\\-1&0&0&1\\0&0&-1&1\end{bmatrix}.$$

例如第一行的 -1 与 1 分别记录边 1 的起点 1 和终点 2。每行恰有一个 -1、一个 1，行和为 0；大图的关联矩阵因此通常很稀疏。

令节点电势列向量为 $x=(x_1,x_2,x_3,x_4)^{\mathsf{T}}$。矩阵乘法逐行给出

$$Ax=\begin{bmatrix} x_2-x_1\\ x_3-x_2\\ x_3-x_1\\ x_4-x_1\\ x_4-x_3 \end{bmatrix}.$$

所以 A 把节点上的数值变成每条边的「终点电势减起点电势」。这一定义固定后，边反向时相应一行整体变号。

从节点走一段路径时，把沿路径方向的边电势升高相加，中间节点的电势会成对抵消，最后只剩「终点电势减起点电势」。因此任何由节点电势产生的边数据，沿任意闭合回路的有向和都必须为 0。这是判断一个边向量能否写成 Ax 的办法：每个回路都要满足电压升高闭合。

■ **例 12.2　从节点电势算边电势差**　在上面的五边图中取 $x=(3,0,1,2)^{\mathsf{T}}$，求 Ax。

解　按边顺序逐个作「终点减起点」：

$$Ax=(0-3,\ 1-0,\ 1-3,\ 2-3,\ 2-1)^{\mathsf{T}}=(-3,1,-2,-1,1)^{\mathsf{T}}.$$

第三个分量为 -2，说明节点 3 的电势比节点 1 低 2。负号表示电势沿所选箭头下降，不表示计算出错。 ♦

■ **例 12.3　边电压差能否来自节点电势**　在五边图上给定边向量 $d=(-3,1,-2,-1,1)^{\mathsf{T}}$。判断它是否能写成 Ax；若能，求一个满足 $x_1=0$ 的节点电势向量。再把第三个分量改成 -1，说明新向量为何不可能来自节点电势。

解　先检查两个基本回路的电压升高和。对回路 $1,2,3$，方向关系给出 $d_1+d_2-d_3=-3+1-(-2)=0$；对回路 $3,4,5$，有 $d_3-d_4+d_5=-2-(-1)+1=0$。两条回路的和都为 0，所以可以从一个节点开始逐边恢复电势。令 $x_1=0$，边 1 给出 $x_2-x_1=-3$，故 $x_2=-3$；边 2 给出 $x_3-x_2=1$，故 $x_3=-2$；边 4 给出 $x_4-x_1=-1$，故 $x_4=-1$。最后边 3 和边 5 分别得到 $x_3-x_1=-2$、$x_4-x_3=1$，与给定数据相符。

若把 d_3 改为 -1，第一条回路的和变成 $-3+1-(-1)=-1$。沿回路走一圈却得到非零电势变化，和起点电势必须等于终点电势矛盾。因此修改后的向量不在 $C(A)$ 中。 ♦

定义也给出边界情形：没有边的图有一个 $0\times n$ 关联矩阵，乘任何节点电势都得到空向量；单条边连接两个节点时，矩阵是 $[-1\quad 1]$。若把该边改向，矩阵变为 $[1\quad -1]$，其秩和零空间不变。

易错提醒　把第 k 行的 -1、1 互换，会把「终点减起点」误写成「起点减终点」。例如边 $1\to 2$ 的行应是 $[-1,1]$；写成 $[1,-1]$ 等于反转了边方向。方向可以任选，但写完后 Ax 与电流的符号约定必须一起跟着改变。

12.3　连通性决定秩与循环数

五边网络的 A 有 4 列，却只有 3 个独立行。边 $1,2,3$ 形成回路，正好有第三行等于前两行之和。这个例子提示：关联矩阵的相关性记录了图里的回路；图是否连通则决定秩。

定理 12.3（连通图的关联矩阵秩与四空间）　设有限无向图连通，含 n 个节点、m 条边，并任意给边定向，关联矩阵为 $A \in \mathbb{R}^{m\times n}$。则

$$\operatorname{rank} A = n-1, \qquad N(A) = \operatorname{span}\{(1,\dots,1)^{\mathsf{T}}\}, \qquad \dim N(A^{\mathsf{T}}) = m-n+1.$$

行空间 $C(A^{\mathsf{T}})$ 是坐标和为 0 的节点向量空间；列空间 $C(A)$ 由所有边电势差 Ax 组成。$N(A^{\mathsf{T}})$ 是满足各节点流量平衡的循环流空间，其维数等于独立回路数 $\mu = m-n+1$。因此

$$n - m + \mu = 1.$$

定理里的四个空间可以用正交关系连在一起。节点向量空间中的 $N(A)$ 与行空间 $C(A^{\mathsf{T}})$ 互为正交补；边向量空间中的循环流空间 $N(A^{\mathsf{T}})$ 与电势差空间 $C(A)$ 也互为正交补：

$$C(A^{\mathsf{T}}) = N(A)^{\perp} = \{z \in \mathbb{R}^n : \mathbf{1}_n^{\mathsf{T}} z = 0\}, \qquad C(A) = N(A^{\mathsf{T}})^{\perp}.$$

第二个等式解释了上一节的回路检验。对每个循环流 c 和边电势差 $d = Ax$，有 $c^{\mathsf{T}}d = c^{\mathsf{T}}Ax = (A^{\mathsf{T}}c)^{\mathsf{T}}x = 0$。也就是说，电势差空间与循环流空间正交。由于两空间维数之和是 $(n-1)+(m-n+1) = m$，正交关系已经刻画了全部的 $C(A)$，没有遗漏其他限制。

树还给出一个构造节点电势的方法。任选一个根节点并把它的电势定为 0，再沿生成树的唯一路径累加边数据 d_k；若走向与箭头相反，就减去 d_k。这样树边一定满足 $d_k = x_{\mathrm{head}(k)} - x_{\mathrm{tail}(k)}$。每条树外边和树路径组成一个基本回路，若这个回路的有向电压和为 0，该树外边的数据也会等于两端节点电势差。因而只需检查一组基本回路，全部通过就能恢复出 $d = Ax$。

子空间	维数	在网络中的含义
$N(A) \subseteq \mathbb{R}^n$	1	全体节点同时平移同一电势
$C(A^{\mathsf{T}}) \subseteq \mathbb{R}^n$	$n-1$	节点坐标和为 0 的向量
$C(A) \subseteq \mathbb{R}^m$	$n-1$	来自节点电势的边电势升高
$N(A^{\mathsf{T}}) \subseteq \mathbb{R}^m$	$m-n+1$	满足 KCL 的循环流

表中的前两行属于节点空间，后两行属于边空间。不能仅凭维数把节点向量和边向量混为一谈；它们分别有 n 个和 m 个坐标。

直观理解　沿一条边令两端电势相等，只能推出这两个节点相等。连通性让任意两节点之间都有路径，于是 $Ax = 0$ 会迫使所有节点电势相同；剩下的一个自由度就是整体加上的常数。

不论图是否连通，常数向量总在 $N(A)$ 中，因为每一行的元素和都是 0。连通性只用于证明零空间里没有其他向量；有多个分支时，每个分支都能各自平移一个常数。

思路　先解 $Ax=0$ 找到零空间，再由秩—零度定理得到秩与左零空间维数。最后选一棵生成树：树边给出独立行，树外每条边恰好闭合一个基本回路。

证明　**零空间与秩。**若 $Ax=0$，每条边两端电势相等。任取节点 i,j，连通性给出从 i 到 j 的路径；沿路径逐边相等，得 $x_i=x_j$。反过来，常数向量在每条边上的差都是 0。因此 $N(A)=\operatorname{span}\{\mathbf{1}_n\}$，零度为 1，秩—零度定理给出 $\operatorname{rank}A=n-1$。

行空间与列空间。每行元素之和为 0，故 $C(A^{\mathsf T})$ 包含在节点坐标和为 0 的超平面中。两边维数都是 $n-1$，所以相等。列空间 $C(A)$ 的向量形如 Ax，其第 k 个分量是边 k 的电势差；沿任意定向回路把这些差按行进方向取正负相加，会逐项抵消为 0。列空间维数也等于 $\operatorname{rank}A=n-1$。

循环流与独立回路。由 $\operatorname{rank}A^{\mathsf T}=\operatorname{rank}A=n-1$，有 $\dim N(A^{\mathsf T})=m-n+1$。取一棵生成树：从一个节点开始，每次加入一条连接已选节点与新节点的边，直到覆盖全图；每条新边只带来一个新节点，所以树边数是 $n-1$。树边对应的 $n-1$ 行也独立：若这些行的线性组合为零，取一片树叶，它只接在一条树边上；查看该叶节点对应的列，就迫使这条边的系数为 0。删去树叶后重复此论证，所有系数都为 0。因此树边行构成行空间的一组基。其余 $m-n+1$ 条边各自与树中唯一路径组成一个基本回路。沿基本回路顺向取 1、逆向取 -1、其余边取 0，所得向量都在 $N(A^{\mathsf T})$。

这些回路向量线性无关，因为每个向量在自己的树外边上非零、在其他树外边上为 0。它们也张成整个 $N(A^{\mathsf T})$：给任一平衡流，按树外边的流量减去相应基本回路流，剩下的平衡流只在树边上非零；树有叶节点，叶节点的平衡式迫使唯一邻边流量为 0。依次删去叶节点，所有树边流量都为 0。所以循环流空间的维数恰为树外边数 $m-n+1$，即独立回路数 μ。代入即得 $n-m+\mu=1$。 ■

连通条件不能去掉。若 4 个节点分成两条互不相连的边 1–2、3–4，关联矩阵秩为 2，而 $n-1=3$；零空间由「前两个坐标相等、后两个坐标相等」的向量组成，维数是 2。一般若有 c 个连通分支，$Ax=0$ 要求每个分支内部电势为常数，各分支的常数可以独立选择，因此零度为 c、秩为 $n-c$，左零空间维数为 $m-n+c$。每个分支分别取生成树，树边总数为 $n-c$；树外边数 $m-n+c$ 仍然等于独立回路数。

同一秩公式也能直接从矩阵行看出。对每个分支取一棵生成树，合起来是含 $n-c$ 条边的生成森林；逐个删叶节点可证明这 $n-c$ 行独立。每条剩余边都与森林中一条唯一路径构成回路，它的关联行是路径各行的有向和，所以其余行不会再增加秩。

边界情形也符合这些计数。只有一个节点且没有边时，A 是 0×1 矩阵，秩为 $0=n-1$，节点电势零空间是一维，循环流空间为零维。若 n 个节点全都孤立，则 $c=n$、秩为 0，每个节点的电势都能独立选择。这些情况说明连通分支数而不是节点总数，决定常数零空间有几个自由度。

■ **例 12.4　五边图的树边与行秩**　求图 12.1 的秩、两个循环流的维数，并指出一棵生成树。

解　记关联矩阵的行依次为 $r_1,\dots,r_5$。直接比较端点系数可得 $r_3=r_1+r_2$、$r_5=r_4-r_1-r_2$，所以全部行由 r_1,r_2,r_4 张成。第一式对应边 1,2 的路径与边 3 的直接连接端点相同；第二式则把边 5 的端点差写成树路径 $3\to2\to1\to4$ 的有向和。每条树外边都能沿生成树路径表示，因此树边行确实控制整张图的行空间。这三行独立：若其线性组合为零，节点 2 的系数先迫使 r_1 的系数为 0，节点 3 再迫使 r_2 的系数为 0，节点 4 最后迫使 r_4 的系数为 0。因此 $\operatorname{rank}A=3$，

$\dim N(A^{\mathsf T}) = 5 - 3 = 2$。

边 $1,2,4$ 连接所有四个节点，且没有回路，构成一棵生成树。注意行对应边、列对应节点；树边对应独立行。行化简时的主元列编号指节点列，不能把主元列称为边。 ♦

■ **例 12.5　方形网络的四个基本空间**　按图 12.2 中的边编号写出关联矩阵 A，求 $N(A)$，并说明 $(1,0,0,0)^{\mathsf T}$ 不在行空间中。

解　照定义逐边写行：

$$A = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 1 \end{bmatrix}, \qquad Ax = (x_2 - x_1,\ x_3 - x_2,\ x_3 - x_4,\ x_4 - x_1)^{\mathsf T}.$$

令 $Ax = 0$，四条边分别给出 $x_1 = x_2$、$x_2 = x_3$、$x_3 = x_4$、$x_4 = x_1$，所以 $N(A) = \operatorname{span}\{(1,1,1,1)^{\mathsf T}\}$。每行和为 0，故行空间中每个向量的坐标和为 0；而 $(1,0,0,0)^{\mathsf T}$ 的坐标和为 1，它不在行空间。矩阵秩为 3，此方形只有一个独立回路。 ♦

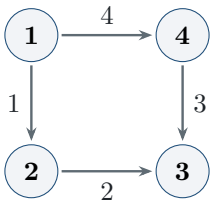


图 12.2　方形网络的方向与原题 12.1 一致；边 $1,2,3,4$ 依次沿左、下、右、上边。

方形例题也是 Problem 12.1 的矩阵计算。节点 1 到 3 的两条路线都给出同一电势差 $x_3 - x_1$；因此四个边电势差不是任意的四元组，它们必须满足绕方形一圈的有向和为 0。

■ **例 12.6　四个基本子空间的维数核对**　对五边图，分别写出四个基本子空间所在的坐标空间和维数，并给出行空间与左零空间的具体向量。

解　关联矩阵有 $m = 5$ 行、$n = 4$ 列、秩 $r = 3$。因此 $C(A^{\mathsf T}) \subseteq \mathbb{R}^4$ 的维数是 3，$N(A) \subseteq \mathbb{R}^4$ 的维数是 $4 - 3 = 1$；$C(A) \subseteq \mathbb{R}^5$ 的维数是 3，$N(A^{\mathsf T}) \subseteq \mathbb{R}^5$ 的维数是 $5 - 3 = 2$。行空间中有 $(-1,1,0,0)^{\mathsf T}$，因为它是 A 的第一行；左零空间中有循环流 $c_1 = (1,1,-1,0,0)^{\mathsf T}$。二者虽在不同的空间，却分别展示两类约束：行空间向量描述节点间的端点差，左零空间向量描述边上的回路流。 ♦

■ **例 12.7　两个连通分支各自计数**　一张图由节点 $1,2,3$ 上的三角形和节点 $4,5$ 之间的一条边组成。求其关联矩阵的秩、两个零空间的维数，并检验分支数与 Euler 计数的关系。

解　三角形三条边依次取方向 $1\to 2$、$2\to 3$、$3\to 1$，末边取方向 $4\to 5$，则

$$A=\begin{bmatrix}-1&1&0&0&0\\0&-1&1&0&0\\1&0&-1&0&0\\0&0&0&-1&1\end{bmatrix}.$$

两个分支各有一个独立常数电势，因此 $\dim N(A)=2$、$\operatorname{rank}A=5-2=3$。三角形提供唯一的独立回路，所以 $\dim N(A^{\mathsf T})=4-3=1$；例如 $(1,1,1,0)^{\mathsf T}$ 是循环流。计数关系为 $n-m+\mu=5-4+1=2=c$。连通时 $c=1$，才化为定理中的 $n-m+\mu=1$。 ♦

12.4　Kirchhoff 电流定律与循环流

沿箭头为正的边电流记作 $y=(y_1,\dots,y_5)^{\mathsf T}$。在节点 2，若 y_1 流入、y_2 流出，电流平衡要求二者相等。把所有节点的平衡式一次写出，正好得到转置矩阵乘法。

定义 12.4（Kirchhoff 电流定律）　在没有外部电流源的节点，流入电流总和等于流出电流总和。若 y_k 按第 k 条边所选方向记号，则所有节点的平衡式可写成

$$A^{\mathsf T}y=0.$$

该方程称为**Kirchhoff 电流定律**（Kirchhoff's Current Law, KCL）。

在图 12.1 中，展开得

$$A^{\mathsf T}y=\begin{bmatrix}-y_1-y_3-y_4\\y_1-y_2\\y_2+y_3-y_5\\y_4+y_5\end{bmatrix}.$$

每一项都是该节点的「流入减流出」。第一项为 0 表示节点 1 处三条边的代数流量平衡；节点 2 的式子则是 $y_1-y_2=0$。因此满足 KCL 的电流向量恰好是左零空间 $N(A^{\mathsf T})$ 中的向量。

把所有节点方程相加，每条边电流都会出现一次 $+y_k$ 和一次 $-y_k$，总和恒为 0。因此连通网络的 n 条节点平衡方程只有 $n-1$ 条独立方程；最后一条由其他节点方程推出。若方程右边包含外部源项，各节点源项也必须总和为 0，否则所有平衡式不可能同时成立。

■ **例 12.8　从 KCL 求出全部循环流**　求图 12.1 上满足 $A^{\mathsf T}y=0$ 的全部电流，并找出一组循环流基。

解　节点 2 给出 $y_2=y_1$，节点 3 给出 $y_5=y_2+y_3$，节点 4 给出 $y_4=-y_5$。节点 1 方程此时自

动成立，因为四条节点方程的总和恒为零。令 $y_1 = t$、$y_3 = s$，于是

$$y = (t, t, s, -t-s, t+s)^{\mathsf{T}} = t\begin{bmatrix}1\\1\\-1\\0\\0\end{bmatrix} + (t+s)\begin{bmatrix}0\\0\\1\\-1\\1\end{bmatrix}.$$

两向量线性无关，故它们构成二维左零空间的一组基。第一向量沿边 $1,2$ 正向、沿边 3 反向，绕节点 $1,2,3$ 流一圈；第二向量沿边 $3,5$ 正向、沿边 4 反向，绕节点 $1,3,4$ 流一圈。代入 $A^{\mathsf{T}}y$ 都得到零向量。 ♦

两个基向量相加得到外圈电流 $(1,1,0,-1,1)^{\mathsf{T}}$。边 3 上一正一负的电流相消，剩下沿外圈循环的流量。若用生成树边 $1,2,4$ 作坐标，树外边 3 和 5 的电流可以分别指定；所有树边电流随后由 KCL 唯一确定。本例参数 (t,s) 与树外边坐标 $(y_3,y_5) = (s,t+s)$ 之间可逆地互相转换，因此它们给出相同的循环流空间。一般每条树外边提供一个自由参数，循环流空间自然有 $m-(n-1) = m-n+1$ 个自由度。这是「多一条边就多一个独立循环流」的矩阵版本。

生成树并不唯一，换一棵树通常会换掉这组基本回路向量；只要树覆盖同一张连通图，回路向量仍张成同一个 $N(A^{\mathsf{T}})$，维数也始终是 $m-n+1$。因此基向量取决于选边方式，循环流空间本身不依赖这项选择。

想一想　为什么任意回路上的单位循环流都满足 $A^{\mathsf{T}}y = 0$?

答　沿闭合路线到达每个节点一次、离开一次。该节点流入与流出的代数贡献相抵，其他节点没有流量，所以每个节点方程都为 0。若某条边的回路行进方向逆着所选箭头，该边坐标必须取 -1。

易错提醒　独立回路数不是简单回路的总个数。图 12.1 有三条简单回路（两个小回路和外圈），但外圈电流是前两个循环流之和，所以独立回路数只有 $2 = m-n+1$。

12.5　Ohm 定律与电路方程

节点电势决定每条边两端的电势差；元件的电导再把电势差转换成电流。这里需要先固定符号：A 的每行是「终点减起点」，所以 Ax 是沿箭头的电势升高；若电流沿箭头为正，Ohm 定律要带一个负号。

设边 k 的正**电导**（conductance）为 c_k，组成对角矩阵 $C = \mathrm{diag}(c_1,\dots,c_m)$。电势高的一端向低的一端流，故

$$e = Ax, \qquad y = -Ce = -CAx.$$

若外部电流源在节点 i 注入 $f_i > 0$，则边电流的「流入减流出」必须等于 $-f_i$。节点平衡写为 $A^{\mathsf{T}}y = -f$，代入 Ohm 定律得到

$$A^{\mathsf{T}}CAx = f.$$

有些电路教材把「流出节点的外部电流」为正，所得源向量与这里的 f 相反，方程右端也会带相反号。物理网络没有改变，改变的是源项的定义。写题时先标明 f_i 表示注入还是流出，再写 $A^Ty=-f$，就能避免只凭「源电流正负」猜符号。

例如只有一条箭头 $1\to 2$ 的电阻边，令 $x_1=0$、$x_2=3$、$c_1=2$。那么 $Ax=3$，而 $y=-c_1Ax=-6$。负号表示实际电流从节点 2 流向节点 1，与箭头相反；它不是负电阻，也不是错误答案。若把箭头反向，则关联矩阵行变号，新的电流坐标为 $+6$，两种记号描述同一物理方向。

同一条边也能展示外部源的符号。令电导为 $c>0$，方向仍为 $1\to 2$，节点 1 注入 $I>0$、节点 2 抽出 I，即 $f=(I,-I)^{\mathsf T}$。接地 $x_2=0$ 后，$L=c\begin{bmatrix}1&-1\\-1&1\end{bmatrix}$，解为 $x=(I/c,0)^{\mathsf T}$，边电流 $y=-c(x_2-x_1)=I$。电流沿箭头从注入点流向抽出点，并满足 $A^Ty=(-I,I)^{\mathsf T}=-f$。这个二节点计算同时核对了电流源符号、接地和 Ohm 定律。

矩阵 $L=A^{\mathsf T}CA$ 称为**加权图拉普拉斯矩阵**（weighted graph Laplacian）。它对称，且

$$x^{\mathsf T}Lx=(Ax)^{\mathsf T}C(Ax)=\sum_{k=1}^{m}c_k(x_{\mathrm{head}(k)}-x_{\mathrm{tail}(k)})^2\ge 0.$$

当图连通且 $c_k>0$ 时，右边为 0 当且仅当所有相邻节点电势相同，也就是 x 为常数。因此 L 的零空间仍是常数向量，秩为 $n-1$。

电导必须严格为正，才能让图中的每条边都传递电流。若两节点间唯一一条边的电导为 0，虽然图论上仍连通，但 L 是零矩阵；平衡源项 $f=(1,-1)^{\mathsf T}$ 也不可能满足 $Lx=f$，接地后另一节点的电势仍无从确定。故结论中的「连通」和「正电导」共同保证电气网络有效连通。

逐项看 L，对角元 L_{ii} 是接在节点 i 上的电导之和；若节点 i,j 由电导为 c 的边相连，则 $L_{ij}=L_{ji}=-c$，没有边相连时该项为 0。若有多条边连接同一对节点，非对角项就是这些电导负值的总和。每行元素和为 0，因此 $L\mathbf{1}_n=0$，再次说明把所有节点电势同时加上常数不会改变边电流。

箭头反向只会把 A 的一行乘以 -1。更一般地，令 D 为对角元取 ±1 的对角矩阵，则一组边反向可写成 $A'=DA$。电导矩阵 C 也是对角矩阵，故 $D^TCD=C$，于是 $A'^TCA'=A^TCA$。这说明电路方程与人为选择的边箭头无关；若更改箭头，边电流坐标也要乘同一个 D，才仍表示同一组物理电流。

由于 $L\mathbf{1}_n=0$，方程 $Lx=f$ 有解的必要条件是 $\sum_i f_i=0$：把方程两边左乘 $\mathbf{1}_n^{\mathsf T}$ 即得。连通且电导为正时，L 的值域是坐标和为 0 的向量空间，所以这也是充分条件。若给节点 j 接地并规定 $x_j=0$，解就唯一：两个解之差在 $N(L)=\mathrm{span}\{\mathbf{1}_n\}$ 中，而第 j 个坐标为 0 迫使这个常数为 0。

唯一性也可直接从能量式看出。若两个满足接地条件的解相减得到 z，则 $Lz=0$ 且 $z_j=0$。于是

$$0=z^TLz=\sum_{k}c_k(z_{\mathrm{head}(k)}-z_{\mathrm{tail}(k)})^2.$$

每项非负且电导为正，故每条边两端的 z 都相同。连通性使全部坐标相同，接地点再迫使它们全为 0。这把「唯一」归结为网络中没有非零且零能量的电势差状态。

把 $Lx=f$ 左乘 $x^{\mathsf T}$，还得到功率平衡

$$x^{\mathsf T}f = x^{T}Lx = \sum_{k=1}^{m} c_k (Ax)_k^2 .$$

左边是外部电流源按节点电势注入网络的功率，右边是各电阻消耗的功率。电导为正时，右边各项非负；这既说明 L 半正定，也给数值解提供了独立核对。

想一想　若电流源向网络总注入 1，却没有任何节点抽出电流，是否存在稳态解？

答　不存在。此时 $\sum_i f_i = 1$，但 $Lx=f$ 左乘 $\mathbf{1}_n^{\mathsf T}$ 会要求 $0=1$。稳态下流入网络的外部电流总量必须等于流出的总量。

注 12.1　原 Session 1.12 摘要的框图写作 $e=Ax$、$y=Ce$，而 Problems 12.2 的题面和解答写作 $y=-CAx$。两处边电流的符号约定相反。本讲沿用题目中的箭头方向和数值解，固定 $Ax=$「终点电势减起点电势」、正电流沿箭头、$y=-CAx$；若把边电压降另记为 $\widetilde{e}=-Ax$，同一关系可写为 $y=C\widetilde{e}$。

■ 例 12.9　**带电流源的方形网络**　按图 12.2 的方向，令 $C=\operatorname{diag}(1,2,2,1)$，并在节点 1 注入单位电流、从节点 3 抽出单位电流，即 $f=(1,0,-1,0)^{\mathsf T}$。求 $A^{\mathsf T}CA$、接地 $x_3=0$ 后的电势 x，以及边电流 y。

解　先把四条边的导电关系相加。由 A 的四行分别得到

$$L = A^{\mathsf T}CA = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 3 & -2 & 0 \\ 0 & -2 & 4 & -2 \\ -1 & 0 & -2 & 3 \end{bmatrix}.$$

方程 $Lx=f$ 只确定电势差，所以可加任意常数。取 $x_3=0$，删去接地节点对应的第三行和第三列，得到

$$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 3 & 0 \\ -1 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

后两行给出 $x_1=3x_2$、$x_1=3x_4$，故 $x_2=x_4$；第一行于是化为 $\frac{4}{3}x_1=1$。所以 $x_1=\frac{3}{4}$、$x_2=x_4=\frac{1}{4}$，并且删去的第三行会自动满足，因为 $f_1+f_2+f_3+f_4=0$。因此

$$x=(\tfrac{3}{4},\tfrac{1}{4},0,\tfrac{1}{4})^{\mathsf T}.$$

边电势差是 $Ax=(-\frac{1}{2},-\frac{1}{4},-\frac{1}{4},-\frac{1}{2})^{\mathsf T}$，所以

$$y=-CAx=(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})^{\mathsf T}.$$

直接核对 $A^{\mathsf T}y=(-1,0,1,0)^{\mathsf T}=-f$：节点 1 的外部电流源注入 1，由两条边各流出 $\frac{1}{2}$；节点 3 收到两条边各 $\frac{1}{2}$，再由电流源抽出 1。这与所选箭头和原题答案一致。

功率平衡也核对了同一组答案。源端净注入功率是 $x^T f = x_1 - x_3 = \frac{3}{4}$；四条边消耗的功率为 $1(\frac{1}{2})^2 + 2(\frac{1}{4})^2 + 2(\frac{1}{4})^2 + 1(\frac{1}{2})^2 = \frac{3}{4}$。如果符号约定、某条边的方向或解出的电流有误，这两个数通常不会相等。 ♦

方法：先定方向，再写方程 给节点和边编号，按「起点 -1、终点 $+1$」写 A；用 Ax 求沿箭头的电势升高；若正电流沿箭头，写 $y = -CAx$；最后用 $A^{\mathsf{T}} y = -f$ 表示外部注入。方向、边序和符号从头到尾保持一致。

本讲小结

- 有向关联矩阵 $A \in \mathbb{R}^{m\times n}$ 每行对应一条边、每列对应一个节点；边 $i \to j$ 的行在 i, j 列分别为 $-1, +1$，且 $(Ax)_k = x_j - x_i$。
- 连通图满足 $\operatorname{rank} A = n - 1$、$N(A) = \operatorname{span}\{\mathbf{1}_n\}$；行空间是节点坐标和为 0 的向量空间。
- 循环流空间是 $N(A^{\mathsf{T}})$，维数为 $\mu = m - n + 1$；树有 $n - 1$ 条边，且 $n - m + \mu = 1$。
- 无外部电流源时 Kirchhoff 电流定律为 $A^{\mathsf{T}} y = 0$；每个独立回路给出一个循环流基向量。
- 若 $C = \operatorname{diag}(c_1, \ldots, c_m)$ 且 $c_k > 0$，沿箭头为正的电流满足 $y = -CAx$；外部注入 f 满足 $A^{\mathsf{T}} y = -f$，从而 $A^{\mathsf{T}} CAx = f$。
- 稳态电流源必须平衡：$\sum_i f_i = 0$。连通网络接地一个节点后，电势唯一；不接地时只差一个整体常数。

练习

12.1 ★ （原题 12.1）按图 12.2 写出 A，求 $N(A)$，并说明 $(1,0,0,0)^{\mathsf{T}}$ 不在行空间中。

12.2 ★ 在图 12.1 中取节点电势 $x = (2, -1, 3, 0)^{\mathsf{T}}$，计算五条边上的 Ax。

12.3 ★ 在五边图上验证 $c_1 = (1, 1, -1, 0, 0)^{\mathsf{T}}$、$c_2 = (0, 0, 1, -1, 1)^{\mathsf{T}}$ 满足 KCL，并写出外圈循环流。

12.4 ★ 连通图有 7 个节点、10 条边。求 $\operatorname{rank} A$、$\dim N(A)$、$\dim N(A^{\mathsf{T}})$ 和独立回路数。

12.5 ★★ 图由两条互不相连的边 1–2、3–4 组成。写出关联矩阵，求秩、零空间维数和左零空间维数，并解释为什么 $n - 1$ 不再是秩。

12.6 ★★ 将图 12.1 的边 3 反向。写出新矩阵 A'，计算 $A'x$，其中 $x = (3, 0, 1, 2)^{\mathsf{T}}$；说明若同一物理电流用新方向记坐标，第三个坐标如何变化。

12.7 ★★ 五边图中给定 $y_1 = 3$、$y_3 = -1$，且 $A^{\mathsf{T}} y = 0$。求其余电流，并用 c_1, c_2 表示 y。

12.8 ★★★ 设图连通且所有边电导都为正。证明 $A^{\mathsf{T}} CAx = f$ 有解当且仅当 $\sum_i f_i = 0$；证明接地一个节点后解唯一。

12.9 ★★★ 在方形图 12.2 上令 $C = I$、$f = (1, 0, -1, 0)^{\mathsf{T}}$，接地 $x_3 = 0$。求 $A^{\mathsf{T}} CA$、x 与 $y = -CAx$，并核对节点平衡。

练习答案与提示

12.1 按四条边的起点、终点写行：

$$A=\begin{bmatrix}-1&1&0&0\\0&-1&1&0\\0&0&1&-1\\-1&0&0&1\end{bmatrix},\qquad N(A)=\operatorname{span}\{(1,1,1,1)^{\mathsf T}\}.$$

解 $Ax=0$ 时，四行依次给出 $x_1=x_2$、$x_2=x_3$、$x_3=x_4$、$x_4=x_1$，所以只有一个自由常数。每行和为 0，故行空间向量的坐标和为 0；$(1,0,0,0)^{\mathsf T}$ 的坐标和为 1，不在行空间。

12.2 $Ax=(x_2-x_1,x_3-x_2,x_3-x_1,x_4-x_1,x_4-x_3)^{\mathsf T}=(-3,4,1,-2,-3)^{\mathsf T}$。

12.3 逐节点检查：$A^{\mathsf T}c_1=(0,0,0,0)^{\mathsf T}$、$A^{\mathsf T}c_2=(0,0,0,0)^{\mathsf T}$。外圈流为 $c_1+c_2=(1,1,0,-1,1)^{\mathsf T}$；边 3 上的一正一负相消，其余四条边沿外圈方向。

12.4 连通图的秩为 $n-1=6$，节点电势零空间维数为 1。边空间有 $m=10$ 个坐标，所以 $\dim N(A^{\mathsf T})=10-6=4$，独立回路数也为 4。核对 Euler 关系 $7-10+4=1$。

12.5 可取 $A=\begin{bmatrix}-1&1&0&0\\0&0&-1&1\end{bmatrix}$，两行非零且支撑互不相交，秩为 2。$N(A)=\{(a,a,b,b)^{\mathsf T}:a,b\in\mathbb{R}\}$，维数为 2；$A^{\mathsf T}u=0$ 的第一行迫使 $u_1=0$、第二行迫使 $u_2=0$，故 $N(A^{\mathsf T})=\{0\}$。图有两个连通分支，秩是 $n-2=2$，不是 $n-1$。

12.6 反向边 3 后得到

$$A'=\begin{bmatrix}-1&1&0&0\\0&-1&1&0\\1&0&-1&0\\-1&0&0&1\\0&0&-1&1\end{bmatrix}.$$

因此 $A'x=(-3,1,2,-1,1)^{\mathsf T}$。同一物理电流沿新箭头的坐标取相反数。令 $D=\operatorname{diag}(1,1,-1,1,1)$，则 $A'=DA$ 且新坐标 $y'=Dy$，只有 y_3 变号；这样 KCL 和物理电流都不变。

12.7 KCL 给出 $y_2=y_1=3$、$y_5=y_2+y_3=2$、$y_4=-y_5=-2$。所以 $y=(3,3,-1,-2,2)^{\mathsf T}=3c_1+2c_2$。

12.8 令 $L=A^{\mathsf T}CA$。对任意 z，$z^TLz=\sum_k c_k(Az)_k^2$；若 $Lz=0$，正电导和连通性使 z 为常数，故 $\ker L=\operatorname{span}\{\mathbf 1_n\}$。由于 L 对称，$\operatorname{Col}(L)=\ker(L)^{\perp}=\{f:\sum_i f_i=0\}$，这给出方程可解的充要条件。必要性也可由 $\mathbf 1_n^{\mathsf T}L=0$ 看出。若 $x,\widetilde{x}$ 都满足接地点 j 的电势为 0，则 $x-\widetilde{x}$ 是常数向量且第 j 坐标为 0，故两解相同。

12.9 $A^{\mathsf T}A=\begin{bmatrix}2&-1&0&-1\\-1&2&-1&0\\0&-1&2&-1\\-1&0&-1&2\end{bmatrix}$. 接地令 $x_3=0$，删除第三行、第三列后解

$$\begin{bmatrix}2&-1&-1\\-1&2&0\\-1&0&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_4\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}.$$

后两行要求 $x_1=2x_2=2x_4$，再由第一行得到 $x_1=1$，所以 $x_2=x_4=\frac12$。代入第三条原方程得到 $-x_2+2x_3-x_4=-1=f_3$。于是 $Ax=(-\frac12,-\frac12,-\frac12,-\frac12)^{\mathsf T}$，$y=-Ax=(\frac12,\frac12,\frac12,\frac12)^{\mathsf T}$，且 $A^{\mathsf T}y=(-1,0,1,0)^{\mathsf T}=-f$。

复习课 1

线性方程、消元与子空间

Linear Equations, Elimination, and Subspaces

对应材料：MIT 18.06SC Unit 1 Exam 及解答（MIT18_06SCF11_ex1.pdf、ex1s.pdf，题 1–4）；Sessions 1.1–1.12 的课时摘要、习题与解答；课程第 1–12 讲讲义。考试复习目录另有 FinalRevsum 和 Final Exam，属于全课程/期末材料，本复习课不把它们当作 Unit 1 原卷。以下标注为 Unit 1 Exam 的例题取自该原卷，其余自测题改编自对应课时习题；没有声称这些自测题是官方考试题。

本讲目标

(1) 能沿着「方程组—消元—主元—子空间—基」的路线选择工具，并检查每步的尺寸与条件。

(2) 能从行最简形读出秩、自由变量、全部解和四个基本子空间的维数。

(3) 能把同一套线性代数语言用于矩阵空间、图距离与电路网络。

(4) 能区分相容与唯一、主元列编号与原矩阵列、图方向与物理电流方向。

R1.1 知识地图

第 1–12 讲的共同问题是：给定线性关系，怎样判断目标能否到达，解有多少自由度，又怎样把这些自由度组织成基？图 R1.1 把课程中的计算步骤和空间结构排在同一条主线上。复习时可先看清所在位置，再选对应的矩阵表示。

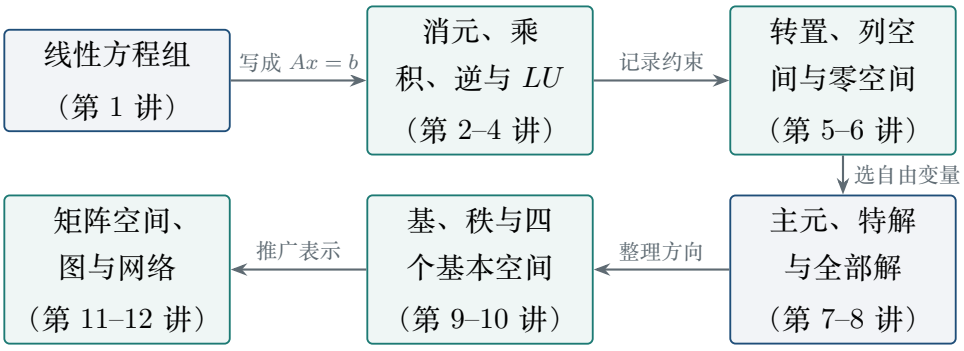


图 R1.1　消元把方程变成主元与自由变量，后者给出解空间、基和四空间，并可推广到矩阵与网络。

遮住表格中间一列，试着自己说出每个概念的定义要点；再遮住最后一列，回忆它通常用来解决什么问题。

概念	定义要点	典型用法
相容	$b \in C(A)$，即 b 是 A 的列的线性组合	判断 $Ax = b$ 是否有解
主元与自由变量	行最简形中主元列定主元变量，其余坐标可自由选	写齐次解的基或非齐次通解
基与维数	线性无关且张成；基向量个数是维数	检查表示是否唯一，计算自由度
四个基本空间	$C(A), N(A^{\mathsf{T}}) \subseteq \mathbb{R}^m$；$C(A^{\mathsf{T}}), N(A) \subseteq \mathbb{R}^n$	读出输入端和输出端的约束
关联矩阵	每行记录一条定向边的起点 -1 与终点 $+1$	把节点电势变成边电势升高

R1.2　核心结论回顾

R1.2.1　方程、消元与分解（第 1–4 讲）

定理（可解性与全部解）　设 $A \in \mathbb{R}^{m\times n}$。方程 $Ax = b$ 有解当且仅当 $b \in C(A)$。若 x_p 是一个特解，则全部解为 $x_p + N(A)$；特别地，齐次系统的解集就是子空间 $N(A)$。

关键　若 x 和 x_p 都满足 $Ax = b$，相减得 $A(x - x_p) = 0$；反过来，对任意 $z \in N(A)$，$A(x_p + z) = b$。完整论证见第 1、6、8 讲。

用法　先用增广矩阵判断相容，再找一个特解和零空间基。两个不同解会沿着同一个零空间方向生成无穷多个解；「有解」本身不代表「唯一」。

定理（消元保持等价系统）　交换两行、把一行乘非零数、或以一行的倍数加到另一行，都必须同步作用于右端；这些可逆操作保持方程组的全部解。若消元得到 $EA = U$ 且 E 可逆，则 $Ax = b$ 等价于 $Ux = Eb$。

关键　每个操作都能撤销，所以不会丢掉原解，也不会新增解。若无行交换且每步主元非零，累计消元可写为 $A = LU$；若需行交换，则写 $PA = LU$。完整构造见第 2、4 讲。

用法　回代前先确认当前主元非零；遇到零主元先检查能否换行，不能直接除以零。求多个右端时，先分解一次，再对每个右端分别做前代与回代。

矩阵乘法要求内维度相同，$(m \times n)(n \times p) = m \times p$；它满足结合律与分配律，但一般不交换。转置和逆的顺序都要反过来：$(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$、$(AB)^{-1} = B^{-1}A^{-1}$（后一式要求 A, B 可逆方阵）。先写出因子顺序，再逐项计算，能避免把「先做什么」误读成乘积从左到右。

R1.2.2　转置、子空间与解的参数（第 5–8 讲）

定理（主元数与自由度）　若 $R = \operatorname{rref}(A)$ 有 r 个主元，则 $\operatorname{rank} A = r$。行化简保留零空间；$N(A)$ 有 $n - r$ 个自由坐标。分别令一个自由变量为 1、其余为 0 得到的特殊解构成 $N(A)$ 的一组基。

关键　把主元变量用自由变量表示，齐次解便是自由参数的线性组合。不同特殊解的自由坐标是不同的标准基向量，因此它们线性无关并张成全部零空间。完整证明见第 7 讲。

用法　读 R 时分别数主元列和零行：自由变量数是 $n-r$，零行数是 $m-r$。前者描述输入解的自由度，后者描述输出端可能出现的相容条件。

> **定理（列空间、零空间与方程组）**　$C(A)=\{Ax : x\in\mathbb{R}^n\}\subseteq\mathbb{R}^m$，$N(A)=\{x : Ax=0\}\subseteq\mathbb{R}^n$。$Ax=b$ 可解当且仅当 $b\in C(A)$；可解时，解集是特解加 $N(A)$。两个子空间的交仍是子空间；它们的和包含两者，而并一般不是子空间。

关键　列空间由列的线性组合构成，零空间由齐次方程刻画；将矩阵竖向叠放时，公共零空间是各零空间的交。完整说明见第 5–6 讲。

用法　若右端不在列空间，增广消元会出现 $[0\ \cdots\ 0\mid c]$ 且 $c\neq 0$。若系统相容，再用零空间判断唯一性：$N(A)=\{0\}$ 才能推出解唯一。

R1.2.3　基、秩与四个基本空间（第 9–10 讲）

> **定理（基与四空间维数）**　若 $A\in\mathbb{R}^{m\times n}$ 且 $\operatorname{rank}A=r$，则 $\dim C(A)=\dim C(A^{\mathsf{T}})=r$，$\dim N(A)=n-r$，$\dim N(A^{\mathsf{T}})=m-r$。$C(A),N(A^{\mathsf{T}})$ 位于 $\mathbb{R}^m$；$C(A^{\mathsf{T}}),N(A)$ 位于 $\mathbb{R}^n$。

关键　秩是列空间和行空间的维数；主元列数给出 r，秩–零度关系给出两个零空间的维数。完整证明见第 9–10 讲。

用法　列空间基要取原矩阵 A 的主元列；主元列编号从 R 读。零空间基取自由变量特殊解；左零空间基可由 $A^{\mathsf{T}}y=0$ 求出。检查答案时先核对每个基向量所在的坐标空间。

R1.2.4　矩阵空间与图网络（第 11–12 讲）

第 11 讲把矩阵本身当作向量；例如 $m\times n$ 矩阵空间维数为 mn，秩 1 矩阵可写成 uv^{T}。第 12 讲把每条定向边作为关联矩阵的一行，边 $i\to j$ 对应的分量是 x_j-x_i。这里都使用矩阵语言，但对象和坐标不同：邻接矩阵幂数游走，关联矩阵乘节点势得边势差。

> **定理（连通图的计数与电路方程）**　连通图有 n 个节点、m 条边时，其关联矩阵 $A\in\mathbb{R}^{m\times n}$ 满足 $\operatorname{rank}A=n-1$，$\dim N(A^{\mathsf{T}})=m-n+1$。若 $C=\operatorname{diag}(c_1,\dots,c_m)$ 且每个 $c_k>0$，沿箭头为正的电流满足 $y=-CAx$；外部注入向量 f 按本讲约定满足 $A^{\mathsf{T}}y=-f$，于是 $A^{\mathsf{T}}CAx=f$。

关键　Ax 是每条边的终点势减起点势，$A^{\mathsf{T}}y$ 是节点的流入减流出；将 $y=-CAx$ 代入节点平衡即可得到网络方程。完整推导见第 12 讲。

用法　先固定边方向和注入符号，再写 A、C 与 f。连通网络的源项需满足 $\sum_i f_i=0$；接地一个节点后，电势解唯一。若换边方向，对应的关联矩阵行和电流坐标要同时变号。

R1.3　题型与例题

R1.3.1　题型一：由相容性和唯一性读出秩

■ **例 R1.1　Unit 1 Exam 第 1 题**　某个 $m\times n$ 矩阵 A 满足：$Ax=(1,1,1)^{\mathsf T}$ 无解，而 $Ax=(0,1,0)^{\mathsf T}$ 恰有一个解。求 $m,n,r=\operatorname{rank}A$ 的所有可能关系，写出 $Ax=0$ 的全部解，并给一个符合条件的矩阵。

解　先从右端坐标和解的个数读取信息。右端有三个分量，所以 $m=3$。向量 $(0,1,0)^{\mathsf T}$ 可达且非零，因此 $\operatorname{rank}A\geq 1$；向量 $(1,1,1)^{\mathsf T}$ 不可达，说明 $C(A)$ 不是整个 $\mathbb{R}^3$，所以 $r<3$。

某个相容右端恰有一个解，意味着 $N(A)=\{0\}$：若存在非零 $z\in N(A)$，一个解 x_p 会连同所有 x_p+tz 都成为解。于是 $n-r=0$，即 $n=r$。综上，$m=3$，$n=r\in\{1,2\}$，且齐次方程只有零解。

可取 $A=(0,1,0)^{\mathsf T}$，它是 3×1 矩阵，秩为 1。此时 $Ax=(0,x,0)^{\mathsf T}$；右端 $(0,1,0)^{\mathsf T}$ 的唯一解是 $x=1$，而 $(1,1,1)^{\mathsf T}$ 不在列空间中。 ♦

点评　「某个右端恰有一个解」给出 $N(A)=\{0\}$，不要求 A 是方阵。最常见的误读是把奇异或矩形矩阵一概判作无解；若将不可达右端换成列空间中的向量，结论可以改变。

R1.3.2　题型二：恢复消元顺序、逆矩阵与 LU

■ **例 R1.2　Unit 1 Exam 第 2 题**　一个 3×3 矩阵 A 依次经过操作 E_{21}（第二行减 4 倍第一行）、E_{31}（第三行减 3 倍第一行）、E_{23}（第二行减第三行）后成为 I。求 A^{-1}、A，并给出 $A=LU$ 中的 L。

解　先按实际作用顺序写左乘积：$E_{23}E_{31}E_{21}A=I$，所以

$$A^{-1}=E_{23}E_{31}E_{21}=\begin{bmatrix}1&0&0\\-1&1&-1\\-3&0&1\end{bmatrix}.$$

撤销时次序反过来，每一步把被减去的倍数加回：

$$A=E_{21}^{-1}E_{31}^{-1}E_{23}^{-1}=\begin{bmatrix}1&0&0\\4&1&1\\3&0&1\end{bmatrix}.$$

直接相乘可验 $A^{-1}A=I$。记录前向消元乘数可取

$$L=\begin{bmatrix}1&0&0\\4&1&0\\3&0&1\end{bmatrix},\qquad U=\begin{bmatrix}1&0&0\\0&1&1\\0&0&1\end{bmatrix}.$$

这组 L 的对角元为 1，且直接相乘得到 $LU=A$。 ♦

点评　连续左乘操作时，最后做的操作写在乘积最左侧；求原矩阵要反序并逐步取逆。分解中的 L 记录前向消元乘数，不能把 Gauss–Jordan 消元矩阵本身当成 L。

R1.3.3 题型三：含参数的矩阵与完整解集

■ **例 R1.3　Unit 1 Exam 第 3 题**　令

$$A_c=\begin{bmatrix}1&1&2&4\\3&c&2&8\\0&0&2&2\end{bmatrix},\qquad b_c=(1,c,0)^{\mathsf T}.$$

分别讨论 $c=3$ 与 $c\neq 3$ 时的列空间基、零空间基和 $A_cx=b_c$ 的全部解。

解　消去第二行的第一列，得到第二行为 $(0,c-3,-4,-4)$，第三行仍为 $(0,0,2,2)$。若 $c\neq 3$，主元列为 $1,2,3$，故列空间基可取原矩阵的三列 a_1,a_2,a_3。若 $c=3$，第二列等于第一列，主元列为 $1,3$，列空间基为 a_1,a_3。对所有 c 都有 $a_4=2a_1+a_3$。

第三行给出 $x_3=-x_4$。当 $c\neq 3$，第二行要求 $x_2=0$，第一行给出 $x_1=-2x_4$。因此

$$N(A_c)=\operatorname{span}\{(-2,0,-1,1)^{\mathsf T}\},\qquad x=(0,1,0,0)^{\mathsf T}+t(-2,0,-1,1)^{\mathsf T}.$$

当 $c=3$，第二行与第三行相关，x_2,x_4 都自由，故

$$N(A_3)=\operatorname{span}\{(-1,1,0,0)^{\mathsf T},(-2,0,-1,1)^{\mathsf T}\},$$

$$x=(1,0,0,0)^{\mathsf T}+s(-1,1,0,0)^{\mathsf T}+t(-2,0,-1,1)^{\mathsf T}.$$

把特解代回 $A_cx=b_c$，把两个零空间方向代回 $A_cx=0$，即可核对全部解。 ♦

点评　参数值 $c=3$ 让第二列与第一列重合，秩从 3 降为 2，因此自由变量增加一个。先找使主元消失的参数，再讨论秩；只在一个一般参数值上消元会漏掉特殊情形。

R1.3.4 题型四：从 R 读出原矩阵的列关系

■ **例 R1.4　Unit 1 Exam 第 4(b) 题**　设 A 为 3×5 矩阵，且行化简后

$$R=\operatorname{rref}(A)=\begin{bmatrix}1&4&0&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}.$$

求 $\operatorname{rank}A$、原矩阵哪些列构成列空间基、$N(A)$ 的一组基，以及列间关系。

解　主元列是第 $1,4,5$ 列，因此 $\operatorname{rank}A=3$，原矩阵的 a_1,a_4,a_5 构成 $C(A)$ 的一组基。第 2 列在 R 中等于第 1 列的 4 倍，第 3 列为零；可逆行变换保留列向量间的线性关系，所以 $a_2=4a_1$、$a_3=0$。

由 $Rx=0$ 得 $x_1+4x_2=0$、$x_4=x_5=0$，而 x_2,x_3 自由。因此

$$N(A)=N(R)=\operatorname{span}\{(-4,1,0,0,0)^{\mathsf T},(0,0,1,0,0)^{\mathsf T}\}.$$

维数为 $2 = 5 - 3$，与两个自由变量一致。 ♦

点评　从 R 读的是主元列编号，实际列向量必须回到 A 中取。行化简会改变列向量本身，但会保留哪些列相关以及相关系数；本题去掉 R 后，仍可由列关系核对答案。

R1.3.5 题型五：矩阵子空间与网络方程

■ **例 R1.5　Unit 1 Exam 第 4(c) 题**　在 3×3 矩阵空间中，所有可能的行最简形 R 所张成的子空间是什么？求其维数。

解　行最简形的每个主元都位于对角线或其上方，因此所有这样的 R 都是上三角矩阵。反过来，六个上三角位置的标准矩阵都能由行最简形线性组合得到：E_{11}, E_{33} 本身是行最简形，$E_{22} = I - E_{11} - E_{33}$；$E_{12}, E_{13}$ 可由第一行分别取 $(1,1,0)$、$(1,0,1)$ 的秩一行最简形与 E_{11} 相减得到；E_{23} 可由两个仅第二行第三列不同的秩二行最简形相减得到。

所以所求空间恰为

$$\left\{ \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} : a,b,c,d,e,f \in \mathbb{R} \right\},$$

其六个自由元素对应一组基，维数为 6。 ♦

点评　题目问的是「所有 R 的张成空间」，不是「所有 R 的集合」。行最简形本身通常不对加法封闭，但它们的线性组合能张成一个熟悉的矩阵子空间。

■ **例 R1.6　改编自 Session 1.11 Problem 11.2：矩阵作用在矩阵空间上**　令

$$G = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix},$$

把变换 $X \mapsto GX$ 看作 3×3 矩阵空间上的线性变换。描述其零空间与值域，并求两者维数。

解　矩阵乘积 GX 按列计算，所以 $GX = 0$ 当且仅当 X 的每一列都属于 $N(G)$。方程 $Gx = 0$ 等价于 $x_1 = x_2 = x_3$，故 $N(G) = \operatorname{span}\{(1,1,1)^{\mathsf{T}}\}$。因此变换的零空间为

$$\left\{ \begin{bmatrix} a & b & c \\ a & b & c \\ a & b & c \end{bmatrix} : a,b,c \in \mathbb{R} \right\},$$

维数为 3。

G 的每列坐标和为 0，且 $\operatorname{rank} G = 2$，所以 $C(G) = \{u \in \mathbb{R}^3 : u_1 + u_2 + u_3 = 0\}$。$GX$ 的三列可独立取在这个二维平面内，值域维数为 $2 \cdot 3 = 6$。两者维数和为 9，恰是定义域 $\mathbf{M}_{3\times 3}$ 的维数。 ♦

点评　矩阵空间中的 X 有三列，左乘 G 是对每一列分别作用。不要把 $GX=0$ 误读成「每一行相同」而不验证；从列属于 $N(G)$ 推出三行相同才是完整理由。

■ **例 R1.7　改编自 Session 1.12 Problem 12.2：加权方形电路**　采用第 12 讲方形图的边方向，令 $C=\mathrm{diag}(1,2,2,1)$，节点注入为 $f=(1,0,-1,0)^{\mathsf T}$。求 $L=A^{\mathsf T}CA$、接地 $x_3=0$ 时的电势，以及沿箭头为正的电流 $y=-CAx$。

解　按四条边分别累加电导贡献，得到

$$L=A^{\mathsf T}CA=\begin{bmatrix}2&-1&0&-1\\-1&3&-2&0\\0&-2&4&-2\\-1&0&-2&3\end{bmatrix}.$$

把第三个节点接地，删去第三行、第三列，解

$$\begin{bmatrix}2&-1&-1\\-1&3&0\\-1&0&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_4\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}.$$

后两式给出 $x_1=3x_2=3x_4$，第一式为 $\frac{4}{3}x_1=1$。所以

$$x=(\tfrac34,\tfrac14,0,\tfrac14)^{\mathsf T},\qquad Ax=(-\tfrac12,-\tfrac14,-\tfrac14,-\tfrac12)^{\mathsf T},$$

$$y=-CAx=(\tfrac12,\tfrac12,\tfrac12,\tfrac12)^{\mathsf T}.$$

核对 $A^{\mathsf T}y=(-1,0,1,0)^{\mathsf T}=-f$；电源净功率 $x^{\mathsf T}f=\frac34$，电阻耗散 $\sum_k c_k(Ax)_k^2=\frac34$。 ♦

点评　注意本讲把 Ax 定义为沿箭头的电势升高，因此电流为 $-CAx$。若把电流正方向换成沿电势下降，边坐标符号和方程必须一起换；只改一个负号会破坏节点平衡。

R1.4　常见错误

把奇异误当成所有右端都无解　错误说法：A 不可逆，所以 $Ax=b$ 无解。反例：$A=\begin{bmatrix}1&0\\0&0\end{bmatrix}$ 对 $b=(1,0)^{\mathsf T}$ 有解 $(1,t)^{\mathsf T}$。正确判断：奇异只说明不是每个右端都能唯一求解；先查 $b\in C(A)$，再查 $N(A)$。

把行最简形的主元列当成列空间基　错误做法：在 $R=\mathrm{rref}(A)$ 中读出主元后，直接把 R 的主元列当作 $C(A)$ 的基。反例：$A=\begin{bmatrix}1&1\\1&1\end{bmatrix}$ 可化为 $R=\begin{bmatrix}1&1\\0&0\end{bmatrix}$；$R$ 的第一列是 $(1,0)^{\mathsf T}$，而 $C(A)$ 的基应取 A 第一列 $(1,1)^{\mathsf T}$。正确做法：用 R 找编号，用 A 取向量。

把零行数当成自由变量数　错误做法：只数消元后的零行来决定 $N(A)$ 的维数。例：3×5 矩阵若秩为 3，零行数为 0，但自由变量数是 $5-3=2$。正确做法：自由变量数看列数减秩，零行数看行数减秩；两者属于不同坐标空间。

把长度为 k 的游走当成最短距离　错误说法：若 $(A^k)_{ij}>0$，两点距离就是 k。反例：图的邻接矩阵满足 $(A^2)_{11}>0$，但 $d(1,1)=0$。正确做法：从 $A^0=I$ 开始，找 $(A^k)_{ij}$ 第一次为正的 k。

本讲小结

- 遇到 $Ax=b$，先查 $b\in C(A)$；若相容，全部解为 $x_p+N(A)$。
- 遇到零主元，先看能否换行；r 个主元留下 $n-r$ 个自由变量，特殊解给出零空间基。
- 由 R 的主元列编号定位原矩阵的列空间基；四空间维数为 $r,r,n-r,m-r$，并先写明各自所在的 $\mathbb{R}^m$ 或 $\mathbb{R}^n$。
- 做乘积与分解时保持顺序：$(AB)^{\mathsf{T}}=B^{\mathsf{T}}A^{\mathsf{T}}$、$(AB)^{-1}=B^{-1}A^{-1}$、$PA=LU$。
- 矩阵空间的元素可按列看作向量；图的邻接矩阵数游走，关联矩阵把节点势映成边势差。
- 电路中若正电流沿箭头，则 $y=-CAx$ 且 $A^{\mathsf{T}}y=-f$；接地固定整体电势自由度。

练习

R1.1 ★　（改编自 Session 1.1 Problem 1.1）令 $w_1=(1,2,3)^{\mathsf{T}}$、$w_2=(4,5,6)^{\mathsf{T}}$、$w_3=(7,8,9)^{\mathsf{T}}$。找一个非平凡线性关系，并求这三列张成空间的维数。

R1.2 ★　（改编自 Session 1.2 Problem 2.1）解方程组 $2x+3y=5$、$6x+15y=12$。写出消元乘数和主元，并代回两式核对。

R1.3 ★　（改编自 Session 1.3 Problem 3.1）令 $A=\begin{bmatrix}1&2\\3&4\end{bmatrix}$、$B=\begin{bmatrix}1&0\\0&0\end{bmatrix}$、$C=\begin{bmatrix}0&0\\5&6\end{bmatrix}$。分别计算 $AB+AC$ 与 $A(B+C)$。

R1.4 ★★　（改编自 Session 1.4 Problem 4.1）对 $A=\begin{bmatrix}1&3&0\\2&4&0\\2&0&1\end{bmatrix}$ 做不换行消元，求 E,U 使 $EA=U$，并写出 $A=LU$。

R1.5 ★　（改编自 Session 1.5 Problem 5.1）取 $P=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}$。求 P^2,P^3，并验证 $P^{\mathsf{T}}=P^{-1}$。

R1.6 ★　（改编自 Session 1.6 Problem 6.2）平面 $x-3y-z=12$ 经过 $(12,0,0)$。写成一个特解加两个方向向量的参数式，并给出对应齐次平面的零空间基。

R1.7 ★★（改编自 Session 1.8 Problem 8.2）令 $U=\begin{bmatrix}1&2&3\\0&0&4\end{bmatrix}$。求 $Ux=0$ 的零空间基，并求 $Ux=(5,8)^{\mathsf T}$ 的全部解。

R1.8 ★★（改编自 Session 1.9 Problem 9.2）求平面 $x-2y+3z=0$ 的一组基、它与 xy 平面的交集的一组基，以及所有垂直于该平面的向量。

R1.9 ★★ 在第 11 讲五节点社交图中，边为 $\{1,2\},\{1,5\},\{2,3\},\{2,4\},\{3,4\},\{4,5\}$。求节点 2 到 5 的距离，并列出所有长度为 2 的游走。

R1.10 ★★（改编自 Session 1.12 Problem 12.1）按第 12 讲方形图的边方向写出关联矩阵 A，求 $N(A)$，并说明 $(1,0,0,0)^{\mathsf T}$ 不在行空间中。

练习答案与提示

R1.1 $w_1-2w_2+w_3=0$，且 w_1,w_2 不成比例；故秩为 2，张成空间是 $\mathbb{R}^3$ 中的过原点平面。

R1.2 $R_2\leftarrow R_2-3R_1$ 得 $6y=-3$，故 $y=-\frac12$、$x=\frac{13}{4}$；主元为 $2,6$，代回第二式得 12。

R1.3 $AB+AC=A(B+C)=\begin{bmatrix}11&12\\23&24\end{bmatrix}$，其中 $AB=\begin{bmatrix}1&0\\3&0\end{bmatrix}$、$AC=\begin{bmatrix}10&12\\20&24\end{bmatrix}$。

R1.4 依次取 $R_2\leftarrow R_2-2R_1$、$R_3\leftarrow R_3-2R_1$、$R_3\leftarrow R_3-3R_2$，得 $U=\begin{bmatrix}1&3&0\\0&-2&0\\0&0&1\end{bmatrix}$。相应 $E=\begin{bmatrix}1&0&0\\-2&1&0\\4&-3&1\end{bmatrix}$，其逆为 $L=\begin{bmatrix}1&0&0\\2&1&0\\2&3&1\end{bmatrix}$。直接相乘 $LU=A$。

R1.5 $P^2=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$、$P^3=I$；又 $P^{\mathsf T}P=I$，所以 $P^{\mathsf T}=P^{-1}=P^2$。

R1.6 参数式为 $(x,y,z)^{\mathsf T}=(12,0,0)^{\mathsf T}+s(3,1,0)^{\mathsf T}+t(1,0,1)^{\mathsf T}$。齐次平面的零空间基为 $(3,1,0)^{\mathsf T},(1,0,1)^{\mathsf T}$。

R1.7 $Ux=0$ 给出 $N(U)=\operatorname{span}\{(-2,1,0)^{\mathsf T}\}$；对右端 $(5,8)^{\mathsf T}$ 得 $x_3=2$、$x_1+2x_2=-1$，全部解为 $(-1,0,2)^{\mathsf T}+t(-2,1,0)^{\mathsf T}$。

R1.8 令 $x=2y-3z$，平面基为 $(2,1,0)^{\mathsf T},(-3,0,1)^{\mathsf T}$；与 xy 平面的交集基为 $(2,1,0)^{\mathsf T}$，法向量为 $(1,-2,3)^{\mathsf T}$。

R1.9 邻接矩阵满足 $A_{25}=0$，而 $(A^2)_{25}=2$，因为共有两个公共邻居 $1,4$。两条游走是 $2\to1\to5$ 和 $2\to4\to5$，所以距离为 2。

R1.10 按边方向，

$$A=\begin{bmatrix}-1&1&0&0\\0&-1&1&0\\0&0&1&-1\\-1&0&0&1\end{bmatrix}.$$

令 $Ax=0$ 得 $x_1=x_2=x_3=x_4$，故 $N(A)=\operatorname{span}\{(1,1,1,1)^{\mathsf T}\}$。

每行坐标和为 0，所以行空间中每个向量坐标和为 0；$(1,0,0,0)^{\mathsf T}$ 的坐标和为 1，不在行空间。

第二部分

正交、行列式与特征值

第 13 讲

正交向量与子空间

Orthogonal Vectors and Subspaces

对应材料：MIT 18.06SC《Orthogonal vectors and subspaces》课时摘要；习题集 4.1 第 7、32 题；Strang《Introduction to Linear Algebra》第 4.1 节。

本讲目标

(1) 会计算点积、长度和夹角余弦，并能用柯西–施瓦茨不等式估计点积。

(2) 会判断向量和子空间是否正交，能从点积展开推出勾股等式。

(3) 会求简单子空间的正交补，并说明其维数为何与原空间互补。

(4) 能证明行空间与零空间、列空间与左零空间互为正交补。

(5) 会用左零空间向量证明方程组无解，并判断 $A^{\mathsf{T}}A$ 是否可逆。

第 10 讲给出了矩阵的四个基本子空间及其维数。本讲补上它们之间的几何关系：行空间和零空间互相垂直，列空间和左零空间也互相垂直。连接代数和几何的工具只有一个：点积。

13.1 点积记录了什么？

在 $\mathbb{R}^2$ 中，$x=(x_1,x_2)$ 与 $y=(y_1,y_2)$ 的点积是 $x_1y_1+x_2y_2$。例如 $(1,2)$ 与 $(2,1)$ 的点积是 4；而 $(1,0)$ 与 $(0,1)$ 的点积是 0。点积把两组坐标逐项相乘再相加，也可以用矩阵乘法写成 $x^{\mathsf{T}}y$。

定义 13.1（点积与长度） 对 $x,y\in\mathbb{R}^n$，定义它们的**点积**（dot product）为

$$\langle x,y\rangle = x_1y_1+\cdots+x_ny_n = x^{\mathsf{T}}y.$$

向量 x 的**欧几里得长度**（Euclidean norm）定义为 $\|x\|=\sqrt{\langle x,x\rangle}$。

例如 $x=(1,2,2)$ 的长度平方是 $1+4+4=9$，所以 $\|x\|=3$；若 $y=(2,0,1)$，则 $\langle x,y\rangle=4$。零向量的长度为 0，且它与任意向量的点积都是 0。

点积满足对称性和线性性：$\langle x,y\rangle=\langle y,x\rangle$，$\langle x,y+z\rangle=\langle x,y\rangle+\langle x,z\rangle$，$\langle x,cy\rangle=c\langle x,y\rangle$。长度平方因此可写成 $\|x\|^2=\langle x,x\rangle$。

$$\boxed{\|x\|^2=\langle x,x\rangle=x^{\mathsf{T}}x}$$

如果 x, y 都非零，它们的夹角 $\theta \in [0, \pi]$ 由

$$\cos\theta = \frac{\langle x, y\rangle}{\|x\|\|y\|}$$

确定。这个公式有意义的关键是右边落在 $[-1, 1]$；柯西–施瓦茨不等式正好保证这一点。

定理 13.2（柯西–施瓦茨不等式）　对任意 $x, y \in \mathbb{R}^n$，有

$$|\langle x, y\rangle| \le \|x\|\|y\|.$$

思路　若 $y \neq 0$，考察 $\|x - ty\|^2$。它对任意实数 t 都非负；选取使其最小的 t，就能把含 $\langle x, y\rangle$ 的项单独留下。若 $y = 0$，不等式两边都是 0。

证明　若 $y = 0$，结论成立。现在设 $y \neq 0$，取 $t = \langle x, y\rangle / \|y\|^2$。由长度平方非负，

$$0 \le \|x - ty\|^2 = \|x\|^2 - 2t\langle x, y\rangle + t^2\|y\|^2 = \|x\|^2 - \frac{\langle x, y\rangle^2}{\|y\|^2}.$$

两边乘以正数 $\|y\|^2$，得到 $\langle x, y\rangle^2 \le \|x\|^2\|y\|^2$。两边开平方即得结论。 ■

当 $y \neq 0$ 时，把同一个二次式配方可见它为何在所选的 t 处最小：

$$\|x - ty\|^2 = \|y\|^2\left(t - \frac{\langle x, y\rangle}{\|y\|^2}\right)^2 + \|x\|^2 - \frac{\langle x, y\rangle^2}{\|y\|^2}.$$

最小值对应的残差 $x - ty$ 与 y 正交；也就是说，从 x 到直线 $\operatorname{span}\{y\}$ 的最短误差垂直于这条直线。

推论 13.3（柯西–施瓦茨不等式的等号条件）　等号 $|\langle x, y\rangle| = \|x\|\|y\|$ 成立，当且仅当 x, y 线性相关。

证明　若两向量线性相关，直接代入即得等号。反过来，若等号成立且 $y \neq 0$，在定理证明中取 $t = \langle x, y\rangle / \|y\|^2$，便有

$$\|x - ty\|^2 = \|x\|^2 - \frac{\langle x, y\rangle^2}{\|y\|^2} = 0.$$

所以 $x = ty$，两向量线性相关。若 $y = 0$，两向量本来就线性相关。 ■

等号条件说明柯西–施瓦茨的边界由同一条直线上的向量达到。前面的 $a = (1, 2)$、$b = (2, 1)$ 不成倍数，因此 $|\langle a, b\rangle| = 4 < 5 = \|a\|\|b\|$，夹角余弦严格小于 1。

■ **例 13.1　柯西–施瓦茨不等式取等号**　令 $a = (1, 2)$、$b = (-2, -4)$。验证这两个向量达到柯西–施瓦茨不等式的等号。

解　有 $\langle a, b\rangle = -10$、$\|a\|^2 = 5$、$\|b\|^2 = 20$，所以

$$|\langle a, b\rangle|^2 = 100 = 5 \cdot 20 = \|a\|^2\|b\|^2.$$

两个向量成比例，方向相反，故夹角为 π；这与等号条件一致。 ♦

绝对值使柯西–施瓦茨同时给出点积的上界和下界。方向相同的非零向量达到正端点，方向相反的非零向量达到负端点；因此夹角公式中的余弦完整落在 $[-1,1]$ 内。

柯西–施瓦茨还能推出长度的三角不等式。它说明两个向量相加后的长度，不会超过分别走完两段所需长度之和。

推论 13.4（三角不等式）　对 $x,y\in\mathbb{R}^n$，有 $\|x+y\|\le\|x\|+\|y\|$。

证明　展开长度平方并对交叉项使用柯西–施瓦茨不等式：

$$\|x+y\|^2=\|x\|^2+2\langle x,y\rangle+\|y\|^2\le\|x\|^2+2\|x\|\|y\|+\|y\|^2=\big(\|x\|+\|y\|\big)^2.$$

两边非负，开平方即得结论。 ■

例如 $x=(1,0)$、$y=(0,1)$ 时，$\|x+y\|=\sqrt{2}<2=\|x\|+\|y\|$。两个非零向量同向时等号成立；方向不同或相反时，起点到终点的直线距离更短。若其中一个向量为零，等号也成立。

■ **例 13.2　由点积求夹角余弦**　求 $a=(1,2)$ 与 $b=(2,1)$ 的夹角余弦，并判断夹角是锐角还是钝角。

解　先算点积和长度：$\langle a,b\rangle=4$，$\|a\|=\|b\|=\sqrt{5}$。所以 $\cos\theta=4/5>0$，夹角是锐角。柯西–施瓦茨不等式保证这个余弦值不会超过 1。 ♦

■ **例 13.3　负点积说明钝角**　令 $u=(1,0)$、$v=(-1,1)$。求它们的夹角余弦，并判断夹角的类型。

解　点积为 -1，长度分别为 1 和 $\sqrt{2}$，所以 $\cos\theta=-1/\sqrt{2}<0$，夹角是钝角。点积为正、零、负，分别对应锐角、直角、钝角；零向量除外，因为它没有定义好的夹角。 ♦

想一想　零向量与每个向量的点积都是 0，能因此说它和每个非零向量的夹角都是直角吗？
答　不能。夹角公式的分母含有 $\|0\|=0$，因此零向量没有定义好的夹角。点积为 0 是正交的代数条件；零向量满足这个等式，但角度的几何说法只用于两个非零向量。

13.2　垂直意味着什么？

向量 $u=(2,1)$ 和 $v=(-1,2)$ 的点积是 $-2+2=0$。它们都不是零向量，却彼此垂直；零点积描述的正是直角。

定义 13.5（正交向量）　若 $u,v\in\mathbb{R}^n$ 满足 $\langle u,v\rangle=0$，就称 u 与 v **正交**（orthogonal），记作 $u\perp v$。

例如 $(2,1)\perp(-1,2)$；$(1,0)$ 与 $(1,1)$ 不正交，因为点积为 1。零向量与每个向量正交，这是

定义直接给出的边界情形。

定理 13.6（勾股等式与正交）　对任意 $u, v \in \mathbb{R}^n$，有

$$u \perp v \quad \Longleftrightarrow \quad \|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

若 $u \perp v$，还成立 $\|u-v\|^2 = \|u\|^2 + \|v\|^2$。

直观理解　把 u 与 v 的尾端都放在原点，若它们垂直，就组成直角三角形的两条直角边。向量 $u+v$ 是平行四边形的对角线，也就是这条三角形的斜边。

思路　展开 $\|u+v\|^2 = \langle u+v, u+v\rangle$。交叉项是 $2\langle u, v\rangle$，所以它为零恰好等价于勾股等式。对 $u-v$ 做同一展开，交叉项只改变符号。

证明　由点积的线性与对称性，

$$\|u+v\|^2 = \|u\|^2 + 2\langle u, v\rangle + \|v\|^2.$$

因此 $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ 当且仅当 $2\langle u, v\rangle = 0$，也就是 $u \perp v$。同理，

$$\|u-v\|^2 = \|u\|^2 - 2\langle u, v\rangle + \|v\|^2,$$

当 $u \perp v$ 时交叉项消失，得到第二个等式。 ■

图 13.1 中，蓝色向量 u 与橙色向量 v 张成一个平行四边形，青色对角线表示 $u+v$。

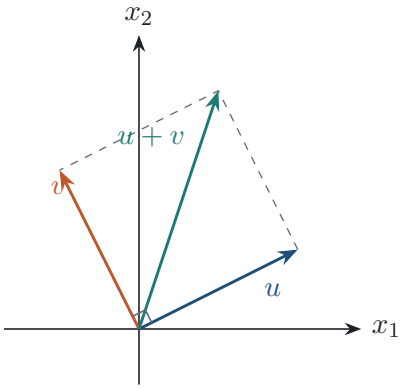


图 13.1　蓝色 $u = (2,1)$ 与橙色 $v = (-1,2)$ 正交；青色对角线 $u+v = (1,3)$ 的长度平方为 $10 = 5+5$。

从图中可见，对角线把平行四边形分成两个全等三角形。两条从原点出发的边垂直，故任一三角形的斜边长度平方等于两条直角边长度平方之和。这是向量版勾股定理。

■ **例 13.4　用勾股等式检验正交**　设 $u = (2,1)$、$v = (-1,2)$。用长度平方检验它们是否正交，并求 $\|u+v\|$。

解　先算 $\|u\|^2=5$、$\|v\|^2=5$、$u+v=(1,3)$ 和 $\|u+v\|^2=10$。因为 $10=5+5$，由定理 13.6 得 $u\perp v$，且 $\|u+v\|=\sqrt{10}$。这个检验和直接算 $\langle u,v\rangle=0$ 一致。 ♦

> **定理 13.7（非零两两正交向量组线性无关）**　若 $v_1,\dots,v_k$ 都非零，且 $i\neq j$ 时 $v_i\perp v_j$，则 $v_1,\dots,v_k$ 线性无关。

证明　设 $c_1v_1+\cdots+c_kv_k=0$。固定 j，两边与 v_j 取点积。由两两正交，左边只剩 $c_j\|v_j\|^2$，所以 $c_j\|v_j\|^2=0$。因 $v_j\neq0$，有 $\|v_j\|^2>0$，故 $c_j=0$。对每个 j 都成立，因此所有系数为零。 ■

例如 $v_1=(1,1,0)$、$v_2=(1,-1,0)$、$v_3=(0,0,2)$ 两两正交且都非零，故线性无关。逐对点积分别为 $1-1=0$、0、0。

> **易错提醒**　「两两正交」本身不能推出向量组线性无关。任取非零向量 v，向量组 $\{0,v\}$ 两两正交，却因含有零向量而线性相关。定理中的「每个向量非零」不能删去。

13.3　一个子空间的正交补是什么？

在 $\mathbb{R}^3$ 中，方程 $x_1+2x_2+2x_3=0$ 的解构成一个过原点的平面。每个解向量都与法向量 $(1,2,2)$ 点积为零，因此这个平面垂直于由该法向量张成的直线。

> **定义 13.8（正交补）**　设 W 是 $\mathbb{R}^n$ 的子空间。W 的**正交补**（orthogonal complement）定义为
>
> $$W^{\perp}=\{\,x\in\mathbb{R}^n\mid \langle x,w\rangle=0\ \text{对每个}\ w\in W\,\}.$$

定义中的「每个」不能换成「某个」。例如 $\mathbb{R}^2$ 中的一条直线只和与它垂直的直线构成正交子空间；它与整个平面不正交，因为平面里含有直线上方向的向量。零子空间的正交补是整个 $\mathbb{R}^n$，而 $\mathbb{R}^n$ 的正交补只有零向量。

两个子空间正交，要求其中任意一个向量都与另一个子空间里的每个向量正交。例如，令

$$S=\operatorname{span}\{(1,0,0),(0,1,0)\},\qquad T=\operatorname{span}\{(0,1,0),(0,0,1)\}.$$

它们分别是 xy 平面和 yz 平面。两平面都含有 $e_2=(0,1,0)$，而 $e_2^{\mathsf T}e_2=1$，所以不正交。虽然 $e_1\in S$ 与 $e_3\in T$ 正交，但只找到这样一对向量还不够。一般地，若两个子空间有非零交集，取交集中的非零向量 v，正交性会要求 $v^{\mathsf T}v=0$，与 $v\neq0$ 矛盾。

实际检验两个子空间是否正交，只需检查它们的基向量。令 $\{u_i\}_{i=1}^{p}$、$\{v_j\}_{j=1}^{q}$ 分别为 U、V 的一组基。若每个 $\langle u_i,v_j\rangle=0$，那么任意 $u\in U$、$v\in V$ 都可写为

$$u=\sum_i a_iu_i,\qquad v=\sum_j b_jv_j,$$

并满足

$$\langle u,v\rangle=\sum_i\sum_j a_ib_j\langle u_i,v_j\rangle=0.$$

这把定义中对无穷多个向量的检查，化为有限次基向量点积。

定理 13.9（正交补的维数）　若 $W \subseteq \mathbb{R}^n$ 且 $\dim W = k$，则 $W^\perp$ 是子空间，且

$$\dim W^\perp = n - k.$$

思路　先用点积的线性性验证子空间条件，再把 $W^\perp$ 写成一个齐次线性系统的零空间。系统的行由 W 的一组基组成，因此有 k 个独立方程；秩—零度定理给出维数。

证明　先证 $W^\perp$ 是子空间。零向量与 W 中每个向量点积为零。若 $x, y \in W^\perp$，$a, b \in \mathbb{R}$，则对每个 $w \in W$，

$$\langle ax + by, w\rangle = a\langle x, w\rangle + b\langle y, w\rangle = 0,$$

所以 $ax + by \in W^\perp$。

取 W 的一组基 $w_1, \dots, w_k$，令矩阵 B 的第 i 行为 w_i^{T}。一个向量 x 属于 $W^\perp$，当且仅当 $w_i^{\mathsf{T}} x = 0$ 对每个 i 成立，也就是 $Bx = 0$。故 $W^\perp = N(B)$。矩阵 B 的 k 行线性无关，秩为 k；由秩—零度定理，$\dim W^\perp = n - k$。 ■

■ **例 13.5　求一个平面的正交补**　设 $W = \operatorname{span}\{(1, 2, 2)\} \subseteq \mathbb{R}^3$。求 $W^\perp$ 的一组基和维数。

解　若 $x = (x_1, x_2, x_3) \in W^\perp$，它须满足

$$x_1 + 2x_2 + 2x_3 = 0, \qquad x_1 = -2x_2 - 2x_3.$$

令 $x_2 = s$、$x_3 = t$，便有

$$x = s(-2, 1, 0) + t(-2, 0, 1).$$

两个生成向量都与 $(1, 2, 2)$ 点积为零，且彼此不成倍数，所以构成 $W^\perp$ 的一组基。故 $\dim W^\perp = 2 = 3 - \dim W$。 ♦

■ **例 13.6　平面的正交补是一条直线**　设 $W = \operatorname{span}\{(1, 1, 0), (0, 1, 1)\} \subseteq \mathbb{R}^3$。求 $W^\perp$ 的维数和一组基。

解　两个生成向量不成倍数，故 $\dim W = 2$。向量 $x = (x_1, x_2, x_3)$ 属于 $W^\perp$，当且仅当它与两个基向量都正交：

$$x_1 + x_2 = 0, \qquad x_2 + x_3 = 0.$$

令 $x_2 = -1$，得到 $x = (1, -1, 1)$；所有解都是它的倍数。两个约束独立，因此 $\dim W^\perp = 1 = 3 - \dim W$。代入可见 $(1, -1, 1)$ 与两个生成向量的点积均为 0。 ♦

想一想　为什么 $W^\perp$ 的维数会少于 $\mathbb{R}^n$ 的维数？
答　若 W 有 k 个独立方向，就有 k 条独立的齐次条件 $w_i^{\mathsf{T}} x = 0$。每条独立条件少一个自由度，故剩下 $n - k$ 个自由度。零空间边界情形给出 $\dim\{0\}^\perp = n$。

推论 13.10（双重正交补）　对 $\mathbb{R}^n$ 的任意子空间 W，都有 $(W^\perp)^\perp = W$。

证明　若 $w \in W$，它与每个 $W^\perp$ 中的向量都正交，所以 $W \subseteq (W^\perp)^\perp$。两边维数相同：右边维数为 $n-(n-k)=k=\dim W$。有限维子空间若包含且维数相等，便相等。 ■

推论 13.11（正交分解）　每个 $x \in \mathbb{R}^n$ 都能唯一写成 $x = w + z$，其中 $w \in W$、$z \in W^\perp$。也就是说，$\mathbb{R}^n = W + W^\perp$，且两部分的交只有零向量。

证明　若 $v \in W \cap W^\perp$，则 $\langle v, v\rangle = 0$，所以 $v = 0$。取 W 和 $W^\perp$ 的基，两组向量合在一起有 $k + (n-k) = n$ 个向量。它们线性无关：若两部分的组合相加为零，则一部分等于另一部分的相反数，落在交集中，只能是零；两组基的系数于是全为零。$\mathbb{R}^n$ 中 n 个线性无关向量构成一组基，因此张成整个 $\mathbb{R}^n$。基表示唯一，故分解唯一。 ■

■ **例 13.7　正交分解的具体计算**　取 $W = \operatorname{span}\{u\}$，其中 $u = (1,2,2)$，并把 $x = (1,2,3)$ 写成 $w + z$，要求 $w \in W$、$z \in W^\perp$。

解　因为 $w \in W$，可写成 $w = tu$。条件 $z = x - w \in W^\perp$ 等价于 $\langle x - tu, u\rangle = 0$，于是

$$t = \frac{\langle x, u\rangle}{\|u\|^2} = \frac{11}{9}, \qquad w = \left(\frac{11}{9}, \frac{22}{9}, \frac{22}{9}\right), \qquad z = \left(-\frac{2}{9}, -\frac{4}{9}, \frac{5}{9}\right).$$

直接核对 $\langle z, u\rangle = (-2 - 8 + 10)/9 = 0$，且 $w + z = x$。所以这确实是所需分解。前一个例题给出的 $W^\perp$ 基也能表示这个 z：

$$z = -\frac{4}{9}(-2,1,0) + \frac{5}{9}(-2,0,1).$$

分解定理还保证这样的 w, z 只有这一组；它们分别是沿 W 方向的部分和与 W 垂直的剩余部分。 ♦

这个分解还说明 w 是 W 中离 x 最近的向量。任取 $q \in W$，有

$$x - q = z + (w - q), \qquad z \perp (w - q), \qquad \|x - q\|^2 = \|z\|^2 + \|w - q\|^2 \geq \|z\|^2.$$

当且仅当 $q = w$ 时取到最小值。这给出下一讲「投影到子空间」的几何动机。

13.4　四个基本子空间怎样互相垂直？

设 A 是 $m \times n$ 矩阵。A 的所有行向量的线性组合构成**行空间**（row space）$C(A^{\mathsf{T}})$；它们属于 $\mathbb{R}^n$，而 $Ax = 0$ 的解也属于 $\mathbb{R}^n$。因此可以直接问：零空间中的向量是否垂直于每一行？A 的列向量张成**列空间**（column space）$C(A) \subseteq \mathbb{R}^m$；满足 $A^{\mathsf{T}} y = 0$ 的向量构成**左零空间**（left nullspace），也属于 $\mathbb{R}^m$。

定理 13.12（四个基本子空间的正交关系）　对任意实矩阵 $A \in \mathbb{R}^{m\times n}$，

$$C(A^{\mathsf{T}})^{\perp} = N(A), \qquad C(A)^{\perp} = N(A^{\mathsf{T}}).$$

若 $\operatorname{rank} A = r$，则两对空间的维数分别为 r 与 $n-r$，以及 r 与 $m-r$。

思路　先证两个等式：零空间方程逐行写开，就是与每一行的点积为零；左零空间同理对应每一列。维数部分再用行秩等于列秩和秩—零度定理。

证明　把 A 的行记作 $a_1^{\mathsf{T}}, \dots, a_m^{\mathsf{T}}$。向量 x 属于 $N(A)$，当且仅当

$$a_1^{\mathsf{T}}x = 0, \quad \dots, \quad a_m^{\mathsf{T}}x = 0.$$

这表示 x 与 A 的每一行正交，所以它与这些行的所有线性组合正交，即 $x \in C(A^{\mathsf{T}})^{\perp}$。反过来，若 $x \in C(A^{\mathsf{T}})^{\perp}$，它与每一行正交，以上各等式成立，故 $Ax = 0$。因此 $N(A) = C(A^{\mathsf{T}})^{\perp}$。

同样，$A^{\mathsf{T}}y = 0$ 当且仅当 y 与 A 的每一列正交，这恰好说明 $y \in C(A)^{\perp}$。故 $N(A^{\mathsf{T}}) = C(A)^{\perp}$。

设 $\operatorname{rank} A = r$。列空间维数是 r；行空间 $C(A^{\mathsf{T}})$ 的维数也是 r，因为行秩等于列秩。秩—零度定理给出 $\dim N(A) = n - r$ 与 $\dim N(A^{\mathsf{T}}) = m - r$，于是两对正交补的维数相加分别等于 n 和 m。 ■

图 13.2 把两个向量空间中的配对分开画出：左侧属于 $\mathbb{R}^n$，右侧属于 $\mathbb{R}^m$。

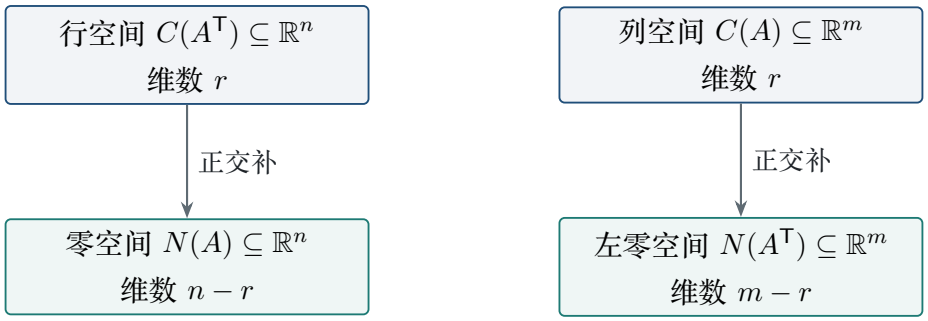


图 13.2　行空间与零空间互为正交补，列空间与左零空间也互为正交补；每对维数之和等于所在空间的维数。

从图中看，左边两空间的维数相加为 $r + (n-r) = n$，右边相加为 $r + (m-r) = m$。因此每一对正交子空间刚好互补，不会漏掉所在的整个空间。

■ 例 13.8　一行独立信息的四个子空间　设

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix}.$$

求它的行空间、零空间、列空间和左零空间的一组基及维数。

解　第二行是第一行的 2 倍，所以行空间的基可取 $(1, 2, 5)$，秩为 1。解 $x_1 + 2x_2 + 5x_3 = 0$，令

自由变量 x_2, x_3 分别取基方向，得到

$$N(A) = \operatorname{span}\{(-2,1,0),(-5,0,1)\}, \qquad \dim N(A) = 2.$$

回代可核对两个基向量：$1(-2)+2(1)+5(0)=0$，$1(-5)+2(0)+5(1)=0$；它们不成倍数，所以确实张成整个二维零空间。列向量分别是 $(1,2)$、$2(1,2)$、$5(1,2)$，故列空间基为 $(1,2)$，维数为 1。左零空间方程为 $y_1+2y_2=0$，所以基可取 $(-2,1)$，维数也为 1。两组正交关系可直接检查：$(1,2,5)^{\mathsf{T}}(-2,1,0)=0$，$(1,2)^{\mathsf{T}}(-2,1)=0$。四个维数依次是 $1,2,1,1$。 ♦

这里 $(-2,1)\in N(A^{\mathsf{T}})$ 还记录了两行的依赖关系：$-2(1,2,5)+(2,4,10)=0$。一般地，$N(A)$ 中的系数描述列向量的线性组合何时为零；$N(A^{\mathsf{T}})$ 中的系数描述行向量之间的线性关系。行系数有 m 个，所以左零空间属于 $\mathbb{R}^m$。

易错提醒　由 $Ax=0$ 得到的是 x 与 A 的行空间正交，不是与列空间正交。以刚才的 2×3 矩阵为例，$x\in\mathbb{R}^3$，列空间却在 $\mathbb{R}^2$ 中，两者中的向量不能直接取点积。

若 $Ax=b$ 无解，如何用一个短等式说明矛盾？行运算会发现矛盾行；左零空间给出同一件事的向量写法。

定理 13.13（Fredholm 择一定理）　对 $A\in\mathbb{R}^{m\times n}$ 和 $b\in\mathbb{R}^m$，下面两件事恰有一件成立：

(1) 存在 $x\in\mathbb{R}^n$ 使 $Ax=b$；

(2) 存在 $y\in\mathbb{R}^m$ 使 $A^{\mathsf{T}}y=0$ 且 $y^{\mathsf{T}}b=1$。

第二种情形中的 y 称为方程组无解的证书。

思路　两种情形不能同时发生，因为左乘 y^{T} 会把 $Ax=b$ 变成 $0=1$。若 b 不在列空间，则双重正交补定理保证它不可能同时垂直于整个左零空间；找到一个点积非零的左零空间向量后，再缩放到点积为 1。这个结论称为**Fredholm 择一定理**（Fredholm alternative）。

证明　**至多一种成立。**若 $Ax=b$ 且 $A^{\mathsf{T}}y=0$、$y^{\mathsf{T}}b=1$，则

$$1 = y^{\mathsf{T}}b = y^{\mathsf{T}}Ax = (A^{\mathsf{T}}y)^{\mathsf{T}}x = 0,$$

矛盾。

至少一种成立。若第一种不成立，则 $b\notin C(A)$。由定理 13.12 和推论 13.10，

$$C(A) = \big(C(A)^{\perp}\big)^{\perp} = N(A^{\mathsf{T}})^{\perp}.$$

因此存在 $z\in N(A^{\mathsf{T}})$ 使 $z^{\mathsf{T}}b\neq 0$；否则 b 会属于 $N(A^{\mathsf{T}})^{\perp}=C(A)$。令 $y=z/(z^{\mathsf{T}}b)$，则 $A^{\mathsf{T}}y=0$ 且 $y^{\mathsf{T}}b=1$。第二种成立。 ■

■ **例 13.9　用左零空间指出矛盾**　考虑方程组

$$x_1-x_2=1, \qquad x_2-x_3=1, \qquad x_1-x_3=1.$$

求一组系数，把三条方程相加后得到矛盾。

解　按系数 $y=(1,1,-1)$ 把第一式加第二式再减第三式。左边化为

$$(x_1-x_2)+(x_2-x_3)-(x_1-x_3)=0,$$

右边为 $1+1-1=1$，所以方程组会要求 $0=1$，不可能有解。对应地，$A^{\mathsf T}y=0$ 而 $y^{\mathsf T}b=1$；这组系数就是 Fredholm 择一定理给出的无解证书。 ♦

■ **例 13.10　左零空间辨别相容的右端项**　令 $A=\begin{bmatrix}1\\2\\3\end{bmatrix}$，并取 $y=(-1,-1,1)$。先检查 $b_1=(2,4,6)$ 是否与 $b_2=(2,4,7)$ 相容。

解　有 $A^{\mathsf T}y=-1-2+3=0$，所以 y 属于左零空间。对 b_1，取 $x=2$ 就有 $Ax=b_1$，且 $y^{\mathsf T}b_1=-2-4+6=0$。对 b_2，则 $y^{\mathsf T}b_2=-2-4+7=1$；若方程有解，左乘 $y^{\mathsf T}$ 会得到 $0=1$，故 $Ax=b_2$ 无解。这个例子也说明：相容的右端项必须与整个左零空间正交，而非相容时至少有一个左零空间向量能给出非零点积。 ♦

13.5　为什么 $A^{\mathsf T}A$ 记住了零空间?

测量数据含有误差时，$Ax=b$ 可能没有精确解。即使 A 是长方矩阵，$A^{\mathsf T}A$ 仍是方阵；它会出现在后续最小二乘方程中。先看一个不依赖最小二乘理论的事实：$A^{\mathsf T}A$ 不会增加或减少零空间中的向量。

若 A 的列为 $a_1,\dots,a_n$，矩阵乘法给出

$$(A^{\mathsf T}A)_{ij}=a_i^{\mathsf T}a_j=\langle a_i,a_j\rangle.$$

因此对角线上的数是各列长度的平方，非对角线上的数记录两列的点积；交换 i,j 不改变点积，所以 $A^{\mathsf T}A$ 对称。以上面稍后要用的 A 为例，它的两列是 $(1,1,1)$ 与 $(1,2,5)$，列间点积依次为 $3,8,30$，故 $A^{\mathsf T}A=\begin{bmatrix}3&8\\8&30\end{bmatrix}$。这个矩阵把列向量之间的长度和夹角信息记录在一个 $n\times n$ 方阵中。

对任意 $x\in\mathbb{R}^n$，还有

$$x^{\mathsf T}A^{\mathsf T}Ax=(Ax)^{\mathsf T}(Ax)=\|Ax\|^2\ge 0.$$

因此这个二次型不会取负值；它等于零恰好表示 $Ax=0$。若 A 的列线性无关，非零 x 不可能使 $Ax=0$，这个量便对每个非零 x 都严格为正。下面的零空间定理把这个观察转成可逆性的判据。

> **定理 13.14（$A^{\mathsf T}A$ 与 A 有相同的零空间）**　对任意实矩阵 $A\in\mathbb{R}^{m\times n}$，有
>
> $$N(A^{\mathsf T}A)=N(A).$$

思路　若 $Ax=0$，立刻有 $A^{\mathsf T}Ax=0$。反方向从 $A^{\mathsf T}Ax=0$ 出发，在等式左边再乘 $x^{\mathsf T}$；得到 Ax 的长度平方为零，只能有 $Ax=0$。

证明　若 $Ax=0$，则 $A^{\mathsf T}Ax=0$，故 $N(A)\subseteq N(A^{\mathsf T}A)$。反过来，若 $A^{\mathsf T}Ax=0$，则

$$0=x^{\mathsf T}A^{\mathsf T}Ax=(Ax)^{\mathsf T}(Ax)=\|Ax\|^2.$$

长度平方为零只可能发生在零向量上，所以 $Ax=0$，即 $x\in N(A)$。两边包含关系合起来得到零空间相等。 ■

> **推论 13.15（$A^{\mathsf T}A$ 的秩与可逆性）**　对 $A\in\mathbb{R}^{m\times n}$，有 $\operatorname{rank}(A^{\mathsf T}A)=\operatorname{rank}A$。此外，$A^{\mathsf T}A$ 可逆当且仅当 A 的列线性无关。

证明　两个矩阵有相同的零空间，且都有 n 列；由秩—零度定理，它们的秩相同。$A^{\mathsf T}A$ 是 $n\times n$ 方阵，可逆当且仅当其零空间只有零向量；由定理 13.14，这等价于 $Ax=0$ 只有零解，也就是 A 的列线性无关。 ■

例 13.11　满列秩与相关列的对照　先看

$$A=\begin{bmatrix}1&1\\1&2\\1&5\end{bmatrix}.$$

计算 $A^{\mathsf T}A$，并判断它是否可逆。再把两列改成相关的矩阵

$$D=\begin{bmatrix}1&2\\2&4\end{bmatrix},$$

检查 $D^{\mathsf T}D$ 的零空间。

解　矩阵 A 的两列不成倍数，因而线性无关；直接相乘得

$$A^{\mathsf T}A=\begin{bmatrix}3&8\\8&30\end{bmatrix},\qquad \det(A^{\mathsf T}A)=3\cdot30-8\cdot8=26\neq0.$$

对两列 a_1,a_2，这个行列式可写成

$$\det(A^{\mathsf T}A)=\|a_1\|^2\|a_2\|^2-\langle a_1,a_2\rangle^2.$$

柯西–施瓦茨保证它非负；等号恰在两列相关时成立。本例的两列不相关，所以行列式严格为正。所以 $A^{\mathsf T}A$ 可逆。矩阵 D 的第二列是第一列的 2 倍，并且

$$D^{\mathsf T}D=\begin{bmatrix}5&10\\10&20\end{bmatrix}.$$

向量 $(-2,1)$ 同时满足 $D(-2,1)=0$ 和 $D^{\mathsf T}D(-2,1)=0$，故两者的零空间都是 $\operatorname{span}\{(-2,1)\}$，$D^{\mathsf T}D$ 不可逆。 ♦

> **易错提醒**　「$A^{\mathsf T}A$ 总是可逆」是错的。上面的 D 有相关列，$D^{\mathsf T}D$ 的零空间含有非零向量 $(-2,1)$。正确判据是：$A^{\mathsf T}A$ 可逆当且仅当 A 的列线性无关。

想一想　若 A 不是方阵，$A^{\mathsf{T}}A$ 仍可能可逆吗?

答　可以。上面的 A 是 3×2 矩阵，列线性无关；所以 $A^{\mathsf{T}}A$ 是可逆的 2×2 方阵。可逆性取决于列是否线性无关，不要求 A 本身是方阵。

方法：检查 $A^{\mathsf{T}}A$　判断 $A^{\mathsf{T}}A$ 是否可逆，不必先算逆矩阵或行列式；先问 $Ax=0$ 是否只有零解。若只有零解，A 列线性无关，$A^{\mathsf{T}}A$ 可逆；若有非零解，它也会落在 $N(A^{\mathsf{T}}A)$ 中。

前后联系　下一讲会把「误差向量垂直于列空间」转成投影计算。本讲的正交关系也会说明，最小二乘误差满足 $A^{\mathsf{T}}e=0$。

本讲小结

- 点积 $\langle x,y\rangle = x^{\mathsf{T}}y$；长度平方 $\|x\|^2 = x^{\mathsf{T}}x$；柯西–施瓦茨给出 $|x^{\mathsf{T}}y|\le\|x\|\|y\|$。
- $u\perp v \iff \|u+v\|^2=\|u\|^2+\|v\|^2$；此时 $\|u-v\|^2$ 也等于两长度平方之和。
- 非零且两两正交的向量组线性无关；零向量虽与所有向量正交，却不能放进这样的线性无关组。
- 若 $\dim W=k$，则 $W^{\perp}$ 是子空间，$\dim W^{\perp}=n-k$，并且 $(W^{\perp})^{\perp}=W$。
- $N(A)=C(A^{\mathsf{T}})^{\perp}$，$N(A^{\mathsf{T}})=C(A)^{\perp}$；两对维数分别相加为 n 与 m。
- $Ax=b$ 无解当且仅当存在 y 使 $A^{\mathsf{T}}y=0$、$y^{\mathsf{T}}b=1$；并且 $N(A^{\mathsf{T}}A)=N(A)$、$\operatorname{rank}(A^{\mathsf{T}}A)=\operatorname{rank}A$。

练习

13.1 ★　令 $u=(3,-2,6)$、$v=(1,4,0)$。求 $u^{\mathsf{T}}v$、$\|u\|^2$、$\|v\|^2$，并核对柯西–施瓦茨不等式。

13.2 ★　令 $u=(3,1)$、$v=(-1,3)$。验证 $u\perp v$，并用勾股等式计算 $\|u+v\|$。

13.3 ★　判断真伪并说明理由：只要存在一对非零向量 $s\in S, t\in T$ 满足 $s\perp t$，两个子空间 S,T 就正交。

13.4 ★★　证明：非零两两正交的向量组线性无关。指出「非零」条件不能删除的原因。

13.5 ★★　设 $W=\operatorname{span}\{(1,2,2)\}\subseteq\mathbb{R}^3$。求 $W^{\perp}$ 的方程、一组基和维数。

13.6 ★★　对矩阵 $B=\begin{bmatrix}1&2&3\\2&4&6\\1&1&1\end{bmatrix}$，求四个基本子空间的一组基及维数。

13.7 ★★　设 $D=\begin{bmatrix}1&2\\2&4\end{bmatrix}$。计算 $D^{\mathsf{T}}D$，求 $N(D)$，并说明 $D^{\mathsf{T}}D$ 是否可逆。

13.8 ★★★　对方程组 $x_1-x_2=1$、$x_2-x_3=1$、$x_1-x_3=1$，找出 Fredholm 无解证书，并说明证书如何导出矛盾。

13.9 ★★★ 在 $\mathbb{R}^2$ 中给定非零向量 r, n, c, ℓ。写出它们分别成为 $C(A^{\mathsf{T}}), N(A), C(A), N(A^{\mathsf{T}})$ 的基的充要条件，并用 $r=(1,2)$、$n=(-2,1)$、$c=(2,1)$、$\ell=(1,-2)$ 构造一个矩阵 A。

练习答案与提示

13.1 $u^{\mathsf{T}}v = 3-8 = -5$，$\|u\|^2 = 9+4+36 = 49$，$\|v\|^2 = 1+16 = 17$。于是 $|u^{\mathsf{T}}v|^2 = 25 \le 49\cdot 17 = 833$。

13.2 $u^{\mathsf{T}}v = -3+3 = 0$。有 $u+v = (2,4)$，$\|u+v\|^2 = 20$；而 $\|u\|^2+\|v\|^2 = 10+10 = 20$，所以 $\|u+v\| = \sqrt{20} = 2\sqrt{5}$。

13.3 假。取 $S = T = \mathbb{R}^2$，其中 $(1,0)\perp(0,1)$，但 $(1,0)$ 与自身点积为 1，故 S, T 并非正交子空间。正交子空间要求每一对向量都正交。

13.4 若 $\sum_i c_i v_i = 0$，固定 j 并与 v_j 取点积，得 $c_j\|v_j\|^2 = 0$。因 $v_j \neq 0$，所以 $c_j = 0$；对所有 j 均成立，故向量组线性无关。若允许零向量，集合 $\{0, v\}$ 两两正交却线性相关。

13.5 $W^{\perp}$ 满足 $x_1+2x_2+2x_3 = 0$。令 $x_2 = s, x_3 = t$，则 $x = s(-2,1,0)+t(-2,0,1)$。因此一组基是 $\{(-2,1,0),(-2,0,1)\}$，维数为 2。

13.6 B 的第二行是第一行的两倍，第三行与第一行独立，所以行空间基可取 $(1,2,3),(1,1,1)$，维数为 2。列空间基可取第一、三列 $(1,2,1),(3,6,1)$，也有维数 2。直接代入得 $N(B) = \operatorname{span}\{(1,-2,1)\}$、$N(B^{\mathsf{T}}) = \operatorname{span}\{(-2,1,0)\}$，两者维数均为 1。

13.7 $D^{\mathsf{T}}D = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix}$。因 $D(-2,1) = 0$ 且 D 的秩为 1，$N(D) = \operatorname{span}\{(-2,1)\}$。由 $N(D^{\mathsf{T}}D) = N(D)$，$D^{\mathsf{T}}D$ 也有非零零空间，故不可逆。

13.8 取 $y = (1,1,-1)$。三式按 $1,1,-1$ 加权后，左边是 $(x_1-x_2)+(x_2-x_3)-(x_1-x_3) = 0$，右边是 $1+1-1 = 1$，产生矛盾。也就是 $A^{\mathsf{T}}y = 0$、$y^{\mathsf{T}}b = 1$。

13.9 充要条件是 $r^{\mathsf{T}}n = 0$ 且 $c^{\mathsf{T}}\ell = 0$；由于四个向量非零，每一对正交向量就在 $\mathbb{R}^2$ 中张成互为正交补的一维空间。取 $A = cr^{\mathsf{T}}$，得到 $A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$。它的行空间是 $\operatorname{span}\{r\}$，列空间是 $\operatorname{span}\{c\}$；其零空间与左零空间分别是 $\operatorname{span}\{n\}$、$\operatorname{span}\{\ell\}$。

第 14 讲

投影到子空间

Projections onto Subspaces

对应材料：MIT 18.06SC《Projections onto subspaces》课时摘要；习题 15.1–15.2；Strang《Introduction to Linear Algebra》第 4.2 节及 ZoomNotes Part 4 第 4.2 节

本讲目标

(1) 会用正交条件求向量到直线或子空间的最近点，并核对残差。

(2) 理解勾股分解为何保证投影点存在、唯一且最近。

(3) 会构造投影矩阵，并说明 $P^2 = P$ 与 $P^{\mathsf{T}} = P$ 分别表达什么。

(4) 会在基向量线性无关时求 $C(A)$ 上的投影，并推导法方程。

(5) 能区分投影向量的唯一性与坐标系数的唯一性。

14.1 投影点为什么是最近点？

在平面中取方向向量 $a = (2,1)$，并给定 $b = (3,4)$。向量 $p = 2a = (4,2)$ 落在直线 $L = \operatorname{span}\{a\}$ 上，剩下的误差是 $e = b - p = (-1,2)$。直接计算得到 $a^{\mathsf{T}}e = 2(-1) + 1(2) = 0$，所以这段误差垂直于直线。为什么这就说明 p 是 L 上离 b 最近的点？

定义 14.1（正交投影） 设 $S \subseteq \mathbb{R}^m$ 是子空间，$b \in \mathbb{R}^m$。若 $p \in S$ 且 $b - p \in S^{\perp}$，就称 p 是 b 到 S 的正交投影（orthogonal projection），记作 $\operatorname{proj}_S b$。

例如上面的 $p = (4,2)$ 满足 $p \in L$ 且 $e \perp L$。若 $b \in S$，则投影就是 b；若 $S = \{0\}$，投影只能是 0；若 $S = \mathbb{R}^m$，投影仍是 b。这三个边界情形分别对应零误差、只有零向量可选和整个空间都可选。

定理 14.2（正交误差刻画最近点） 设 $S \subseteq \mathbb{R}^m$ 是子空间，$b \in \mathbb{R}^m$。存在唯一的 $p \in S$ 使 $\|b - p\|$ 最小；它满足

$$b - p \in S^{\perp}.$$

反过来，若 $p \in S$ 且 $b - p \in S^{\perp}$，则 p 就是这个唯一最近点。

思路　第 13 讲的正交分解定理先给出 $b = p + e$，其中 $p \in S$、$e \in S^{\perp}$。勾股等式说明这个 p 最

近；反方向可沿 S 中任意方向微调 p，若误差仍有平行分量，距离就还能减小。

证明　由第 13 讲的正交分解定理，唯一写成 $b = p + e$，其中 $p \in S$、$e \in S^{\perp}$。对任意 $q \in S$，$p - q \in S$，所以 $e \perp (p - q)$，从而

$$\|b - q\|^2 = \|e + (p - q)\|^2 = \|e\|^2 + \|p - q\|^2 \geq \|e\|^2.$$

等号当且仅当 $q = p$，故 p 是唯一最近点。

反过来，设 p 是最近点，记 $e = b - p$。若存在 $v \in S$ 使 $\langle e, v\rangle \neq 0$，则 $v \neq 0$，并且 $p + tv \in S$ 对每个实数 t 都成立。取 $t = \langle e, v\rangle / \|v\|^2$，便有

$$\|b - (p + tv)\|^2 = \|e - tv\|^2 = \|e\|^2 - \frac{\langle e, v\rangle^2}{\|v\|^2} < \|e\|^2,$$

这与 p 最近矛盾。因此 e 与 S 中每个向量正交，即 $b - p \in S^{\perp}$。 ■

■ **例 14.1**　**比较直线上的两个候选点**　对 $a = (2, 1)$、$b = (3, 4)$，验证 $p = 2a$ 是 b 到 $L = \operatorname{span}\{a\}$ 的投影，并与另一个候选点 $q = 3a$ 比较距离。

解　用正交误差条件检验 p，再用距离平方比较。$p = (4, 2)$，所以 $b - p = (-1, 2)$，并且 $a^{\mathsf{T}}(b - p) = 0$。$q = (6, 3)$ 也在 L 上，但 $b - q = (-3, 1)$。

$$\|b - p\|^2 = (-1)^2 + 2^2 = 5, \qquad \|b - q\|^2 = (-3)^2 + 1^2 = 10.$$

因此 p 比 q 更近；定理 14.2 说明 L 上没有比 p 更近的点。 ♦

还可以用夹角读出投影长度。若 θ 是 a 与 b 的夹角，则

$$\cos\theta = \frac{a^{\mathsf{T}} b}{\|a\|\|b\|} = \frac{10}{\sqrt{5}\cdot 5} = \frac{2}{\sqrt{5}}, \qquad \|p\| = 2\sqrt{5} = \|b\|\cos\theta.$$

因此点积公式不仅给出投影的方向，也给出 b 沿直线方向的有符号长度；当 $a^{\mathsf{T}} b < 0$ 时，系数为负，投影落在 a 的反方向。

图 14.1 把刚才的向量关系画在同一坐标系里：橙色箭头是 b，蓝色箭头是投影 p，灰色箭头从 p 指向 b，表示误差 $e = b - p$。

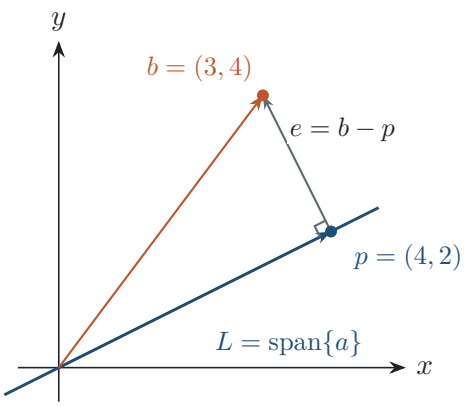


图 14.1　向量 b 分解为沿直线 L 的投影 p 与垂直于 L 的误差 e，直角记号标出 $p \perp e$。

投影点 p 必须落在 L 上，而误差箭头从 p 出发并与 L 垂直。若沿着 L 把 p 移到另一个点 q，新的误差由原来的垂直分量 e 和沿 L 的位移 $p-q$ 组成；两者互相垂直，所以新的距离平方会多出 $\|p-q\|^2$。

想一想　如果把直线方向向量 a 换成 $3a$，投影点会变吗?

答　不会，因为 $\mathrm{span}\{a\}=\mathrm{span}\{3a\}$，投影由子空间本身决定，与用哪个非零向量表示方向无关。下一节从点积条件推导可直接计算的公式。

14.2　怎样计算投影到一条直线?

设 $L=\mathrm{span}\{a\}$，其中 $a\neq 0$。投影 p 一定有形式 $p=xa$；误差垂直于 a，所以 $a^{\mathsf{T}}(b-xa)=0$。这条线性方程决定唯一的系数 x。

定理 14.3（投影到直线的公式）　设 $a\in\mathbb{R}^m$ 且 $a\neq 0$，$b\in\mathbb{R}^m$。则 b 到 $L=\mathrm{span}\{a\}$ 的正交投影为

$$p=\mathrm{proj}_L\, b=\frac{a^{\mathsf{T}}b}{a^{\mathsf{T}}a}\,a,\qquad e=b-p\in L^{\perp}.$$

直观理解　分子 $a^{\mathsf{T}}b$ 衡量 b 沿 a 方向的分量，分母 $a^{\mathsf{T}}a=\|a\|^2$ 把方向向量的长度校正回来。若 b 已沿着 a，公式会把它原样保留；与 a 垂直的分量则落进误差 e。

思路　令 $p=xa$，先用 $a^{\mathsf{T}}(b-p)=0$ 解出 x，再用定理 14.2 验证这个 p 正是唯一最近点。条件 $a\neq 0$ 保证 $a^{\mathsf{T}}a=\|a\|^2>0$，因此除法合法。

证明　写 $p=xa$。由 $b-p\perp L$ 必须有

$$0=a^{\mathsf{T}}(b-xa)=a^{\mathsf{T}}b-x(a^{\mathsf{T}}a),$$

所以 $x=(a^{\mathsf{T}}b)/(a^{\mathsf{T}}a)$。因为 $a\neq 0$，有 $a^{\mathsf{T}}a=\|a\|^2>0$，这个系数存在且唯一。代回可得 $a^{\mathsf{T}}(b-p)=0$；L 中每个向量都是 a 的倍数，故 $b-p\perp L$。定理 14.2 于是给出唯一最近点。 ■

条件 $a\neq 0$ 不能去掉：若 $a=0$，分母 $a^{\mathsf{T}}a=0$，公式没有定义；此时 $L=\{0\}$，投影仍存在，只是应直接取 $p=0$。

把公式写成矩阵乘法，可以一次作用于任意输入 b：

$$P_a=\frac{aa^{\mathsf{T}}}{a^{\mathsf{T}}a},\qquad p=P_a b.$$

这里 aa^{T} 是 $m\times m$ 矩阵，P_a 称为到直线 L 的投影矩阵（projection matrix）。转置后

$$(aa^{\mathsf{T}})_{ij}=a_i a_j.$$

因此 P_a 的第 j 列为 $(a_j/(a^{\mathsf{T}}a))a$，也是 a 的倍数。这说明 aa^{T} 是外积矩阵，而 $a^{\mathsf{T}}a$ 才是标量点积；把二者的位置写反会造成维数不匹配，也会丢掉「输出沿 a 方向」的结构。若要同时投影多

个输入，可把它们作为矩阵 B 的列，一次计算 P_aB；结果的每一列就是对应输入的投影。这种写法适合直线方向固定、需要比较多组输入的情形：投影空间不变，只需更换输入列。输入列之间是否独立，不影响按列投影的正确性。

$$P_a^{\mathsf{T}} = \frac{(aa^{\mathsf{T}})^{\mathsf{T}}}{a^{\mathsf{T}}a} = \frac{aa^{\mathsf{T}}}{a^{\mathsf{T}}a} = P_a.$$

而

$$P_a^2 = \frac{aa^{\mathsf{T}}aa^{\mathsf{T}}}{(a^{\mathsf{T}}a)^2} = \frac{a(a^{\mathsf{T}}a)a^{\mathsf{T}}}{(a^{\mathsf{T}}a)^2} = P_a.$$

前者表示矩阵对称，后者表示投影后的向量再投一次不会移动。并且 $P_ab \in L$，而 $P_ay = y$ 对每个 $y \in L$ 成立，所以 P_a 的像恰好是 L。

公式还说明投影对输入是线性的：$P_a(\alpha b_1 + \beta b_2) = \alpha P_a b_1 + \beta P_a b_2$。特别地，放大输入会按同一比例放大投影；若给输入增加一个垂直于 a 的向量 z，则 $P_a(b+z) = P_ab$，因为 $a^{\mathsf{T}}z = 0$。因此投影只读取沿直线方向的分量，垂直分量全部留在残差中。

若 $P_ab = 0$，则 $a^{\mathsf{T}}b = 0$；反过来，$a^{\mathsf{T}}b = 0$ 也会使 $P_ab = 0$。因此 $N(P_a) = L^{\perp}$，而 $\operatorname{rank}(P_a) = 1$。同时，$(I - P_a)b = b - P_ab$ 是误差，属于 $L^{\perp}$；若 $y \in L^{\perp}$，则 $(I - P_a)y = y$。所以 $I - P_a$ 正交投到直线的正交补。

■ **例 14.2**　取 $a = (1,2,2)$、$b = (1,1,1)$。求 b 到 $L = \operatorname{span}\{a\}$ 的投影和误差，并检验勾股分解。

解　由直线投影公式，$a^{\mathsf{T}}b = 5$、$a^{\mathsf{T}}a = 9$，因此

$$p = \frac{5}{9}(1,2,2) = \left(\frac{5}{9}, \frac{10}{9}, \frac{10}{9}\right), \qquad e = b - p = \left(\frac{4}{9}, -\frac{1}{9}, -\frac{1}{9}\right).$$

验算 $a^{\mathsf{T}}e = (4-2-2)/9 = 0$，且 $\|p\|^2 = 25/9$、$\|e\|^2 = 2/9$，所以 $\|b\|^2 = 3 = 25/9 + 2/9$。误差给出了垂直于直线的分量。对应的投影矩阵是

$$P_a = \frac{1}{9}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}, \qquad P_ab = p.$$

♦

■ **例 14.3　投影系数可以为负**　令 $a = (1,1)$、$b = (-2,1)$。求 b 到 $\operatorname{span}\{a\}$ 的投影，并解释投影系数的符号。

解　先用点积公式计算系数：$a^{\mathsf{T}}b = -1$、$a^{\mathsf{T}}a = 2$，所以 $x = -1/2$。因此

$$p = -\frac{1}{2}(1,1) = (-1/2, -1/2), \qquad e = b - p = (-3/2, 3/2).$$

验算 $a^{\mathsf{T}}e = 0$，并且 $\|b\|^2 = 5 = 1/2 + 9/2 = \|p\|^2 + \|e\|^2$。系数为负表示投影点落在 a 的反方向；它仍然属于同一条直线，且仍是离 b 最近的点。♦

■ **例 14.4　投影到三维中的一条对角线**　令 $a = (1,1,1)^{\mathsf{T}}$、$b = (6,0,0)^{\mathsf{T}}$。求 b 到 $L = \operatorname{span}\{a\}$ 的投影与残差。

解　由 $a^{\mathsf{T}}b=6$、$a^{\mathsf{T}}a=3$，投影系数为 2，所以

$$p=2a=(2,2,2)^{\mathsf{T}},\qquad e=b-p=(4,-2,-2)^{\mathsf{T}}.$$

验证 $a^{\mathsf{T}}e=4-2-2=0$；并且 $\|p\|^2=12$、$\|e\|^2=24$，故 $\|b\|^2=36=12+24$。点 p 是沿三个坐标方向均匀分量组成的对角线上离 b 最近的点，残差则在这条线的正交补中。 ♦

方法：投影到直线　遇到 b 到 $\mathrm{span}\{a\}$ 的最近点问题，先令 $p=xa$，再用 $a^{\mathsf{T}}(b-p)=0$ 解 x；最后检查残差垂直条件。

14.3　怎样投影到列空间?

在三维空间里，两条不平行的直线方向张成一个过原点的平面。把这两个方向作为矩阵 A 的列，平面就是 $C(A)$；要在平面内找最接近 b 的点，只需让误差垂直于两列。

定理 14.4（列空间投影与法方程）　设 $A\in\mathbb{R}^{m\times n}$ 的列线性无关，$b\in\mathbb{R}^m$。则存在唯一的系数向量 $\hat{x}$，使 $p=A\hat{x}$ 是 b 到 $C(A)$ 的正交投影。它满足

$$A^{\mathsf{T}}A\hat{x}=A^{\mathsf{T}}b,\qquad \hat{x}=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}b,\qquad p=A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}b.$$

其中 $A^{\mathsf{T}}A\hat{x}=A^{\mathsf{T}}b$ 称为**法方程**（normal equations）。**残差**（residual）$e=b-p$ 满足 $A^{\mathsf{T}}e=0$，即 $e\in C(A)^{\perp}=N(A^{\mathsf{T}})$；这里 $N(A^{\mathsf{T}})$ 称为 A 的**左零空间**（left nullspace）。

第 13 讲的正交补维数公式还告诉我们，若 A 的秩为 r，残差所在空间 $N(A^{\mathsf{T}})$ 的维数是 $m-r$。列独立时 $r=n$，所以左零空间有 $m-n$ 个自由方向：方阵可逆时没有可留给残差的方向，每个 b 都能精确表示；若观测维数 m 大于模型参数数 n，正交补则为不相容数据提供了容纳误差的空间。三点拟合例中 $m=3,n=2$，因此所有可行残差只沿一个方向变化。

直观理解　每一列 a_j 都是平面内的方向。误差要垂直于整个列空间，恰好等价于 $a_j^{\mathsf{T}}e=0$ 对每一列都成立；把这些点积条件叠在一起就是 $A^{\mathsf{T}}e=0$。

思路　令 $p=A\hat{x}$，先把 $A^{\mathsf{T}}(b-A\hat{x})=0$ 改写成方程组。列独立保证 $A^{\mathsf{T}}A$ 可逆，故系数唯一；再用定理 14.2 将正交残差转成最近点结论。

证明　先说明 $A^{\mathsf{T}}A$ 可逆。若 $x\neq0$，列线性无关给出 $Ax\neq0$，于是

$$x^{\mathsf{T}}A^{\mathsf{T}}Ax=\|Ax\|^2>0.$$

因此 $A^{\mathsf{T}}A$ 的零空间只有零向量，作为方阵可逆。令 $\hat{x}=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}b$，并取 $p=A\hat{x}$、$e=b-p$。代入得

$$A^{\mathsf{T}}e=A^{\mathsf{T}}b-A^{\mathsf{T}}A\hat{x}=0.$$

所以 e 与 A 的每一列正交，也与其所有线性组合正交，即 $e\in C(A)^{\perp}$。定理 14.2 说明 p 是唯一最近点。列独立还保证 $A\hat{x}=p$ 的系数 $\hat{x}$ 唯一。 ■

这个证明还给出一个可用于检查计算的恒等式。对任意 $x \in \mathbb{R}^n$，令 $\delta = x - \hat{x}$。由于 $b - Ax = e - A\delta$，其中 $e \in C(A)^\perp$ 而 $A\delta \in C(A)$，两项正交，故

$$\|b - Ax\|^2 = \|e\|^2 + \|A\delta\|^2.$$

因此改变系数只能在原有误差平方上增加 $\|A\delta\|^2$。列独立时，$\delta \neq 0$ 会推出 $A\delta \neq 0$，所以只有 $x = \hat{x}$ 达到最小值；这同时解释了投影点和系数向量为何都唯一。这里垂直的是残差 e 与列空间中的向量，特别是 $e \perp p$；一般不能说 $e \perp b$。由 $b = p + e$ 可得 $b^\mathsf{T} e = \|e\|^2$，除非残差为零，这个点积通常不为零。

设 A 是 $m \times n$ 矩阵，则 $\hat{x}$ 有 n 个分量，p 与 b 有 m 个分量，投影矩阵 $P_A = A(A^\mathsf{T}A)^{-1}A^\mathsf{T}$ 是 $m \times m$。按乘法次序，A^T 先把输入 $b \in \mathbb{R}^m$ 变成 n 个点积，$(A^\mathsf{T}A)^{-1}$ 在系数空间 $\mathbb{R}^n$ 中调整坐标，再由 A 把系数送回 $\mathbb{R}^m$。因此 P_A 的输入和输出维数相同。这里通常不能把 P_A 简化为 $AA^{-1}(A^\mathsf{T})^{-1}$：$A$ 可能不是方阵，也未必可逆。若 A 是可逆方阵，则 $C(A) = \mathbb{R}^m$，每个 b 已经在投影子空间中，$P_A = I$。

可以把这个乘法过程读成三步：先比较 b 与每个列方向，再根据列之间的夹角修正各方向的系数，最后把这些系数组合成列空间中的向量。若把输入改为 $b + z$，其中 $z \in N(A^\mathsf{T})$，第一步得到的点积完全不变，所以投影点也不变；变化的只有正交补中的残差。

行数与列数也给出一个快速判断：当 $n > m$ 时，A 有比环境维数更多的列，列必相关，因此不能使用含 $(A^\mathsf{T}A)^{-1}$ 的公式；最近点仍然存在且唯一，只是系数不一定唯一。当 $m > n$ 时，列独立是可能的，但不由形状自动保证；即使列独立，$C(A)$ 通常只是 $\mathbb{R}^m$ 的真子空间，b 仍可能落在它外面。

还可直接看出 P_A 的像是 $C(A)$：每个 $P_A b$ 都等于 A 乘某个系数向量；反过来，对 $y = Az \in C(A)$，

$$P_A y = A(A^\mathsf{T}A)^{-1}A^\mathsf{T}Az = Az = y.$$

所以 P_A 会保留列空间里的向量，并把其余方向送到列空间。这个结论也说明，投影矩阵虽然作用在 $\mathbb{R}^m$ 上，输出仍受 A 的列所张成的空间约束。

矩阵 $A^\mathsf{T}A$ 的第 (i, j) 项是两列 a_i 与 a_j 的点积。对角项记录每列长度的平方，非对角项记录不同列之间的夹角信息。若列互相垂直，这些非对角项为零，法方程中的系数可以逐个计算；列不垂直时，系数会通过这些项相互影响。

列两两正交时还可以把一般公式拆成若干条直线投影。若 $A = [a_1 \ \cdots \ a_n]$ 且每列非零、任意两列互相垂直，则 $A^\mathsf{T}A$ 是对角矩阵，因而

$$P_A = \sum_{j=1}^{n} \frac{a_j a_j^\mathsf{T}}{a_j^\mathsf{T} a_j}.$$

每一项只留下沿一个方向的分量，所有这些分量相加就是列空间中的投影。若列不正交，直接把各列方向上的投影相加一般会重复计算交叠的分量；这时需要保留法方程里的非对角项。若这些正交列还都已单位化，则 $A^\mathsf{T}A = I$，上式进一步成为 $P_A = AA^\mathsf{T}$。省略归一化因子需要列向量长度为 1；仅有互相垂直还不够。

例 14.5　投影到三维空间中的平面　令

$$A=\begin{bmatrix}1&0\\1&1\\1&2\end{bmatrix},\qquad b=(6,0,0)^{\mathsf T}.$$

求 b 到 $C(A)$ 的投影，并计算误差向量。

解　先确认列独立，再解法方程。记两列为 $a_1=(1,1,1)^{\mathsf T}$ 和 $a_2=(0,1,2)^{\mathsf T}$；若 $\alpha a_1+\beta a_2=0$，第一坐标给出 $\alpha=0$，第二坐标随即给出 $\beta=0$，所以它们独立。逐列取点积得到

$$A^{\mathsf T}A=\begin{bmatrix}3&3\\3&5\end{bmatrix},\qquad (A^{\mathsf T}A)^{-1}=\frac{1}{6}\begin{bmatrix}5&-3\\-3&3\end{bmatrix},\qquad A^{\mathsf T}b=\begin{bmatrix}6\\0\end{bmatrix}.$$

所以 $\hat{x}=(5,-3)^{\mathsf T}$，进而

$$p=A\hat{x}=(5,2,-1)^{\mathsf T},\qquad e=b-p=(1,-2,1)^{\mathsf T}.$$

对应的 3×3 投影矩阵为

$$P_A=\frac{1}{6}\begin{bmatrix}5&2&-1\\2&2&2\\-1&2&5\end{bmatrix},\qquad P_Ab=p.$$

两列 $a_1=(1,1,1)^{\mathsf T}$、$a_2=(0,1,2)^{\mathsf T}$ 都与 e 正交：$a_1^{\mathsf T}e=1-2+1=0$，$a_2^{\mathsf T}e=-2+2=0$。因此 e 垂直于整个平面。勾股检验为 $\|b\|^2=36=30+6=\|p\|^2+\|e\|^2$。这个平面也可写成 $x_1-2x_2+x_3=0$；法向量 $n=(1,-2,1)^{\mathsf T}$ 与两列都正交。本例的误差正好是 $e=n$，而 $I-P_A=\frac{1}{6}nn^{\mathsf T}$，这和下一节的补空间公式一致。这里 A 有两个独立列，所以 $C(A)$ 是 $\mathbb{R}^3$ 中的二维平面；其正交补 $N(A^{\mathsf T})$ 只有一个自由方向，正由法向量 n 张成。因此本例的残差只能沿 n 的方向变化。 ♦

图 14.2 把列空间从侧面看成一条线，突出显示 p 在子空间内而误差 e 沿法向离开子空间。

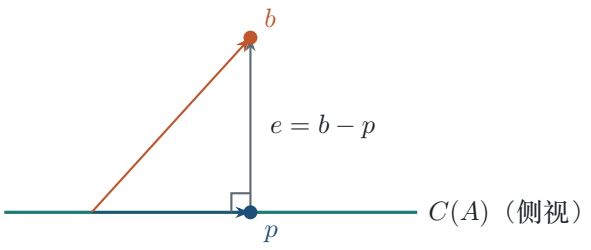


图 14.2　侧视图中，投影 p 位于 $C(A)$，残差 $e=b-p$ 垂直于该子空间。

图中从原点到 b 的向量分解为 $b=p+e$。这幅侧视图只表达「子空间方向」和「正交误差」的关系；在实际计算中，$A^{\mathsf T}e=0$ 同时检查了 e 与平面内的每个基方向是否垂直。

若 A 的列相关，投影点仍唯一，但系数可能不唯一。比如第二列是第一列的 2 倍时，两个系数组合可以表示同一个平面内的向量。

例 14.6　列相关时投影仍唯一　令

$$A=\begin{bmatrix}1&2\\2&4\\2&4\end{bmatrix},\qquad b=(1,1,1)^{\mathsf T}.$$

求 b 到 $C(A)$ 的投影，并说明 $A\hat{x}=p$ 的解是否唯一。

解　两列张成同一条直线，取 $u=(1,2,2)^{\mathsf T}$ 作为方向。投影公式给出

$$p=\frac{u^{\mathsf T}b}{u^{\mathsf T}u}u=\frac{5}{9}u=\left(\frac{5}{9},\frac{10}{9},\frac{10}{9}\right)^{\mathsf T},\qquad e=b-p=\left(\frac{4}{9},-\frac{1}{9},-\frac{1}{9}\right)^{\mathsf T}.$$

验算 $u^{\mathsf T}e=(4-2-2)/9=0$，所以这是到该直线的正交投影。这里 $A^{\mathsf T}A=\begin{bmatrix}9&18\\18&36\end{bmatrix}$ 是奇异矩阵，不能使用含 $(A^{\mathsf T}A)^{-1}$ 的公式。方程 $A\hat{x}=p$ 只要求 $\hat{x}_1+2\hat{x}_2=5/9$，所以全部解可写成 $\hat{x}=(5/9-2t,t)^{\mathsf T}$，其中 $t\in\mathbb{R}$；例如 $t=0$ 与 $t=5/18$ 分别给出 $(5/9,0)^{\mathsf T}$ 和 $(0,5/18)^{\mathsf T}$。几何投影 p 唯一，系数表示则不唯一。 ♦

这不是本例偶然出现的差别。任意子空间 $C(A)$ 都有唯一的正交投影点，因为最近点定理只依赖它是子空间；若列相关，变化的是把这个点写成 $A\hat{x}$ 时的系数自由度。只要 $z\in N(A)$，就有 $A(\hat{x}+z)=A\hat{x}$，所以一旦找到一个系数解，给它加上任意零空间向量仍得到同一个投影点。若需要使用法方程的逆公式，可以先从 A 的列中选出一组基，组成列独立矩阵 B，再对 $C(B)=C(A)$ 使用 $B(B^{\mathsf T}B)^{-1}B^{\mathsf T}$。

想一想　为什么列独立是法方程显式逆公式的条件？

答　若列相关，则存在非零 z 使 $Az=0$，于是 $z^{\mathsf T}A^{\mathsf T}Az=\|Az\|^2=0$，所以 $A^{\mathsf T}A$ 有非零零空间，不能求逆。投影仍可通过删去重复方向或直接找 $C(A)$ 的一组基来计算。

法方程的计算顺序　给定 A 和 b，先检查矩阵尺寸，再确认列是否独立；独立时计算 $G=A^{\mathsf T}A$ 与 $c=A^{\mathsf T}b$，解线性方程 $G\hat{x}=c$，最后算 $p=A\hat{x}$ 和 $e=b-p$。验算时至少检查 $A^{\mathsf T}e=0$；若还要比较距离，可使用

$$\|b-Ax\|^2=\|e\|^2+\|A(x-\hat{x})\|^2.$$

通常直接解法方程比先求出整个 G^{-1} 再相乘更省计算，也能避免把「求解线性方程」误写成「求逆矩阵」。

14.4　补空间如何表示？

投影矩阵保留 $C(A)$ 内的分量。剩下的部分 $I-P_A$ 应当保留误差并消去投影；在满列秩条件下，它恰好投到左零空间。

定理 14.5（正交补的投影矩阵）　设 $A \in \mathbb{R}^{m\times n}$ 的列线性无关，令 $P_A = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$。则

$$P_A^{\mathsf{T}} = P_A, \qquad P_A^2 = P_A,$$

且 $I - P_A$ 是到 $N(A^{\mathsf{T}}) = C(A)^{\perp}$ 的正交投影矩阵。

思路　先把 P_A 的转置和平方按矩阵乘法展开，利用 $A^{\mathsf{T}}A$ 对称及其逆矩阵消去中间因子。再验证 $I - P_A$ 的像包含在 $N(A^{\mathsf{T}})$，并检查它固定该零空间里的每个向量。

证明　记 $G = A^{\mathsf{T}}A$。G 对称且可逆，故 G^{-1} 也对称。于是

$$P_A^{\mathsf{T}} = (AG^{-1}A^{\mathsf{T}})^{\mathsf{T}} = AG^{-1}A^{\mathsf{T}} = P_A,$$

并且

$$P_A^2 = AG^{-1}(A^{\mathsf{T}}A)G^{-1}A^{\mathsf{T}} = AG^{-1}GG^{-1}A^{\mathsf{T}} = P_A.$$

令 $Q = I - P_A$。则 $Q^{\mathsf{T}} = Q$，且 $Q^2 = I - 2P_A + P_A^2 = Q$。此外

$$A^{\mathsf{T}}Q = A^{\mathsf{T}} - A^{\mathsf{T}}AG^{-1}A^{\mathsf{T}} = A^{\mathsf{T}} - GG^{-1}A^{\mathsf{T}} = 0,$$

所以 $Qb \in N(A^{\mathsf{T}})$ 对每个 b 都成立。反过来，若 $y \in N(A^{\mathsf{T}})$，则 $P_A y = 0$，因而 $Qy = y$。所以 Q 的像恰为 $N(A^{\mathsf{T}})$；它对称、幂等，确实是该空间的正交投影。 ■

这两个投影互相消去：$P_A(I-P_A) = (I-P_A)P_A = 0$。对 $y \in C(A)$，$P_A y = y$、$(I-P_A)y = 0$；对 $z \in N(A^{\mathsf{T}})$，$P_A z = 0$、$(I-P_A)z = z$。因此每个 b 都按

$$b = P_A b + (I - P_A)b$$

唯一分解成列空间分量和正交补分量。这不仅是矩阵恒等式，也把投影与误差分别对应到两个互相垂直的子空间。

若子空间是一个过原点的超平面 $S = \{x : n^{\mathsf{T}}x = 0\}$，其中 $n \neq 0$，它的正交补只有法向量方向 $\operatorname{span}\{n\}$。因此先把 b 投到法向量，再从 b 中减去该分量，就得到

$$\operatorname{proj}_S b = b - \frac{n^{\mathsf{T}}b}{n^{\mathsf{T}}n}n, \qquad P_S = I - \frac{nn^{\mathsf{T}}}{n^{\mathsf{T}}n}.$$

代入点积可验 $n^{\mathsf{T}}\operatorname{proj}_S b = 0$，残差则是 $\frac{n^{\mathsf{T}}b}{n^{\mathsf{T}}n}n$。若 $b \in S$，这个系数为零，投影保留 b；若 b 平行于 n，投影为零。练习中的平面是这个公式在三维的一个实例。

易错提醒　仅有 $P^2 = P$ 不能保证 Pb 是最近点。矩阵 $Q = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ 满足 $Q^2 = Q$，却不对称；它把 $(1,2)$ 投到 $(3,0)$，距离平方为 8，而横轴上最近点 $(1,0)$ 的距离平方只有 4。正交投影还需要 $P^{\mathsf{T}} = P$。

■ **例 14.7　投影到四维坐标子空间**　令 A 为 4×4 单位矩阵删去最后一列所得的 4×3 矩阵。求投影矩阵 P_A，并把 $b = (1,2,3,4)^{\mathsf{T}}$ 分解为 $p + e$。

解　$C(A)$ 由前三个坐标方向张成，所以

$$P_A=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix},\qquad p=P_Ab=(1,2,3,0)^{\mathsf T}.$$

因此 $e=b-p=(0,0,0,4)^{\mathsf T}$，而

$$I-P_A=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{bmatrix}$$

把向量投到最后一个坐标轴，也就是 $N(A^{\mathsf T})$。这正是「先保留前三个坐标，再留下垂直误差」的分解。 ♦

一般地，若 A 的列是若干个标准基向量 e_j，那么 P_A 是对角矩阵：对应这些列的坐标位置填 1，其余位置填 0。作用在 b 上时，它只保留选中的坐标；$I-P_A$ 则保留剩余坐标。没有选中任何坐标时，投影矩阵是零矩阵；选中全部坐标时，它是单位矩阵。坐标子空间是最容易直接看出投影效果的一类情形。

14.5　方程无解时为什么仍要投影?

三组数据点 $(1,1)$、$(2,2)$、$(3,2)$ 不在同一条直线上。若用 $y=C+Dt$ 通过三点，就要同时满足

$$C+D=1,\qquad C+2D=2,\qquad C+3D=2.$$

前两式给出 $D=1$，后两式给出 $D=0$，所以精确方程组无解。尽管如此，所有向量 Ax 都在 $C(A)$ 中，我们仍能取 b 到 $C(A)$ 的唯一最近点 p，再找 x 使 $Ax=p$。这把不相容方程转成了**最小二乘问题**（least-squares problem）。

这里投影发生在观测向量所在的 $\mathbb{R}^3$ 中。矩阵 A 把参数 (C,D) 送到三次预测值 $(C+D,C+2D,C+3D)$；这些预测值组成 $C(A)$，其中的零向量对应 $C=D=0$。图上的拟合直线可以不经过原点，但它在观测空间里对应的预测向量集合仍是一个过原点的线性子空间。要投影的是实测向量 $(1,2,2)^{\mathsf T}$，不是平面上的三组点本身。

若观测恰好落在 $C(A)$，则投影就是 b，残差为零，法方程与原方程有相同解；若数据不满足模型，投影给出离 b 最近的可预测向量。列独立时系数唯一，列相关时最佳预测向量仍唯一，但产生它的参数表示可能不止一个。

■ **例 14.8**　**三点直线拟合的法方程**　把三点写成 $A\binom{C}{D}=b$，其中

$$A=\begin{bmatrix}1&1\\1&2\\1&3\end{bmatrix},\qquad b=(1,2,2)^{\mathsf T}.$$

写出投影条件 $A^{\mathsf{T}}A\hat{x}=A^{\mathsf{T}}b$。

解　每一行对应一个数据点，第一列乘 C，第二列乘 D。$A^{\mathsf{T}}A$ 的四个位置分别是两列的点积：$1+1+1=3$、$1+2+3=6$、$1+2+3=6$、$1+4+9=14$；$A^{\mathsf{T}}b$ 的两个位置则是 $1+2+2=5$ 与 $1+4+6=11$。所以

$$A^{\mathsf{T}}A=\begin{bmatrix}3&6\\6&14\end{bmatrix},\qquad A^{\mathsf{T}}b=\begin{bmatrix}5\\11\end{bmatrix}.$$

因此法方程为 $\begin{bmatrix}3&6\\6&14\end{bmatrix}\hat{x}=\binom{5}{11}$。它要求残差 $b-A\hat{x}$ 与 A 的每一列正交；本讲解释了方程从何而来，下一讲再解出拟合系数并研究残差。 ♦

这里两列的含义可以直接读出残差条件。第一列全是 1，所以 $A^{\mathsf{T}}e=0$ 的第一条方程表示三个残差之和为零；第二列记录横坐标 t，第二条方程表示用 t 加权的残差之和也为零。残差因此不会整体偏高或偏低，也不会沿数据横坐标呈现线性趋势。

更具体地，若残差记为 $(r_1,r_2,r_3)^{\mathsf{T}}$，两个条件是

$$r_1+r_2+r_3=0,\qquad r_1+2r_2+3r_3=0.$$

相减后得到 $r_2=-2r_3$，再代回第一式得到 $r_1=r_3$，所以残差只能沿 $(1,-2,1)^{\mathsf{T}}$ 方向。拟合直线调整的是 $C(A)$ 内的分量，正交补中的分量则无法靠改变 C,D 消除；这正是方程不相容时仍能定义最佳近似的几何原因。

前后联系　第 15 讲将进一步整理 P_A 的对称、幂等性质，并用投影来完整求解最小二乘问题。这里需要记住的几何起点是：最佳拟合向量 p 位于 $C(A)$，误差 $b-p$ 垂直于 $C(A)$。

本讲小结

- **最近点刻画**：$p=\mathrm{proj}_S b \iff p\in S$ 且 $b-p\in S^{\perp}$；这个 p 是 S 中唯一离 b 最近的向量。
- **直线公式**：$a\neq 0$ 时，$\mathrm{proj}_{\mathrm{span}\{a\}}b=\dfrac{a^{\mathsf{T}}b}{a^{\mathsf{T}}a}a$，误差垂直于 a。
- **投影矩阵**：$a\neq 0$ 时 $P_a=\dfrac{aa^{\mathsf{T}}}{a^{\mathsf{T}}a}$ 满足 $P_a^{\mathsf{T}}=P_a$、$P_a^2=P_a$；它投到 $\mathrm{span}\{a\}$。
- **法方程**：A 列独立时，$\hat{x}=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}b$，$p=A\hat{x}$，且 $A^{\mathsf{T}}(b-p)=0$。
- **秩亏边界**：列相关时最近点仍唯一，但 $\hat{x}$ 可能不唯一；此时不能使用 $(A^{\mathsf{T}}A)^{-1}$。
- **正交补投影**：满列秩时，$I-P_A$ 投到 $N(A^{\mathsf{T}})=C(A)^{\perp}$；最佳拟合误差属于该正交补。

练习

14.1 ★　令 $a=(1,1,1)$、$b=(1,2,6)$。求 b 到 $\mathrm{span}\{a\}$ 的投影 p 与误差 e，并核对 $\|b\|^2=\|p\|^2+\|e\|^2$。

14.2 ★　令 $a=(3,4)$。求 P_a，计算 $P_a(5,0)^{\mathsf{T}}$；再说明为什么 $P_a(-4,3)^{\mathsf{T}}=0$。

14.3 ★　证明把非零方向向量 a 换成 ca（$c\neq 0$）不会改变投影矩阵 P_a。

14.4 ★ 令 $A=\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}$、$b=(1,1,0)^{\mathsf T}$。求法方程解、投影和残差，并核对 $A^{\mathsf T}e=0$。

14.5 ★★ 令 A 是 I_4 删除最后一列得到的矩阵。写出 P_A、$I-P_A$，并计算它们作用于 $(1,2,3,4)^{\mathsf T}$ 的结果。

14.6 ★★ 若 $P^2=P$，证明 $(I-P)^2=I-P$。这是否足以说明 $I-P$ 是正交投影？

14.7 ★★ 对 $A=\begin{bmatrix}1&2\\2&4\\2&4\end{bmatrix}$、$b=(1,1,1)^{\mathsf T}$，求投影点并描述所有满足 $A\hat{x}=p$ 的系数向量。

14.8 ★★ 对数据 $(1,1)$、$(2,2)$、$(3,2)$，写出 $A^{\mathsf T}A\hat{x}=A^{\mathsf T}b$；说明精确插值方程为何无解。

14.9 ★★★ 令 $Q=\begin{bmatrix}1&1\\0&0\end{bmatrix}$、$b=(1,2)^{\mathsf T}$。验证 $Q^2=Q$，并比较 Qb 与 b 到 x 轴的最近点，说明 Q 不是正交投影。

14.10 ★★★ 平面 $S=\{x\in\mathbb{R}^3:(1,-2,1)^{\mathsf T}x=0\}$ 的法向量为 $n=(1,-2,1)$。求 $b=(6,0,0)$ 到 S 的投影，并验证误差沿 n。

练习答案与提示

14.1 $a^{\mathsf T}b=9$、$a^{\mathsf T}a=3$，所以 $\hat{x}=3$、$p=(3,3,3)^{\mathsf T}$、$e=(-2,-1,3)^{\mathsf T}$。有 $a^{\mathsf T}e=0$，且 $\|b\|^2=41$、$\|p\|^2=27$、$\|e\|^2=14$，故 $41=27+14$。

14.2 $P_a=\dfrac{1}{25}\begin{bmatrix}9&12\\12&16\end{bmatrix}$，因此 $P_a(5,0)^{\mathsf T}=(9/5,12/5)^{\mathsf T}$。又 $\langle(3,4),(-4,3)\rangle=0$，该向量已垂直于投影直线，所以投影为零。

14.3 $P_{ca}=\dfrac{(ca)(ca)^{\mathsf T}}{(ca)^{\mathsf T}(ca)}=\dfrac{c^2aa^{\mathsf T}}{c^2a^{\mathsf T}a}=P_a$。分母非零是因为 $c\neq0$ 且 $a\neq0$。

14.4 解法方程得 $\hat{x}=(1/3,1/3)^{\mathsf T}$，故 $p=(1/3,1/3,2/3)^{\mathsf T}$、$e=(2/3,2/3,-2/3)^{\mathsf T}$；代入得 $A^{\mathsf T}e=(0,0)^{\mathsf T}$。

14.5 $P_A=\operatorname{diag}(1,1,1,0)$，$I-P_A=\operatorname{diag}(0,0,0,1)$。两者分别把 $(1,2,3,4)^{\mathsf T}$ 送到 $(1,2,3,0)^{\mathsf T}$ 与 $(0,0,0,4)^{\mathsf T}$。

14.6 $(I-P)^2=I-2P+P^2=I-P$。但这只说明补矩阵也幂等；正交性还需要对称。例 $Q=\begin{bmatrix}1&1\\0&0\end{bmatrix}$ 幂等却不对称，且将 $(1,2)^{\mathsf T}$ 送到 $(3,0)^{\mathsf T}$，不是横轴上的最近点。

14.7 取 $u=(1,2,2)^{\mathsf T}$，则 $p=\dfrac{u^{\mathsf T}b}{u^{\mathsf T}u}u=\dfrac{5}{9}u=(5/9,10/9,10/9)^{\mathsf T}$。因为 $A\hat{x}=(\hat{x}_1+2\hat{x}_2)u$，所有系数解正是 $\hat{x}_1+2\hat{x}_2=5/9$。

14.8 法方程为 $\begin{bmatrix}3&6\\6&14\end{bmatrix}\hat{x}=(5,11)^{\mathsf T}$。前两数据方程给 $D=1$，后两式给 $D=0$，矛盾，故无精确解。

14.9 $Q^2=Q$ 且 $Qb=(3,0)^{\mathsf T}$，其距离平方为 $(-2)^2+2^2=8$。横轴上最近点为 $(1,0)^{\mathsf T}$，距离平方为 $0^2+2^2=4$。此外 $Q^{\mathsf T}\neq Q$，所以它不是正交投影。

14.10 $e=\dfrac{n^{\mathsf T}b}{n^{\mathsf T}n}n=n$，故 $p=b-e=(5,2,-1)^{\mathsf T}$；并且 $n^{\mathsf T}p=5-4-1=0$，所以 $p\in S$，误差沿法向。

第 15 讲

投影矩阵与最小二乘

Projection Matrices and Least Squares

对应材料：MIT 18.06SC Session 2.3 「Projection matrices and least squares」课时摘要；习题 16.1–16.6 及解答；ZoomNotes Part 4 第 4.3 节。

本讲目标

(1) 会从正交投影构造矩阵，并验证对称、幂等和补投影性质。

(2) 会把最小二乘问题改写成法方程，说明残差为什么垂直于列空间。

(3) 能证明 $N(A^{\mathsf{T}}A)=N(A)$，据此判断 $A^{\mathsf{T}}A$ 何时可逆。

(4) 会完整计算一组直线拟合的系数、拟合值（fitted value）和误差向量。

(5) 能区分满列秩与秩亏时系数解的唯一性，以及拟合向量的唯一性。

15.1 投影矩阵有哪些代数特征？

向量 $b=(3,0)^{\mathsf{T}}$ 到直线 $y=x$ 的最近点是什么？直线方向向量可取 $a=(1,1)^{\mathsf{T}}$。第 14 讲的投影公式给出

$$p=\frac{a^{\mathsf{T}}b}{a^{\mathsf{T}}a}a=\frac{3}{2}(1,1)^{\mathsf{T}}=\left(\frac{3}{2},\frac{3}{2}\right)^{\mathsf{T}}.$$

这个计算可以写成对 b 乘一个固定矩阵：$p=Pb$。固定矩阵让同一个投影操作能直接用于一批向量，也把几何性质变成可检验的代数等式。

定义 15.1（投影矩阵） 实方阵 P 若满足 $P^2=P$，就称为**投影矩阵**（projection matrix）。若还满足 $P^{\mathsf{T}}=P$，则称为**正交投影矩阵**（orthogonal projection matrix）。

例如 $P=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}$ 满足 $P^2=P=P^{\mathsf{T}}$，把向量投到 $y=x$；$P=0$ 的像为 $\{0\}$，$P=I$ 不改变任何向量。矩阵 $\begin{bmatrix}1&1\\0&0\end{bmatrix}$ 也满足 $P^2=P$，但它不对称，是**斜投影**（oblique projection）而非正交投影。

易错提醒　不能只凭 $P^2=P$ 就断定投影方向与投影平面垂直。上面的 $\begin{bmatrix}1&1\\0&0\end{bmatrix}$ 幂等却不对称；正交投影矩阵还必须满足 $P^{\mathsf T}=P$。

命题 15.2（对称幂等矩阵确实给出正交投影）　若 $P^{\mathsf T}=P$ 且 $P^2=P$，则每个 b 都能分解为 $b=Pb+(I-P)b$，其中 $Pb\in C(P)$、$(I-P)b\in C(P)^{\perp}$。因此 Pb 是 b 到 $C(P)$ 的正交投影。

证明　令 $p=Pb$、$e=(I-P)b$。由于 p 是 P 作用在 b 上所得的向量，$p\in C(P)$；又 $Pe=P(I-P)b=0$，故 $e\in N(P)$。下面证明 $N(P)=C(P)^{\perp}$。

若 $e\in N(P)$，对任意 $y=Pu\in C(P)$，有 $e^{\mathsf T}y=e^{\mathsf T}Pu=(Pe)^{\mathsf T}u=0$。反过来，若 $e\perp C(P)$，则 $Pe\in C(P)$，所以 $e^{\mathsf T}Pe=0$。利用对称与幂等，

$$\|Pe\|^2=e^{\mathsf T}P^{\mathsf T}Pe=e^{\mathsf T}P^2e=e^{\mathsf T}Pe=0,$$

于是 $Pe=0$，即 $e\in N(P)$。因此 $e\perp C(P)$。对任意 $q\in C(P)$，$p-q$ 仍在 $C(P)$，所以

$$\|b-q\|^2=\|e+(p-q)\|^2=\|e\|^2+\|p-q\|^2\geq\|e\|^2.$$

等号只有 $q=p$ 时成立，故 p 是唯一最近点。 ■

上面的斜投影矩阵 $P=\begin{bmatrix}1&1\\0&0\end{bmatrix}$ 把 $b=(0,1)^{\mathsf T}$ 送到 $p=(1,0)^{\mathsf T}$，误差为 $(-1,1)^{\mathsf T}$。误差与横轴方向 $(1,0)^{\mathsf T}$ 的点积是 -1；原点才是横轴上离 b 最近的点，距离为 1，而斜投影的距离为 $\sqrt{2}$。这给出了「幂等但非对称」不够用的具体原因。

设 A 的列向量线性无关。第 14 讲已经把投影到 $C(A)$ 的计算写成矩阵乘法；下面列出这个矩阵的结构特征。它们说明为什么投影一次后再投影不会改变结果。

记 $N(A^{\mathsf T})$ 为 A 的左零空间（left nullspace），它由所有与 A 的每一列都垂直的向量组成。

定理 15.3（列空间的正交投影矩阵）　设 $A\in\mathbb{R}^{m\times n}$ 的列线性无关，并令

$$P=A(A^{\mathsf T}A)^{-1}A^{\mathsf T}.$$

则 $P^{\mathsf T}=P$、$P^2=P$、$C(P)=C(A)$，且 $N(P)=N(A^{\mathsf T})$。矩阵 $I-P$ 是到 $N(A^{\mathsf T})=C(A)^{\perp}$ 的正交投影。

证明　先说明逆矩阵存在。对任意 $x\neq0$，列独立给出 $Ax\neq0$，于是

$$x^{\mathsf T}A^{\mathsf T}Ax=\|Ax\|^2>0.$$

所以 $A^{\mathsf T}A$ 的零空间只有零向量，作为方阵可逆。

对称性与幂等性。因 $A^{\mathsf T}A$ 对称，其逆也对称，故 $P^{\mathsf T}=P$。直接相乘得

$$P^2=A(A^{\mathsf T}A)^{-1}(A^{\mathsf T}A)(A^{\mathsf T}A)^{-1}A^{\mathsf T}=P.$$

像空间。每个 Pb 都是 A 的列的线性组合，所以 $C(P)\subseteq C(A)$。反过来，对任意 $y=Ax$，

$$Py=A(A^{\mathsf T}A)^{-1}A^{\mathsf T}Ax=Ax=y,$$

故 $y\in C(P)$，于是 $C(P)=C(A)$。

零空间与补投影。若 $A^{\mathsf T}b=0$，公式给出 $Pb=0$。若 $Pb=0$，因 A 的列独立，$A(A^{\mathsf T}A)^{-1}A^{\mathsf T}b=0$ 只能有 $(A^{\mathsf T}A)^{-1}A^{\mathsf T}b=0$，进而 $A^{\mathsf T}b=0$。所以 $N(P)=N(A^{\mathsf T})$。此外 $P(I-P)=0$，故 $(I-P)b$ 属于 $N(P)$；反过来若 $y\in N(P)$，则 $(I-P)y=y$。因此 $I-P$ 的像空间恰为 $N(A^{\mathsf T})$，且它对称、幂等，是该空间的正交投影。 ■

两个边界情形也符合这些结论：投影到零子空间的矩阵是 0，投影到整个 $\mathbb{R}^m$ 的矩阵是 I。若 A 是可逆方阵，则 $C(A)=\mathbb{R}^m$，公式 $A(A^{\mathsf T}A)^{-1}A^{\mathsf T}$ 化为 I，每个向量都已在投影子空间中。

投影到一个子空间的矩阵不应取决于我们选了哪一组基。下面把这个事实从公式中直接算出。

> **推论 15.4（投影矩阵与基的选择无关）**　若 $R\in\mathbb{R}^{n\times n}$ 可逆且 $B=AR$，则由 B 构造的正交投影矩阵与由 A 构造的矩阵相同：$P_B=P_A$。

证明　记 $G=A^{\mathsf T}A$。由于 $B^{\mathsf T}B=R^{\mathsf T}GR$，

$$P_B=AR(R^{\mathsf T}GR)^{-1}R^{\mathsf T}A^{\mathsf T}=AR\big(R^{-1}G^{-1}(R^{\mathsf T})^{-1}\big)R^{\mathsf T}A^{\mathsf T}=AG^{-1}A^{\mathsf T}=P_A.$$

所以换一组基会改变系数，却不会改变投影后的向量。 ■

在平面例中，A 的两列是 $(1,1,1)^{\mathsf T}$ 和 $(0,1,2)^{\mathsf T}$。若把第二列换成两列之和，就得到另一组基 $(1,1,1)^{\mathsf T},(1,2,3)^{\mathsf T}$；推论 15.4 保证两种写法算出的同一个平面投影矩阵相同。

■ **例 15.1　投影到三维空间中的平面**　令 $A=\begin{bmatrix}1&0\\1&1\\1&2\end{bmatrix}$，其两列张成 $\mathbb{R}^3$ 中的一个平面。求投影矩阵，并将 $b=(6,0,0)^{\mathsf T}$ 分解为 $p+e$。

解　先算 Gram 矩阵及其逆：

$$A^{\mathsf T}A=\begin{bmatrix}3&3\\3&5\end{bmatrix},\qquad (A^{\mathsf T}A)^{-1}=\frac{1}{6}\begin{bmatrix}5&-3\\-3&3\end{bmatrix}.$$

代入 $P=A(A^{\mathsf T}A)^{-1}A^{\mathsf T}$ 得

$$P=\begin{bmatrix}\frac{5}{6}&\frac{1}{3}&-\frac{1}{6}\\[2pt]\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\\[2pt]-\frac{1}{6}&\frac{1}{3}&\frac{5}{6}\end{bmatrix},\quad p=Pb=(5,2,-1)^{\mathsf T},\quad e=b-p=(1,-2,1)^{\mathsf T}.$$

直接检查 $A^{\mathsf T}e=(0,0)^{\mathsf T}$，故误差垂直于平面；同时 $Pp=p$。本例把矩阵公式、投影值和补空间分量放在同一组计算里。 ♦

例 15.2　把一个向量投到直线　令 $a=(1,1)^{\mathsf T}$，$P=\frac{aa^{\mathsf T}}{a^{\mathsf T}a}$。计算 $b=(3,0)^{\mathsf T}$ 的投影 p 和误差 $e=b-p$。

解　先写出投影矩阵，再分别作用到 b：

$$P=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix},\qquad p=Pb=\left(\frac{3}{2},\frac{3}{2}\right)^{\mathsf T},\qquad e=b-p=\left(\frac{3}{2},-\frac{3}{2}\right)^{\mathsf T}.$$

验算 $a^{\mathsf T}e=\frac{3}{2}-\frac{3}{2}=0$，所以误差垂直于投影直线。这里 $Pb=p$，而 $Pp=p$，正是 $P^2=P$ 的几何含义。 ♦

图 15.1 把 $b=p+e$ 画在同一坐标系中：蓝色箭头 p 沿直线 $y=x$，灰色线段 e 与直线垂直。

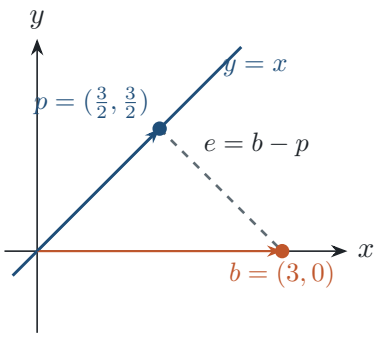


图 15.1　投影 p 是直线上的最近点，误差 $e=b-p$ 垂直于直线。

从图中看，b 可以拆成沿投影子空间的分量 p 和垂直分量 e。P 保留前者并消去后者；$I-P$ 则保留 e、消去 p。

想一想　把 b 投影一次后再投影，结果会变吗？
答　不会。第一次得到 $p\in C(P)$，而定理 15.3 的证明已说明 P 固定 $C(P)=C(A)$ 中的每个向量，所以 $Pp=p$；矩阵写法就是 $P^2b=Pb$。

15.2　无解时怎样找最近的列组合？

三个测量点 $(1,1)$、$(2,2)$、$(3,2)$ 不在同一条直线上。若用 $C+Dt$ 逐点拟合，就要解

$$C+D=1,\qquad C+2D=2,\qquad C+3D=2.$$

前两式相减得到 $D=1$，后两式相减得到 $D=0$，所以原方程组无解。可是可以找一条直线，使三个预测值整体上最接近给定数据。

设 $A\in\mathbb{R}^{m\times n}$，$b\in\mathbb{R}^m$。向量 Ax 总在列空间 $C(A)$ 中；若 $b\notin C(A)$，就不能令 $Ax=b$。我们改为最小化误差的平方 $\|Ax-b\|^2$，称为**最小二乘法**（least squares）。因为 $\|Ax-b\|\ge 0$，最小化它与最小化平方有完全相同的解。令 p 是 b 在 $C(A)$ 上的正交投影，并把 $e=b-p$ 称为**残差**（residual）。上节的结论给出 $e\in N(A^{\mathsf T})$。

定理 15.5（最小二乘法与法方程）　对任意实矩阵 A 和向量 b，向量 $\hat{x}$ 最小化 $\|Ax-b\|^2$，当且仅当

$$A^{\mathsf{T}}A\hat{x}=A^{\mathsf{T}}b.$$

这个方程称为**法方程**（normal equations）。最小化的拟合向量 $A\hat{x}=p$ 唯一；系数向量 $\hat{x}$ 是否唯一，取决于 $N(A)$。

若方程 $Ax=b$ 有精确解，则 $b\in C(A)$，投影就是 $p=b$，残差为零，最小二乘法退化为原方程的精确求解。若方程不相容，e 是 b 中无法由列组合表示的分量；改变系数只能沿 $C(A)$ 移动拟合值，不能消去这个垂直分量。精确可解也不保证系数唯一：只有同时满足 $N(A)=\{0\}$，一个拟合向量才只对应一组系数。若把 A 的列写为 $a_1,\dots,a_n$，则法方程的残差形式是

$$A^{\mathsf{T}}e=(a_1^{\mathsf{T}}e,\dots,a_n^{\mathsf{T}}e)^{\mathsf{T}}=0.$$

也就是说，每一列各自给出一条正交条件；只检查某个总和为零可能漏掉其余条件。

直观理解　先在 $C(A)$ 中找离 b 最近的点 p，再解 $Ax=p$。从 p 到 b 的残差 e 垂直于列空间，因此也垂直于每一列；把这些点积合在一起就是 $A^{\mathsf{T}}e=0$。

思路　存在性沿用第 14 讲的投影定理：子空间 $C(A)$ 中有唯一最近点 p。先证 p 对应的系数满足法方程，再用勾股分解说明这确实最小；反过来从法方程推出残差正交，也就得到最小性。

证明　**投影点满足法方程。**取 $\hat{x}_0$ 使 $A\hat{x}_0=p$，这是可能的，因为 $p\in C(A)$。残差 $e=b-p$ 垂直于 $C(A)$，特别垂直于 A 的每一列，所以 $A^{\mathsf{T}}e=0$。代入 $e=b-A\hat{x}_0$，得到 $A^{\mathsf{T}}A\hat{x}_0=A^{\mathsf{T}}b$。

法方程解给出最小误差。对任意 x，向量 $e=b-p$ 属于 $C(A)^{\perp}$，而 $p-Ax$ 属于 $C(A)$。勾股定理给出

$$\|b-Ax\|^2=\|e+(p-Ax)\|^2=\|e\|^2+\|p-Ax\|^2.$$

因此误差平方至少为 $\|e\|^2$，且当 $Ax=p$ 时取到。若 $\hat{x}$ 满足法方程，则 $b-A\hat{x}$ 与 $C(A)$ 正交，故 $A\hat{x}$ 正是投影点 p；反之，$A\hat{x}=p$ 的每个系数都满足法方程。

拟合向量唯一。子空间中的最近点 p 唯一，所以所有最小二乘解都满足 $A\hat{x}=p$。若 x_0 是一个解，则其他解恰为 x_0+z，其中 $z\in N(A)$。这说明系数唯一当且仅当 $N(A)=\{0\}$。 ■

■ **例 15.3　一个二维列空间的最小二乘解**　令 $A=\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}$，$b=(1,2,0)^{\mathsf{T}}$。求最小二乘系数、拟合向量和残差。

解　先写法方程。这里

$$A^{\mathsf{T}}A=\begin{bmatrix}2&1\\1&2\end{bmatrix},\qquad A^{\mathsf{T}}b=(1,2)^{\mathsf{T}},$$

所以 $2x_1+x_2=1$、$x_1+2x_2=2$，解为 $\hat{x}=(0,1)^{\mathsf{T}}$。于是 $p=A\hat{x}=(0,1,1)^{\mathsf{T}}$，$e=b-p=(1,1,-1)^{\mathsf{T}}$。核对 $A^{\mathsf{T}}e=(0,0)^{\mathsf{T}}$，说明误差垂直于两列；$\|e\|^2=3$。若改用系数 $x=(1,1)^{\mathsf{T}}$，则 $Ax=(1,1,2)^{\mathsf{T}}$，

$p-Ax=(-1,0,-1)^{\mathsf T}$，且 $b-Ax=(0,1,-2)^{\mathsf T}=e+(p-Ax)$。两分量点积为零，平方误差 5 正好等于 $\|e\|^2+\|p-Ax\|^2=3+2$。这个具体比较展示了法方程的勾股分解。 ♦

易错提醒　直线拟合中，残差之和为零还不够。向量 $e=(1,-1,0)^{\mathsf T}$ 的分量和为零，但对采样时刻 $(-1,1,2)$，其加权和 $-1-1+0=-2$，所以它不垂直于斜率列；最小二乘要求 $A^{\mathsf T}e=0$ 的每个分量都为零。

15.3　$A^{\mathsf T}A$ 什么时候可逆?

矩阵 $A^{\mathsf T}A$ 汇集了 A 各列的内积，称为 A 的**Gram 矩阵**（Gram matrix）。法方程总能找到最小二乘解，但把它写成 $\hat x=(A^{\mathsf T}A)^{-1}A^{\mathsf T}b$ 时多了一个条件：逆矩阵必须存在。这个条件正好由 A 的列是否独立决定。

定理 15.6（Gram 矩阵的零空间与秩）　对任意实矩阵 $A\in\mathbb{R}^{m\times n}$，有

$$N(A^{\mathsf T}A)=N(A),\qquad \operatorname{rank}(A^{\mathsf T}A)=\operatorname{rank}(A).$$

因此 $A^{\mathsf T}A$ 可逆，当且仅当 A 的列线性无关；等价地，$A^{\mathsf T}A$ 正定，当且仅当 A 满列秩。

证明　若 $Ax=0$，左乘 $A^{\mathsf T}$ 得 $A^{\mathsf T}Ax=0$。反过来，若 $A^{\mathsf T}Ax=0$，则

$$0=x^{\mathsf T}A^{\mathsf T}Ax=(Ax)^{\mathsf T}(Ax)=\|Ax\|^2,$$

从而 $Ax=0$。两零空间相同。它们都是从 $\mathbb{R}^n$ 出发的线性映射，秩-零化度定理便给出两者秩相同。

此外 $x^{\mathsf T}A^{\mathsf T}Ax=\|Ax\|^2\ge 0$，所以 $A^{\mathsf T}A$ 半正定；它正定恰好表示对每个 $x\neq 0$ 都有 $Ax\neq 0$，也就是列独立。作为 $n\times n$ 方阵，满秩与可逆等价。 ■

当 A 满列秩时，法方程可以直接解为

$$\hat x=(A^{\mathsf T}A)^{-1}A^{\mathsf T}b,\qquad p=A\hat x=A(A^{\mathsf T}A)^{-1}A^{\mathsf T}b=Pb.$$

因此，系数公式要同时用到两件事：$A^{\mathsf T}b$ 把数据与每一列的内积收集起来，$(A^{\mathsf T}A)^{-1}$ 再消除列之间的相互影响。P 把这两步合成一个作用在 $\mathbb{R}^m$ 上的矩阵。若法方程有一个解 x_0，其他解与它的差满足 $A^{\mathsf T}A(x-x_0)=0$。由定理 15.6，这等价于 $x-x_0\in N(A)$；反过来给 x_0 加上任意零空间向量也仍解法方程。因此解集精确地是 $x_0+N(A)$，而不只是「可能有重复解」。

条件「列线性无关」不能从逆公式中删去：若 A 有重复列，$A^{\mathsf T}A$ 会有非零零空间，不能取逆。但这不表示最小二乘问题没有答案。下面的例子同时展示了唯一的拟合值和不唯一的系数。

■ **例 15.4　重复列造成系数不唯一**　令 $A=\begin{bmatrix}1&1\\2&2\end{bmatrix}$、$b=(1,0)^{\mathsf T}$。求全部最小二乘系数，并计算唯一的拟合向量。

解　先解法方程。矩阵和右端为

$$A^{\mathsf{T}}A=\begin{bmatrix}5&5\\5&5\end{bmatrix},\qquad A^{\mathsf{T}}b=(1,1)^{\mathsf{T}},$$

所以条件只有 $x_1+x_2=\frac{1}{5}$。全部最小二乘系数是

$$\hat{x}=\left(\frac{1}{5},0\right)^{\mathsf{T}}+t(1,-1)^{\mathsf{T}},\qquad t\in\mathbb{R}.$$

虽然系数不唯一，$A\hat{x}$ 总等于 $p=(\frac{1}{5},\frac{2}{5})^{\mathsf{T}}$，误差 $e=b-p=(\frac{4}{5},-\frac{2}{5})^{\mathsf{T}}$。直接计算 $(1,2)e=0$，故误差垂直于列空间。取一组基 $B=(1,2)^{\mathsf{T}}$，则到该列空间的投影矩阵是 $B(B^{\mathsf{T}}B)^{-1}B^{\mathsf{T}}=\frac{1}{5}\begin{bmatrix}1&2\\2&4\end{bmatrix}$；它作用到 b 仍给出 p。♦

想一想　$A^{\mathsf{T}}A$ 奇异是否意味着最小二乘问题没有最优解？

答　不意味着。例中法方程有无穷多个系数解，且它们全映到唯一的 p。若 $A=0$，唯一拟合向量是 0，投影矩阵也是 0；此时不需要对空列基求逆。

判断可逆与唯一　先检验 A 的列是否独立。满列秩时 $A^{\mathsf{T}}A$ 可逆、系数唯一；秩亏时法方程仍有解，但所有系数解相差一个 $N(A)$ 中的向量，拟合向量不变。

15.4　怎样拟合三个测量点？

测量点 $(1,1)$、$(2,2)$、$(3,2)$ 无法由同一条直线 $b=C+Dt$ 精确通过。把三个观测值排成向量 b 后，寻找最佳直线就是在列空间 $C(A)$ 中寻找离 b 最近的向量；每一行 $[1,t_i]$ 分别记录截距项和斜率项。

■ **例 15.5**　**三点的最小二乘直线**　对数据 $(t_i,b_i)=(1,1),(2,2),(3,2)$，求最小二乘直线 $b=C+Dt$，并求三个拟合值与残差。

解　先把三个方程写成 $Ax\approx b$，其中

$$A=\begin{bmatrix}1&1\\1&2\\1&3\end{bmatrix},\qquad x=\begin{bmatrix}C\\D\end{bmatrix},\qquad b=\begin{bmatrix}1\\2\\2\end{bmatrix}.$$

依定理 15.5，$\hat{x}$ 满足 $A^{\mathsf{T}}A\hat{x}=A^{\mathsf{T}}b$。逐列点积得到

$$A^{\mathsf{T}}A=\begin{bmatrix}3&6\\6&14\end{bmatrix},\qquad A^{\mathsf{T}}b=\begin{bmatrix}5\\11\end{bmatrix}.$$

法方程是

$$3C+6D=5,\qquad 6C+14D=11.$$

第二式减去第一式的 2 倍给出 $2D=1$，所以 $D=\frac{1}{2}$；代回第一式得到 $C=\frac{2}{3}$。最佳拟合直线为 $b=\frac{2}{3}+\frac{1}{2}t$。

代入 $t=1,2,3$，拟合值分别是 $p_1=\frac{7}{6}$、$p_2=\frac{5}{3}$、$p_3=\frac{13}{6}$。残差按 $e_i=b_i-p_i$ 定义，故

$$e=\left(-\frac{1}{6},\frac{1}{3},-\frac{1}{6}\right)^{\mathsf{T}},\qquad A^{\mathsf{T}}e=\begin{bmatrix}0\\0\end{bmatrix},\qquad \|e\|^2=\frac{1}{36}+\frac{1}{9}+\frac{1}{36}=\frac{1}{6}.$$

这说明误差平方确为所有坐标残差平方之和；残差向量与两列都垂直。表 15.1 列出三个观测、拟合值与误差。作比较，直线 $b=1+\frac{1}{2}t$ 给出预测值 $(\frac{3}{2},2,\frac{5}{2})^{\mathsf{T}}$，残差为 $(-\frac{1}{2},0,-\frac{1}{2})^{\mathsf{T}}$，平方和是 $\frac{1}{2}$，大于最小值 $\frac{1}{6}$。定理 15.5 不只给出一个看似合理的拟合，还保证没有别的直线误差更小。 ♦

t_i	b_i（观测）	p_i（拟合）	$e_i=b_i-p_i$
1	1	$\frac{7}{6}$	$-\frac{1}{6}$
2	2	$\frac{5}{3}$	$\frac{1}{3}$
3	2	$\frac{13}{6}$	$-\frac{1}{6}$

表 15.1　最小二乘直线的三个预测值与残差，其残差平方和为 1/6。

图 15.2 把三个观测值和拟合线放在 (t,b) 平面里。灰色竖线段表示每个观测与预测之间的坐标差，蓝点是投影后的拟合值。

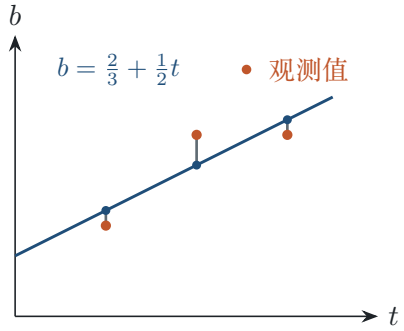


图 15.2　蓝线使观测点的竖直残差平方和最小，灰线段表示各自的残差。

从图上能看到，第二个点高于拟合线，另外两个低于拟合线；这些是各坐标方向上的残差。法方程中的正交性发生在 $\mathbb{R}^3$ 的数据向量空间：e 与常数列、时刻列都垂直，不表示图中的每段灰线都垂直于蓝线。

想一想　图中的残差线段看起来是竖直的，为什么正文又说 e 垂直于 $C(A)$？
答　图画在 (t,b) 平面中，竖线段展示每个观测值与预测值之差。把三个差按顺序组成 $e\in\mathbb{R}^3$ 后，$A^{\mathsf{T}}e=0$ 表示它垂直于 A 的两个列向量；这是数据空间中的正交关系，和图上单条线段的方向不同。

一般的**线性回归**（linear regression）也可直接由法方程写出。对 m 组数据 (t_i,b_i)，模型 $b_i=C+Dt_i$ 的法方程是

$$mC+\left(\sum_i t_i\right)D=\sum_i b_i,\qquad \left(\sum_i t_i\right)C+\left(\sum_i t_i^2\right)D=\sum_i t_ib_i.$$

记 $\bar t=\frac{1}{m}\sum_i t_i$、$\bar b=\frac{1}{m}\sum_i b_i$。第一式给出 $C=\bar b-D\bar t$；将第二式减去第一式的 $\bar t$ 倍，得到

$$D=\frac{\sum_i (t_i-\bar t)(b_i-\bar b)}{\sum_i (t_i-\bar t)^2},\qquad C=\bar b-D\bar t,$$

第一条法方程除以 m 还给出 $C+D\bar t=\bar b$，所以最佳拟合直线必过数据质心 $(\bar t,\bar b)$。在三点例中质心是 $(2,\frac{5}{3})$，而拟合线在 $t=2$ 处的值也为 $\frac{5}{3}$。前提是分母不为零。三点例中 $\bar t=2$、$\bar b=\frac{5}{3}$，分子为 1、分母为 2，所以 $D=\frac{1}{2}$、$C=\frac{2}{3}$，与直接解法方程一致。这个公式的分母是 $\sum_i t_i^2-m\bar t^2$，分子是 $\sum_i t_i b_i-m\bar t\bar b$；展开平均数定义即可得到它们分别等于中心化后的平方和与乘积和。它把斜率的来源说得很清楚：数据偏离平均时刻的方向，若与观测值偏离平均值的方向同向，斜率为正；若反向，斜率为负。

对直线模型，矩阵 A 的两列分别是常数列和时刻列。因此 $A^{\mathsf T}e=0$ 等价于两条条件

$$\sum_i e_i=0,\qquad \sum_i t_i e_i=0.$$

第一条要求正负残差总和抵消，第二条还要求残差与时刻一起加权后抵消。三点例中 $-\frac{1}{6}+\frac{1}{3}-\frac{1}{6}=0$，且 $-\frac{1}{6}+2\cdot\frac{1}{3}-3\cdot\frac{1}{6}=0$；只满足第一条并不能保证最小二乘。

分母为零恰好表示所有 t_i 相同。这个秩条件也能从 2×2 的 Gram 矩阵看出：

$$\det(A^{\mathsf T}A)=m\sum_{i=1}^{m}t_i^2-\Big(\sum_{i=1}^{m}t_i\Big)^2=m\sum_{i=1}^{m}(t_i-\bar t)^2.$$

右端是非负的；只要有两个不同的采样时刻，它就严格为正，$A^{\mathsf T}A$ 可逆，截距和斜率都能唯一确定。所有 $t_i=t_0$ 时，截距列和斜率列成比例，矩阵奇异，法方程只约束预测值 $c=C+t_0D$。此时

$$\sum_i (b_i-c)^2=\sum_i (b_i-\bar b)^2+m(c-\bar b)^2,$$

故唯一最佳预测值是 $c=\bar b$，拟合向量为 $(\bar b,\dots,\bar b)^{\mathsf T}$；满足 $C+t_0D=\bar b$ 的系数仍有无穷多组。

例 15.6　采样时刻重复时的边界　两次测量都在 $t=2$，观测值分别为 1 和 3。求最小二乘预测值，并说明系数是否唯一。

解　两行设计矩阵都是 $(1,2)$，所以 $C(A)=\operatorname{span}\{(1,1)^{\mathsf T}\}$。向量 $b=(1,3)^{\mathsf T}$ 投影到这条直线上得到 $p=(2,2)^{\mathsf T}$，残差 $e=(-1,1)^{\mathsf T}$ 与 $(1,1)^{\mathsf T}$ 垂直。法方程的 Gram 矩阵为 $\begin{bmatrix}2&4\\4&8\end{bmatrix}$，右端为 $(4,8)^{\mathsf T}$，只约束 $C+2D=2$。所以所有这样的直线在 $t=2$ 都给出同一预测值 2，而系数有无穷多组。 ♦

15.5　拟合值与误差分别属于哪里？

一组最小二乘计算会同时给出三个对象：系数 $\hat x$、拟合值 $p=A\hat x$、残差 $e=b-p$。原题 16.1–16.5 用一条直线的数据说明它们落在不同空间中。

例 15.7　三个时刻的拟合　测量值为 $b=7$（$t=-1$）、$b=7$（$t=1$）、$b=21$（$t=2$）。求最佳拟合直线，并解释其拟合向量与误差。

解　把截距和斜率作为未知数，得到

$$A=\begin{bmatrix}1&-1\\1&1\\1&2\end{bmatrix},\quad b=\begin{bmatrix}7\\7\\21\end{bmatrix},\quad A^{\mathsf T}A=\begin{bmatrix}3&2\\2&6\end{bmatrix},\quad A^{\mathsf T}b=\begin{bmatrix}35\\42\end{bmatrix}.$$

法方程给出 $3C+2D=35$、$2C+6D=42$，解为 $\hat{x}=(9,4)^{\mathsf T}$。因此最佳拟合直线为 $b=9+4t$，其三个拟合值是

$$p=A\hat{x}=(5,13,17)^{\mathsf T},\qquad e=b-p=(2,-6,4)^{\mathsf T}.$$

残差检验为 $A^{\mathsf T}e=(2-6+4,\,-2-6+8)^{\mathsf T}=(0,0)^{\mathsf T}$。所以 $p\in C(A)$，而 $e\in N(A^{\mathsf T})=C(A)^{\perp}$。若 $P=A(A^{\mathsf T}A)^{-1}A^{\mathsf T}$，则 $Pe=0$，因为 $A^{\mathsf T}e=0$；同时 $Pp=p$。 ♦

本例的 A 满列秩，故 $N(A)=\{0\}$，且 $C(A^{\mathsf T})=\mathbb{R}^2$。所以参数向量 $\hat{x}=(9,4)^{\mathsf T}$ 属于行空间 $C(A^{\mathsf T})$。秩亏时不能把这一点照搬到任意系数解：上一节的 $A=\begin{bmatrix}1&1\\2&2\end{bmatrix}$ 中，解 $(\frac{1}{5},0)^{\mathsf T}$ 不在行空间 $\operatorname{span}\{(1,1)\}$ 中。

■ **例 15.8**　**数据本身就是误差**　仍用上面的三个时刻，若观测向量换成 $b=(2,-6,4)^{\mathsf T}$，最小二乘拟合是什么？

解　先看它与 A 的列点积：$A^{\mathsf T}b=(2-6+4,\,-2-6+8)^{\mathsf T}=(0,0)^{\mathsf T}$。因此 $b\in N(A^{\mathsf T})$，它垂直于整个列空间，投影值是 $p=0$。由于本例 $A^{\mathsf T}A$ 可逆，法方程给出 $\hat{x}=0$，最佳拟合直线为 $b=0$，残差仍为原向量 $e=b$。 ♦

■ **例 15.9**　**数据恰好在模型空间中**　若观测向量为 $b=(5,13,17)^{\mathsf T}$，求最佳拟合直线和误差。

解　直接代入 $A(9,4)^{\mathsf T}=(5,13,17)^{\mathsf T}=b$。因此 $b\in C(A)$，精确解 $\hat{x}=(9,4)^{\mathsf T}$ 已使误差为零；平方误差不可能小于零，所以这也是唯一的最小二乘系数，$e=0$。 ♦

四个基本子空间在这里各有位置：$e\in N(A^{\mathsf T})$，$p\in C(A)$，参数向量位于 $C(A^{\mathsf T})=\mathbb{R}^2$，而 $N(A)=\{0\}$。具体地，$e=2(1,-3,2)^{\mathsf T}$，并且 $A^{\mathsf T}(1,-3,2)^{\mathsf T}=0$，所以 $N(A^{\mathsf T})=\operatorname{span}\{(1,-3,2)^{\mathsf T}\}$。于是 $C(A)$ 是 $\mathbb{R}^3$ 中二维平面，左零空间是一维直线；在参数空间 $\mathbb{R}^2$ 中，行空间占满整个平面，零空间只有零向量。两对正交子空间的维数分别相加为 3 和 2。

先投影，再求系数　把 b 分成 $p+e$，其中 $p\in C(A)$、$e\in N(A^{\mathsf T})$。先确定唯一的 p，再求 $A\hat{x}=p$；只有当 A 满列秩时，$\hat{x}$ 才唯一。

前后联系　第 14 讲建立了子空间上的正交投影；本讲把它写成投影矩阵并用于最小二乘。下一讲将讨论正交矩阵和 Gram–Schmidt 正交化，用正交列简化这些计算。

本讲小结

- **正交投影矩阵**满足 $P^{\mathsf{T}}=P$、$P^2=P$；满列秩 A 给出 $P=A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$，且 $C(P)=C(A)$、$N(P)=N(A^{\mathsf{T}})$。
- $I-P$ 投影到 $C(A)^{\perp}=N(A^{\mathsf{T}})$；任意 b 可分解为 $b=p+e$，其中 $p=Pb$、$e=(I-P)b$ 且 $p\perp e$。
- $\hat{x}$ 是最小二乘解，当且仅当 $A^{\mathsf{T}}A\hat{x}=A^{\mathsf{T}}b$；等价地，残差 $e=b-A\hat{x}$ 满足 $A^{\mathsf{T}}e=0$。
- 对任意 x，$\|b-Ax\|^2=\|e\|^2+\|p-Ax\|^2$；所以拟合向量 p 唯一，最小误差由正交投影刻画。
- $N(A^{\mathsf{T}}A)=N(A)$ 且 $\operatorname{rank}(A^{\mathsf{T}}A)=\operatorname{rank}(A)$；$A^{\mathsf{T}}A$ 可逆当且仅当 A 满列秩。
- 若 x_0 是法方程解，则全部系数解为 $x_0+N(A)$；秩亏时系数可不唯一，拟合值仍唯一。

练习

15.1 ★ 判断真假：任意满足 $P^2=P$ 的实矩阵 P 都表示正交投影。若为假，给出一个 2×2 反例并核对幂等性。

15.2 ★ 令 $a=(1,2)^{\mathsf{T}}$、$b=(4,-1)^{\mathsf{T}}$。求到 $\operatorname{span}\{a\}$ 的投影矩阵、投影 p 和误差 e，验证 $a^{\mathsf{T}}e=0$。

15.3 ★ 令 $A=\begin{bmatrix}1&0\\1&1\\1&2\end{bmatrix}$、$b=(6,0,0)^{\mathsf{T}}$。解最小二乘系数并求 p、e。

15.4 ★★ 设 A 的列线性无关，$P=A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$。证明 $P^{\mathsf{T}}=P$、$P^2=P$、$C(P)=C(A)$，并说明 $I-P$ 投影到哪个空间。

15.5 ★★ 对任意实矩阵 A，证明 $N(A^{\mathsf{T}}A)=N(A)$，并推出 $\operatorname{rank}(A^{\mathsf{T}}A)=\operatorname{rank}(A)$ 及 $A^{\mathsf{T}}A$ 可逆的条件。

15.6 ★★ 求最小二乘直线 $C+Dt$ 拟合数据 $(-2,4),(-1,2),(0,-1),(1,0),(2,0)$。给出法方程、系数、拟合值与残差。

15.7 ★★ 对时刻 $t=-1,1,2$ 的观测 $b=(7,7,21)^{\mathsf{T}}$，求 $\hat{x}$、p 和 e，并说明为何投影矩阵满足 $Pe=0$。

15.8 ★★ 令 $A=\begin{bmatrix}1&1\\2&2\end{bmatrix}$、$b=(3,6)^{\mathsf{T}}$。求所有最小二乘系数；说明此时拟合误差为何为零而系数不唯一。

15.9 ★★★ 若 $A=\begin{bmatrix}1&1\\0&0\end{bmatrix}$，说明公式 $A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$ 为什么失效，并写出投影到 $C(A)$ 的正交投影矩阵。

练习答案与提示

15.1 假。取 $P=\begin{bmatrix}1&1\\0&0\end{bmatrix}$，则

$$P^2=\begin{bmatrix}1&1\\0&0\end{bmatrix}=P,\qquad P^{\mathsf T}=\begin{bmatrix}1&0\\1&0\end{bmatrix}\neq P.$$

所以它是斜投影，不是正交投影。

15.2 $a^{\mathsf T}a=5$，故

$$P=\frac{aa^{\mathsf T}}{a^{\mathsf T}a}=\frac15\begin{bmatrix}1&2\\2&4\end{bmatrix},\quad p=Pb=\left(\frac25,\frac45\right)^{\mathsf T},\quad e=b-p=\left(\frac{18}{5},-\frac95\right)^{\mathsf T}.$$

$a^{\mathsf T}e=\frac{18}{5}-\frac{18}{5}=0$，验证误差垂直于投影直线。

15.3 $A^{\mathsf T}A=\begin{bmatrix}3&3\\3&5\end{bmatrix}$、$A^{\mathsf T}b=(6,0)^{\mathsf T}$。解得 $\hat x=(5,-3)^{\mathsf T}$，于是 $p=A\hat x=(5,2,-1)^{\mathsf T}$，$e=b-p=(1,-2,1)^{\mathsf T}$；乘 $A^{\mathsf T}$ 得 $A^{\mathsf T}e=(0,0)^{\mathsf T}$。

15.4 由 $A^{\mathsf T}A$ 对称，其逆也对称，所以 $P^{\mathsf T}=P$。相乘得

$$P^2=A(A^{\mathsf T}A)^{-1}(A^{\mathsf T}A)(A^{\mathsf T}A)^{-1}A^{\mathsf T}=P.$$

每个 Pb 都在 $C(A)$；若 $y=Ax$，则 $Py=Ax=y$，所以 $C(P)=C(A)$。由定理 15.3，$I-P$ 投影到 $N(A^{\mathsf T})=C(A)^{\perp}$。

15.5 若 $A^{\mathsf T}Ax=0$，则 $0=x^{\mathsf T}A^{\mathsf T}Ax=\|Ax\|^2$，故 $Ax=0$；反方向由左乘 $A^{\mathsf T}$ 立即得到。因此 $N(A^{\mathsf T}A)=N(A)$。由秩-零化度定理，两者秩相同；$A^{\mathsf T}A$ 是 $n\times n$ 方阵，所以它可逆当且仅当 $\operatorname{rank}(A^{\mathsf T}A)=n$，也就是 $\operatorname{rank}(A)=n$、A 满列秩。

15.6 取 A 的五行为 $(1,-2),(1,-1),(1,0),(1,1),(1,2)$。逐列求点积得

$$A^{\mathsf T}A=\begin{bmatrix}5&0\\0&10\end{bmatrix},\qquad A^{\mathsf T}b=(5,-10)^{\mathsf T}.$$

所以 $C=1,D=-1$，最佳直线是 $1-t$。拟合值为 $(3,2,1,0,-1)^{\mathsf T}$，残差为 $(1,0,-2,0,1)^{\mathsf T}$，且 $A^{\mathsf T}e=(0,0)^{\mathsf T}$。

15.7 法方程为 $\begin{bmatrix}3&2\\2&6\end{bmatrix}\hat x=(35,42)^{\mathsf T}$，解得 $\hat x=(9,4)^{\mathsf T}$。于是 $p=(5,13,17)^{\mathsf T}$、$e=(2,-6,4)^{\mathsf T}$，且 $A^{\mathsf T}e=(0,0)^{\mathsf T}$。由 $P=A(A^{\mathsf T}A)^{-1}A^{\mathsf T}$，$Pe=A(A^{\mathsf T}A)^{-1}(A^{\mathsf T}e)=0$。

15.8 $A^{\mathsf T}A=\begin{bmatrix}5&5\\5&5\end{bmatrix}$、$A^{\mathsf T}b=(15,15)^{\mathsf T}$，所以法方程只有条件 $x_1+x_2=3$。全部解为 $(3,0)^{\mathsf T}+t(1,-1)^{\mathsf T}$，$t\in\mathbb{R}$；每个解都给出 $A\hat x=(3,6)^{\mathsf T}=b$，故误差为零，而 $N(A)=\operatorname{span}\{(1,-1)^{\mathsf T}\}$ 使系数不唯一。

15.9 $A^{\mathsf T}A=\begin{bmatrix}1&1\\1&1\end{bmatrix}$，行列式为 0，所以逆矩阵不存在。$C(A)=\operatorname{span}\{(1,0)^{\mathsf T}\}$，到该直线的正交投影矩阵为 $P=\begin{bmatrix}1&0\\0&0\end{bmatrix}$；它满足 $P^{\mathsf T}=P$、$P^2=P$ 且 $PA=A$。

第 16 讲

正交矩阵与 Gram–Schmidt 正交化

Orthogonal Matrices and Gram–Schmidt

对应材料：MIT 18.06SC Session 2.4 summary；Exercises 17.1–17.2；ZoomNotes Part 4, Section 4.4。

本讲目标

(1) 会判断一组向量是否正交、正交归一，并能说明零向量为何不能属于正交归一组。

(2) 会用 $Q^{\mathsf{T}}Q=I$ 判断矩阵的列是否正交归一，并能区分方阵与矩形矩阵。

(3) 能证明正交矩阵保持内积、长度和距离，且 $Q^{-1}=Q^{\mathsf{T}}$。

(4) 会逐步执行 Gram–Schmidt 正交化，写出每个投影系数与残差。

(5) 能解释线性无关条件的作用，识别零残差，并由 Gram–Schmidt 写出 $A=QR$。

16.1 正交与正交归一：垂直之外还要单位长度

向量 $(3,4)$ 与 $(-4,3)$ 的内积是 $-12+12=0$，所以它们互相垂直；但两者长度都是 5。把它们各除以 5，长度变成 1，垂直关系不变。下面区分这两个要求。

定义 16.1（正交与正交归一） 向量 u,v 若满足 $u^{\mathsf{T}}v=0$，称它们**正交**（orthogonal）。向量组 $q_1,\dots,q_k$ 若满足

$$q_i^{\mathsf{T}}q_j=\begin{cases}1, & i=j,\\ 0, & i\neq j,\end{cases}$$

称为**正交归一**（orthonormal）向量组。

例如 $(3,4)$ 与 $(-4,3)$ 正交，因为内积为 0；除以 5 后得到正交归一组。向量 $(1,0)$ 与 $(1,1)$ 不正交，因为内积为 1。零向量与任何向量内积为 0，因此可出现在正交组中；但它长度为 0，不能出现在正交归一组中。

易错提醒 两个单位向量不一定正交。例如 $(1,0)$ 与 $\frac{1}{\sqrt{2}}(1,1)$ 长度都为 1，内积却是 $1/\sqrt{2}$。正交归一同时要求「互相垂直」和「各自长度为 1」。

■ **例 16.1 检验正交归一** 判断 $q_1=\frac{1}{5}(3,4)$ 与 $q_2=\frac{1}{5}(-4,3)$ 是否构成正交归一组。

解　先分别计算长度平方和内积：$q_1^{\mathsf T}q_1=(9+16)/25=1$，$q_2^{\mathsf T}q_2=(16+9)/25=1$，且 $q_1^{\mathsf T}q_2=(-12+12)/25=0$。因此它们构成正交归一组。这个检验可以写成内积矩阵 $[q_1\ q_2]^{\mathsf T}[q_1\ q_2]=I_2$。 ♦

命题 16.2（非零正交向量组线性无关）　若 $v_1,\dots,v_k$ 两两正交且都非零，则它们线性无关。

证明　假设 $c_1v_1+\cdots+c_kv_k=0$。两边与 v_j 取内积后，所有交叉项都为零，留下 $c_j(v_j^{\mathsf T}v_j)=0$。因 $v_j\neq 0$，有 $v_j^{\mathsf T}v_j=\|v_j\|^2>0$，所以 $c_j=0$。对每个 j 都如此，只有平凡线性组合。 ■

非零条件不可省略：正交组 $(0,e_1)$ 的两向量内积为 0，但它们线性相关。正交归一组自动满足非零条件，因此也自动线性无关。

想一想　为什么正交组允许有零向量，正交归一组却不允许？

答　正交只要求不同向量的内积为 0；零向量与任何向量的内积都是 0。正交归一还要求每个向量与自身的内积为 1，而 $0^{\mathsf T}0=0$，所以零向量不合格。

16.2　正交矩阵：长度、内积与距离保持

矩阵 $Q=\frac{1}{5}\begin{bmatrix}3&-4\\4&3\end{bmatrix}$ 把 $(1,2)$ 变成 $(-1,2)$。输入输出长度平方都是 5。这不是巧合：Q 的两列正交且长度为 1，因此 $Q^{\mathsf T}Q=I$。

定义 16.3（列正交归一矩阵与正交矩阵）　若矩阵 $Q=[q_1\ \cdots\ q_n]$ 的列向量构成正交归一组，就称 Q 的列正交归一，等价地 $Q^{\mathsf T}Q=I_n$。当这样的 Q 是方阵时，称为**正交矩阵**（orthogonal matrix）。

这个等价关系可以逐项读出：$(Q^{\mathsf T}Q)_{ij}=q_i^{\mathsf T}q_j$，所以 $Q^{\mathsf T}Q=I_n$ 正是「第 i 列与第 j 列的内积等于 δ_{ij}」。如果 Q 是 $m\times n$ 矩阵，$Q^{\mathsf T}Q$ 是 $n\times n$；方阵时它与 $QQ^{\mathsf T}$ 才有相同大小。

方阵 $Q=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ 的列是 $(0,1)$ 与 $(-1,0)$，故为正交矩阵；矩形矩阵 $R=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}$ 的两列也正交归一，但 R 不是方阵。零矩阵的列互相正交，却没有单位长度，所以不满足 $Q^{\mathsf T}Q=I$。

定理 16.4（正交列保持内积）　设 Q 是 $m\times n$ 矩阵且 $Q^{\mathsf T}Q=I_n$。对任意 $x,y\in\mathbb{R}^n$，

$$(Qx)^{\mathsf T}(Qy)=x^{\mathsf T}y.$$

因此 $\|Qx\|=\|x\|$，且 $\|Qx-Qy\|=\|x-y\|$。若 Q 还是方阵，则 $Q^{-1}=Q^{\mathsf T}$，并且 $QQ^{\mathsf T}=I$。

直观理解　列正交归一意味着坐标轴被送到一组新的互相垂直的单位方向。沿这些新方向组合坐标时，分量间没有交叉项，长度平方仍是坐标平方和。方阵时新方向铺满整个空间，所以变换既不拉伸也不压缩，并且可以反向恢复。

思路　先把 $(Qx)^{\mathsf{T}}(Qy)$ 改写成含 $Q^{\mathsf{T}}Q$ 的式子，再代入单位矩阵。长度和距离分别是内积的特例。方阵部分先证明列无关，再用方阵的可逆性说明左逆就是逆。

证明　由转置与乘法的规则，

$$(Qx)^{\mathsf{T}}(Qy) = x^{\mathsf{T}}Q^{\mathsf{T}}Qy = x^{\mathsf{T}}I_n y = x^{\mathsf{T}}y.$$

取 $y = x$ 得 $\|Qx\|^2 = \|x\|^2$；把 x, y 分别看作两个输入，$Qx - Qy = Q(x-y)$，所以长度相同也给出距离保持。

现在设 Q 为 $n \times n$ 方阵。若 $Qx = 0$，则 $x = Q^{\mathsf{T}}Qx = 0$，所以 Q 的零空间只有零向量，Q 可逆。因为 $Q^{\mathsf{T}}Q = I_n$，Q^{T} 是 Q 的左逆；可逆矩阵的逆唯一，故 $Q^{-1} = Q^{\mathsf{T}}$，从而 $QQ^{\mathsf{T}} = I_n$。后一个等式说明 Q 的行也构成正交归一组。 ■

若 x, y 都非零，它们夹角的余弦是 $x^{\mathsf{T}}y/(\|x\|\|y\|)$。内积和长度都保持，因此 Qx, Qy 的夹角也相同。这个结论对矩形列正交归一矩阵同样成立；变换可以把向量嵌入更高维空间，却不会改变夹角。

> **推论 16.5（正交矩阵的乘积与转置）**　若 Q_1, Q_2 是同阶正交矩阵，则 Q_1Q_2 也是正交矩阵；Q^{T} 也是正交矩阵。

证明　利用 $Q_i^{\mathsf{T}}Q_i = I$，

$$(Q_1Q_2)^{\mathsf{T}}(Q_1Q_2) = Q_2^{\mathsf{T}}(Q_1^{\mathsf{T}}Q_1)Q_2 = Q_2^{\mathsf{T}}Q_2 = I.$$

又 $Q^{-1} = Q^{\mathsf{T}}$，所以 $(Q^{\mathsf{T}})^{\mathsf{T}}Q^{\mathsf{T}} = QQ^{\mathsf{T}} = I$。 ■

乘积仍正交，但次序通常不能交换。令 $S = \operatorname{diag}(1, -1)$ 表示关于 x 轴的反射，$Q = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$ 表示旋转，则

$$QS = \begin{bmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{bmatrix}, \qquad SQ = \begin{bmatrix} 3/5 & -4/5 \\ -4/5 & -3/5 \end{bmatrix}.$$

两者都正交，却不相等；例如它们作用在 e_1 上分别得到 $(3/5, 4/5)$ 与 $(3/5, -4/5)$。矩阵乘积表示按顺序连续变换，先反射再旋转与先旋转再反射可以有不同结果。

■ **例 16.2　置换矩阵**　令 $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$。说明它是正交矩阵，并描述 P^{T}。

解　P 的列是 e_2, e_3, e_1，因此列正交归一，$P^{\mathsf{T}}P = I_3$。它把 (x_1, x_2, x_3) 变为 (x_3, x_1, x_2)。其转置 P^{T} 也正交，并把坐标按反向次序移回；具体地 $P^{-1} = P^{\mathsf{T}}$。 ♦

例如 $P(2, -1, 3) = (3, 2, -1)$，再乘 P^{T} 得回 $(2, -1, 3)$。置换没有改动坐标的大小，只改了顺序，因此长度平方仍是 $2^2 + (-1)^2 + 3^2 = 14$。

列正交归一是保持内积的关键条件。若改为 $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$，则 $D(1, 0) = (2, 0)$，长度平方从 1 变成 4。方阵条件则关系到满射和逆矩阵；矩形 R 满足 $R^{\mathsf{T}}R = I_2$，但 $RR^{\mathsf{T}} = \mathrm{diag}(1, 1, 0) \neq I_3$。

■ **例 16.3　旋转后的向量**　对 $Q = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$，计算 Qx, Qy，其中 $x = (1, 2)$、$y = (2, -1)$，并核对内积与长度。

解　先用 $Q^{\mathsf{T}}Q = I$ 确认这是正交矩阵，再逐行相乘：$Qx = (-1, 2)$，$Qy = (2, 1)$。原向量内积为 $1 \cdot 2 + 2 \cdot (-1) = 0$，像向量内积为 $(-1) \cdot 2 + 2 \cdot 1 = 0$；两组长度平方都是 5。因此原来垂直的两向量仍垂直，长度也未改变。 ♦

若反过来问「哪个输入会被 Q 送到 $(-1, 2)$」，可以直接乘转置：$Q^{\mathsf{T}}(-1, 2) = (1, 2)$。它不用重新解一个方程组，原因是正交矩阵的逆就是转置。

如图 16.1 所示，同一正交变换把标准坐标轴转到新方向，向量 x 的像为 Qx；两者长度相同。

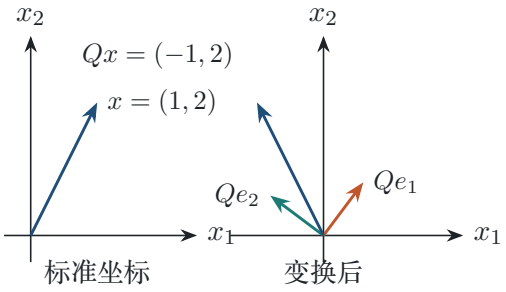


图 16.1　矩阵 Q 把标准正交归一坐标轴送到新方向，并保持 x 的长度 $\sqrt{5}$。

右图的两条坐标轴仍互相垂直且长度为 1，只是方向转了。正交矩阵还可以表示反射，例如 $\mathrm{diag}(1, -1)$；它仍保持内积和距离。

图中 Qe_1, Qe_2 是 Q 的两列。任意 $x = (x_1, x_2)$ 都可写成 $x_1e_1 + x_2e_2$，线性性便给出 $Qx = x_1Qe_1 + x_2Qe_2$。两列的长度为 1、内积为 0，所以组合后的长度平方仍是 $x_1^2 + x_2^2$。

■ **例 16.4　矩形矩阵的列**　令 $R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$。验证它的列正交归一，并计算 $R(2, -1)$。

解　两列分别是 $e_1 = (1, 0, 0)$、$e_2 = (0, 1, 0)$，内积矩阵为 $R^{\mathsf{T}}R = I_2$。直接相乘得 $R(2, -1) = (2, -1, 0)$，长度平方从 $2^2 + (-1)^2 = 5$ 变到 $2^2 + (-1)^2 + 0^2 = 5$。像落在 xy 平面中，不会填满整个 $\mathbb{R}^3$。 ♦

想一想　矩形矩阵的列正交归一，是否意味着 $RR^{\mathsf T}=I$？
答　不意味着。上例中 $RR^{\mathsf T}=\operatorname{diag}(1,1,0)$；第三个坐标方向不在列空间中。列正交归一保证输入空间的长度不变，方阵条件才保证像覆盖整个输出空间。

■ **例 16.5　Hadamard 矩阵**　把 $H=\begin{bmatrix}1&1\\1&-1\end{bmatrix}$ 的列归一化；再验证下列四阶矩阵的列正交归一：

$$H_4=\frac{1}{2}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{bmatrix}.$$

解　H 的两列内积为 0，平方长度各为 2，所以 $H/\sqrt{2}$ 的列正交归一。称这种由 1、-1 组成且经比例缩放后正交的方阵为**Hadamard 矩阵**（Hadamard matrix）。四阶例子中，每列平方长度为 $4/4=1$；任意两列在缩放前都有两个 1 与两个 -1 的乘积和，内积为 0，所以 $H_4^{\mathsf T}H_4=I_4$。 ♦

■ **例 16.6　矩形正交列补成方阵**　验证 $Q=\frac{1}{3}\begin{bmatrix}1&-2\\2&-1\\2&2\end{bmatrix}$ 的列正交归一，并用 $q_3=\frac{1}{3}(2,-2,1)$ 补成方阵。

解　两列平方长度都是 $9/9=1$，内积为 $(-2-2+4)/9=0$。向量 q_3 与两列的未缩放点积分别为 $2-4+2=0$、$-4+2+2=0$，平方长度为 $(4+4+1)/9=1$。所以

$$\widehat{Q}=\frac{1}{3}\begin{bmatrix}1&-2&2\\2&-1&-2\\2&2&1\end{bmatrix}$$

满足 $\widehat{Q}^{\mathsf T}\widehat{Q}=I_3$，是正交矩阵。 ♦

注 16.1　若矩形矩阵 Q 的列正交归一，则 $p=QQ^{\mathsf T}b$ 是 b 在 $\operatorname{span}(Q)$ 上的正交投影：p 属于列空间，且 $Q^{\mathsf T}(b-p)=Q^{\mathsf T}b-Q^{\mathsf T}QQ^{\mathsf T}b=0$。因此坐标系正交归一时，系数直接由 $Q^{\mathsf T}b$ 给出。若 Q 是方阵，列空间已是整个空间，$QQ^{\mathsf T}=I$。

■ **例 16.7　用正交列计算投影**　对 $Q=\frac{1}{3}\begin{bmatrix}1&-2\\2&-1\\2&2\end{bmatrix}$ 与 $b=(1,2,0)$，计算投影系数、投影 p 和误差 $e=b-p$。

解　直接取内积得 $c=Q^{\mathsf T}b=(5/3,-4/3)$。所以

$$p=Qc=(13/9,14/9,2/9),\qquad e=b-p=(-4/9,4/9,-2/9).$$

核对 $Q^{\mathsf{T}}e=(0,0)$，故误差与两列都垂直。又 $\|b\|^2=5$、$\|p\|^2=41/9$、$\|e\|^2=4/9$，满足 $\|b\|^2=\|p\|^2+\|e\|^2$。正交归一列让投影系数不需要解一般的联立方程。 ♦

16.3　两个向量的 Gram–Schmidt 正交化

给定独立向量 a,b，先保留 $A=a$。要从 b 中去掉沿 A 的分量，可以令 $B=b-cA$ 并要求 $A^{\mathsf{T}}B=0$。解这个一元方程即可确定 c。

定义 16.6（Gram–Schmidt 正交化）　对线性无关向量组，依次从每个新向量中减去它在先前正交方向上的分量，得到同一张成空间中的正交向量组；再把每个非零向量除以自身长度，得到正交归一向量组。这一过程称为**Gram–Schmidt 正交化**（Gram–Schmidt process）。

对两个独立向量，计算给出

$$B=b-\frac{A^{\mathsf{T}}b}{A^{\mathsf{T}}A}A,\qquad A=a.$$

分母 $A^{\mathsf{T}}A=\|A\|^2$；只要 $A\neq 0$，就能相除。随后取 $q_1=A/\|A\|$、$q_2=B/\|B\|$。

系数不是凭图猜出来的：条件 $A^{\mathsf{T}}(b-cA)=0$ 展开为 $A^{\mathsf{T}}b-cA^{\mathsf{T}}A=0$，唯一解是 $c=(A^{\mathsf{T}}b)/(A^{\mathsf{T}}A)$。由于 $B=b-cA$ 且 $A^{\mathsf{T}}B=0$，勾股定理还给出

$$\|B\|^2=\|b\|^2-\frac{(A^{\mathsf{T}}b)^2}{A^{\mathsf{T}}A}.$$

被减掉的是 b 沿 A 方向的长度平方，剩下的正是垂直部分的长度平方。

定理 16.7（两向量的正交化）　若 a,b 线性无关，则上式所得 A,B 均非零、彼此正交，且 $\operatorname{span}(A,B)=\operatorname{span}(a,b)$。归一化后 q_1,q_2 是同一平面的正交归一基。

证明　由定义，$A^{\mathsf{T}}B=A^{\mathsf{T}}b-\frac{A^{\mathsf{T}}b}{A^{\mathsf{T}}A}A^{\mathsf{T}}A=0$。若 $A=0$，则 $a=0$，与线性无关矛盾。若 $B=0$，则 $b=\frac{A^{\mathsf{T}}b}{A^{\mathsf{T}}A}A$，所以 b 是 a 的倍数，也与线性无关矛盾。

又 $A=a$，而 $B=b-cA$；反过来 $b=B+cA$。两组向量互相线性表示，因此张成空间相同。A,B 非零，归一化只改变长度，不改变方向或张成空间。 ■

条件「a,b 线性无关」不能去掉：若 $a=(1,0)$、$b=(2,0)$，则 $B=b-2A=0$，无法把 B 归一化。

■ **例 16.8　在 $\mathbb{R}^3$ 中正交化**　对 $a=(1,1,0)$、$b=(1,0,1)$ 执行 Gram–Schmidt，并写出一组正交归一基。

解　先取 $A=a$，于是 $A^{\mathsf{T}}A=2$、$A^{\mathsf{T}}b=1$。因此

$$B=b-\frac{1}{2}A=\left(\frac{1}{2},-\frac{1}{2},1\right),\qquad A^{\mathsf{T}}B=0,\qquad B^{\mathsf{T}}B=\frac{3}{2}.$$

归一化得 $q_1 = \frac{1}{\sqrt{2}}(1,1,0)$、$q_2 = \frac{1}{\sqrt{6}}(1,-1,2)$。核对两者内积为 0、长度平方为 1。此外 $b = B + \frac{1}{2}A$，所以它们仍张成原来的平面。 ♦

输入顺序会改变得到的基，但不改变平面。若先取 b，则 $q_1' = (1,0,1)/\sqrt{2}$；从 a 中减去它的投影后，$q_2' = (1,2,-1)/\sqrt{6}$。这组向量也正交归一，并与原向量组张成同一平面。

■ **例 16.9　MIT 课件的三维例子**　对 $a = (1,1,1)$ 与 $b = (1,0,2)$ 正交化，并说明投影系数如何决定残差。

解　$A = a$ 满足 $A^{\mathsf{T}}A = 3$、$A^{\mathsf{T}}b = 3$，所以 b 在 A 方向上的系数为 1。因此 $B = b - A = (0,-1,1)$，且 $A^{\mathsf{T}}B = 0$、$B^{\mathsf{T}}B = 2$。归一化得 $q_1 = (1,1,1)/\sqrt{3}$、$q_2 = (0,-1,1)/\sqrt{2}$。这里系数为 1，所以新的垂直方向是从 b 中减去整个 A。 ♦

如图 16.2 所示，b 被拆成沿 a 的分量 p 与垂直残差 B；正交化保留前者方向，并把后者作为新的垂直方向。

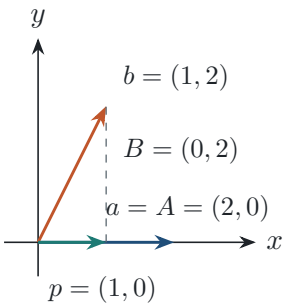


图 16.2　b 分解为沿 A 的投影 $p = (1,0)$ 与正交残差 $B = (0,2)$。

图中的数值来自 $A^{\mathsf{T}}b/A^{\mathsf{T}}A = 2/4 = 1/2$，所以 $p = (1/2)A = (1,0)$；残差 $b - p = (0,2)$ 与 A 的内积为 0。这正是一般算法每一步重复的操作。

想一想　若第二个向量恰好平行于第一个向量，Gram–Schmidt 的第二步会发生什么？
答　第二个向量会被第一个方向完全解释，减去其投影后残差为零。此时原向量组线性相关，没有二维正交基可供归一化；若只想给它们的张成空间找基，保留第一个方向即可。

16.4　多个向量：逐个减去已有方向

三个独立向量 a, b, c 需要先得到 $A = a$，再让 B 垂直于 A，最后让 C 同时垂直于 A, B。每一步都只减去新向量在旧正交方向上的分量。

给定线性无关的 $a_1, \dots, a_n$，递推定义

$$A_1 = a_1, \qquad A_k = a_k - \sum_{j=1}^{k-1} \frac{A_j^{\mathsf{T}} a_k}{A_j^{\mathsf{T}} A_j} A_j \quad (k \ge 2), \qquad q_k = \frac{A_k}{\|A_k\|}.$$

分母是先前残差的长度平方。等价地，已经归一化时可直接写

$$u_k = a_k - \sum_{j=1}^{k-1} (q_j^{\mathsf{T}} a_k) q_j, \qquad q_k = \frac{u_k}{\|u_k\|}.$$

因为旧方向彼此正交，减去 a_k 在 q_i 上的分量，不会改变它在其他 q_j 方向上的分量；所以各系数可以直接从原向量 a_k 计算。三向量例子表明，若把旧的非正交向量误当作坐标轴，减去一个投影后仍可能留下沿另一个方向的分量。

定理 16.8（Gram–Schmidt 定理）　若 $a_1, \dots, a_n$ 线性无关，则上述递推中的每个 A_k 都非零，所得 $q_1, \dots, q_n$ 正交归一，且

$$\operatorname{span}(q_1, \dots, q_k) = \operatorname{span}(a_1, \dots, a_k) \quad (1 \le k \le n).$$

思路　按 k 归纳。对每个新残差，先与任一旧方向取内积，利用旧向量彼此正交消去求和；再用定义说明新旧向量组张成空间相同。最后以线性无关排除残差为零，保证归一化合法。

证明　基步：$A_1 = a_1 \neq 0$，因为线性无关向量组不含零向量。对 $k \ge 2$，假设前 $k-1$ 步已得到互相正交的非零向量 $A_1, \dots, A_{k-1}$，且前 $k-1$ 个向量组张成相同空间。对任意 $i < k$，

$$\begin{aligned} A_i^{\mathsf{T}} A_k &= A_i^{\mathsf{T}} a_k - \sum_{j<k} \frac{A_j^{\mathsf{T}} a_k}{A_j^{\mathsf{T}} A_j} A_i^{\mathsf{T}} A_j \\ &= A_i^{\mathsf{T}} a_k - \frac{A_i^{\mathsf{T}} a_k}{A_i^{\mathsf{T}} A_i} A_i^{\mathsf{T}} A_i = 0. \end{aligned}$$

所以 A_k 与每个旧方向正交。又由递推式，A_k 是 a_k 与旧向量的线性组合；反过来

$$a_k = A_k + \sum_{j<k} \frac{A_j^{\mathsf{T}} a_k}{A_j^{\mathsf{T}} A_j} A_j.$$

故前 k 个 A 与前 k 个 a 张成同一空间。若 $A_k = 0$，则递推式会给出 $a_k \in \operatorname{span}(A_1, \dots, A_{k-1}) = \operatorname{span}(a_1, \dots, a_{k-1})$，与原向量组线性无关矛盾。因此 $A_k \neq 0$，可以归一化。归一化不改变正交性或张成空间，归纳完成。 ■

正交归一向量组必线性无关：若 $Q = [q_1 \ \cdots \ q_n]$ 且 $Q^{\mathsf{T}} Q = I_n$，那么 $Qx = 0$ 蕴含 $x = Q^{\mathsf{T}} Q x = 0$。这个证明不要求 Q 方阵，正好覆盖 Session 2.4 的矩形矩阵情形。

■ **例 16.10　Session 2.4：四维超平面的基**　对 $a = (1, -1, 0, 0)$、$b = (0, 1, -1, 0)$、$c = (0, 0, 1, -1)$ 正交化，并证明原向量与所得向量都是 $d = (1, 1, 1, 1)$ 的正交补的一组基。

解　按递推式逐个减去已有方向。先取 $A = a = (1, -1, 0, 0)$，则 $A^{\mathsf{T}} A = 2$。由于 $A^{\mathsf{T}} b = -1$，

$$B = b + \frac{1}{2} A = \left(\frac{1}{2}, \frac{1}{2}, -1, 0 \right), \qquad B^{\mathsf{T}} B = \frac{3}{2}.$$

又 $A^{\mathsf{T}} c = 0$、$B^{\mathsf{T}} c = -1$，所以

$$C = c + \frac{2}{3} B = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1 \right).$$

直接计算得 $A^{\mathsf T}B = A^{\mathsf T}C = B^{\mathsf T}C = 0$，且 $A^{\mathsf T}A = 2$、$B^{\mathsf T}B = 3/2$、$C^{\mathsf T}C = 4/3$。因此三者非零且正交，归一后即为正交归一组。它们各自与 d 的内积都是 0。

$d^{\mathsf T}x = 0$ 是一个非零线性方程，故 d 的正交补维数为 3。三个非零正交向量线性无关，因此 $\{A, B, C\}$ 是这个正交补的基。原向量也都与 d 正交；若 $\alpha a + \beta b + \gamma c = 0$，比较第一、第四、第三坐标依次得 $\alpha = 0$、$\gamma = 0$、$\beta = 0$，所以 a, b, c 也线性无关，因而是同一空间的基。 ♦

■ **例 16.11　第三步要用正交化后的方向**　对 $a = (1, 1, 0)$、$b = (1, 0, 1)$、$c = (0, 1, 1)$ 执行三步正交化，并检验一种常见的错误做法。

解　$A = a$，$A^{\mathsf T}b = 1$，所以 $B = b - \frac{1}{2}A = (\frac{1}{2}, -\frac{1}{2}, 1)$。对第三个向量，$A^{\mathsf T}c = 1$、$B^{\mathsf T}c = 1/2$、$B^{\mathsf T}B = 3/2$，故

$$C = c - \frac{1}{2}A - \frac{1}{3}B = \left(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right).$$

直接计算 $A^{\mathsf T}B = A^{\mathsf T}C = B^{\mathsf T}C = 0$，三条残差的平方长度为 $2, 3/2, 4/3$。若错误地直接从 c 减去它在原向量 a, b 上的两个投影，会得到 $c - \frac{1}{2}a - \frac{1}{2}b = (-1, 1/2, 1/2)$；它与 a 的内积是 $-1/2$，仍不垂直。 ♦

易错提醒　不能把零残差除以它自己的长度。例如 $a = (1, 0, 0)$、$b = (2, 0, 0)$ 时，第二步残差为 $b - 2a = 0$。它说明新向量没有带来新方向；应跳过该残差，不能写 $0/\|0\|$。

若输入向量组线性相关，第一次出现依赖于先前向量的 a_k 就会产生零残差。跳过零残差后，其余非零正交向量仍张成原向量组的空间；向量数等于该空间的维数。

这条处理规则也适用于多个相关向量，而不仅是两个共线向量。每个非零残差与对应的 a_k 只差先前残差的线性组合，所以加入它不会改变前缀张成空间；若残差为零，则 a_k 已在先前空间里，跳过它也不改变张成空间。归纳每一步，最终保留下来的正交向量组就张成原空间。若列表开头就是零向量，先跳过它；不能把它放在 A_1 后再计算含 $A_1^{\mathsf T}A_1$ 的分母。

16.5　QR 分解：把列写成正交方向的组合

消元把矩阵写成 LU，这里的正交化也能用矩阵语言表达。设 $A = [a_1 \ \cdots \ a_n]$ 的列线性无关，Gram–Schmidt 得到 $Q = [q_1 \ \cdots \ q_n]$。每一列 a_j 都是前 j 个正交方向的组合。

定理 16.9（QR 分解）　若 A 是 $m \times n$ 矩阵且列线性无关，则

$$A = QR,$$

其中 $Q^{\mathsf T}Q = I_n$，R 是对角线为正的 $n \times n$ 上三角矩阵。其元素为 $r_{ij} = q_i^{\mathsf T}a_j$（$i \le j$），而 $r_{ij} = 0$（$i > j$）。

思路　用 Gram–Schmidt 的每步分解 a_j：它是残差 A_j 加上先前方向的线性组合。将残差写成 $r_{jj}q_j$，即可读出 R 的第 j 列。晚于 j 的正交方向系数为零，所以 R 上三角。

证明　第 j 步的递推式可改写为

$$a_j = \sum_{i<j} \frac{A_i^{\mathsf T} a_j}{A_i^{\mathsf T} A_i} A_i + A_j = \sum_{i<j} (q_i^{\mathsf T} a_j) q_i + \|A_j\| q_j .$$

因此令 $r_{ij} = q_i^{\mathsf T} a_j$ $(i<j)$、$r_{jj} = \|A_j\| > 0$，并令 $r_{ij} = 0$ $(i>j)$，就有 $a_j = Q r_j$，其中 r_j 是 R 的第 j 列。对每一列都成立，故 $A = QR$。再左乘 $Q^{\mathsf T}$ 得 $R = Q^{\mathsf T} A$。 ∎

正交因子不改变列组合的内积，因此 QR 也给出

$$A^{\mathsf T} A = (QR)^{\mathsf T}(QR) = R^{\mathsf T} Q^{\mathsf T} Q R = R^{\mathsf T} R .$$

对任意系数向量 x，这等价于 $\|Ax\|^2 = \|Rx\|^2$。换句话说，R 保留了原矩阵各列的内积数据；第 15 讲法方程中的 $A^{\mathsf T}A$ 也可通过 $R^{\mathsf T}R$ 来表达。

上三角性反映了处理顺序：a_j 只用到第 1 到第 j 个正交方向。对角元是每一步新残差的长度，因此列独立时每个对角元都为正。

在 $A = QR$ 中，Q 记录一组新的正交归一方向，R 记录原列在这些方向中的系数。和 LU 分解的消元乘数相比，R 的上三角结构来自「第 j 列不会含有第 j 步之后才造出的方向」。乘 $Q^{\mathsf T}$ 可以一次得到全部这些系数，因为 $Q^{\mathsf T}Q = I$。

■ **例 16.12**　**一个 2×2 QR 分解**　对 $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ 执行 Gram–Schmidt，并写出 Q, R。

解　第一列 $a_1 = (1,1)$，故 $r_{11} = \|a_1\| = \sqrt{2}$、$q_1 = (1,1)/\sqrt{2}$。第二列 $a_2 = (1,0)$ 在 q_1 方向上的系数是 $r_{12} = q_1^{\mathsf T} a_2 = 1/\sqrt{2}$，残差为

$$u_2 = a_2 - r_{12} q_1 = (1/2, -1/2), \qquad r_{22} = \|u_2\| = 1/\sqrt{2}, \qquad q_2 = (1,-1)/\sqrt{2}.$$

因此

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad R = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}.$$

计算 $Q^{\mathsf T}Q = I$，并相乘得 $QR = A$。这也显示 R 的对角元记录了两次正交化残差的长度。取 $x = (2,-1)$，则 $Ax = (1,2)$，而 $Rx = (3/\sqrt{2}, -1/\sqrt{2})$；两者长度平方都为 5。这检查了 QR 对整个列组合的保距，而不只检查 Q 的列。 ♦

若列相关，完整列数就无法成为 Q 中的非零正交方向。删去零残差后，Q 有 $r = \operatorname{rank} A$ 列，$R = Q^{\mathsf T} A$ 有 r 行；它通常是上梯形矩阵。

■ **例 16.13**　**相关列的薄 QR**　对 $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$，说明为何只需两个正交方向，并写出一个分解。

解　三列分别为 $(1,0)$、$(2,0)$、$(0,1)$。第二列被第一列完全解释，残差为零；第三列给出新方向 $(0,1)$。因此可取 $Q = I_2$、$R = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$，直接相乘 $QR = A$。Q 的两列正交归一，矩阵 R 为 2×3 上梯形；A 的秩是 2，并不需要三列的 Q。 ♦

这里第二列的零残差对应「没有新的方向」，而不是算式失败。若强行给这一步归一化，会遇到零分母；若强行要求 Q 有三列，也无法得到正交归一列。定理中的列线性无关条件正是为了保证每一步都能产生一个非零的新方向。

命题 16.10（正交坐标系中的系数）　若方阵 $Q=[q_1\ \cdots\ q_n]$ 正交，任意 $v\in\mathbb{R}^n$ 在这组基下的坐标为 $c=Q^{\mathsf T}v$，即 $c_i=q_i^{\mathsf T}v$。

证明　写 $v=Qc$，左乘 $Q^{\mathsf T}$，得到 $Q^{\mathsf T}v=Q^{\mathsf T}Qc=c$。第 i 个分量正是 $q_i^{\mathsf T}v$。 ■

■ **例 16.14　由正交坐标恢复向量**　令 $Q=\frac{1}{\sqrt2}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$、$v=(3,1)$。求 v 在 Q 的列基下的坐标，再由坐标还原 v。

解　$c=Q^{\mathsf T}v=(4/\sqrt2,2/\sqrt2)=(2\sqrt2,\sqrt2)$。还原时

$$Qc=\frac{1}{\sqrt2}\begin{bmatrix}1&1\\1&-1\end{bmatrix}\begin{bmatrix}2\sqrt2\\\sqrt2\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}=v.$$

坐标 c 的数值不必等于 v 的标准坐标；它表示 v 沿着另一组正交单位方向的系数。 ♦

若 Q 只有正交归一列而不是方阵，$Q^{\mathsf T}b$ 仍给出 b 在这些列方向上的系数；$QQ^{\mathsf T}b$ 是这些分量组成的向量。缺少方阵条件时，b 可能还有列空间之外的分量。

方法：逐步清除旧方向　处理新向量 a_k 时，依次计算它在旧正交单位向量 q_i 上的系数 $q_i^{\mathsf T}a_k$，减去这些分量；剩下的残差垂直于旧空间，归一化后成为新方向。

本讲小结

- 正交归一：$q_i^{\mathsf T}q_j=\delta_{ij}$；零向量可在正交组中，不能在正交归一组中。
- 列正交归一等价于 $Q^{\mathsf T}Q=I$；方阵时 $Q^{-1}=Q^{\mathsf T}$ 且 $QQ^{\mathsf T}=I$。
- 若 $Q^{\mathsf T}Q=I$，则 $(Qx)^{\mathsf T}(Qy)=x^{\mathsf T}y$，从而长度和距离保持。
- Gram–Schmidt：$u_k=a_k-\sum_{i<k}(q_i^{\mathsf T}a_k)q_i$；若 $u_k\neq0$，取 $q_k=u_k/\|u_k\|$。
- 输入向量线性无关时残差非零，且每一步都保持前缀张成空间；线性相关时首个依赖残差为零。
- 列满秩矩阵有 $A=QR$，$Q^{\mathsf T}Q=I$，$R=Q^{\mathsf T}A$ 为对角线正的上三角矩阵。

练习

16.1 ★　判断 $q_1=\frac13(1,2,2)$、$q_2=\frac13(2,1,-2)$ 是否正交归一，写出长度平方和内积。

16.2 ★　对 $Q=\frac15\begin{bmatrix}3&-4\\4&3\end{bmatrix}$，验证 $Q^{\mathsf T}Q=I$，并计算 $Q(1,2)$。

16.3 ★ 判断命题真假：「若矩形矩阵 $R^{\mathsf T}R=I$，则 $RR^{\mathsf T}=I$。」给出反例或证明。

16.4 ★★ 对 $a=(1,1,0)$、$b=(1,0,1)$ 求 Gram–Schmidt 的正交残差与正交归一组。

16.5 ★★ 证明任意正交归一向量组线性无关。按 Session 2.4 的要求，用 $Q^{\mathsf T}$ 证明，不使用 Q^{-1}。

16.6 ★★ 对 $a=(1,0,0)$、$b=(2,0,0)$、$c=(0,1,0)$ 做正交化。指出哪个残差为零，并给出原向量组张成空间的一组正交归一基。

16.7 ★★★ 对 $a=(1,-1,0,0)$、$b=(0,1,-1,0)$、$c=(0,0,1,-1)$ 正交化，并证明两组向量都是 $d=(1,1,1,1)$ 的正交补的一组基。

16.8 ★★ 对 $A=\begin{bmatrix}1&1\\1&0\end{bmatrix}$ 求 QR 分解，验证 $Q^{\mathsf T}Q=I$ 且 $QR=A$。

16.9 ★★ 设 Q,S 是方阵正交矩阵。证明 QS 仍正交，并对本讲的 $Q=\frac15\begin{bmatrix}3&-4\\4&3\end{bmatrix}$ 与 $S=\operatorname{diag}(1,-1)$ 写出 QS。

16.10 ★★ 对上题的 Q,S 求 SQ，并说明为什么 QS 与 SQ 虽然都保持长度，却不必相等。

练习答案与提示

16.1 两个长度平方都是 $(1+4+4)/9=1$；内积为 $(2+2-4)/9=0$，所以正交归一。

16.2 两列长度平方为 $(9+16)/25=1$，内积为 $(-12+12)/25=0$，故 $Q^{\mathsf T}Q=I$。矩阵乘法给 $Q(1,2)=(-1,2)$。

16.3 假。取 $R=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}$，则 $R^{\mathsf T}R=I_2$，但 $RR^{\mathsf T}=\operatorname{diag}(1,1,0)\neq I_3$。

16.4 $A=a=(1,1,0)$，$B=b-\frac12A=(\frac12,-\frac12,1)$。其内积为 0，平方长度为 2、$3/2$。归一化得 $q_1=(1,1,0)/\sqrt2$、$q_2=(1,-1,2)/\sqrt6$。

16.5 令 $Q=[q_1\ \cdots\ q_n]$。若 $Qx=0$，则 $x=I_nx=Q^{\mathsf T}Qx=0$。所以 Q 的零空间只有零向量，列向量线性无关；没有使用 Q^{-1}。

16.6 第一个残差为 $u_1=(1,0,0)$。第二个残差是 $u_2=b-(u_1^{\mathsf T}b)u_1=0$，所以不能归一化。跳过它后，$c=(0,1,0)$ 与 u_1 正交，得到基 $(1,0,0),(0,1,0)$，它们张成 xy 平面。

16.7 $A=a=(1,-1,0,0)$，$B=b+\frac12A=(\frac12,\frac12,-1,0)$，$C=c+\frac23B=(\frac13,\frac13,\frac13,-1)$。由 $B=b+\frac12A$、$C=c+\frac23B$，两组向量张成同一空间。A,B,C 两两正交且非零，并都垂直于 d，所以线性无关。由于 $d\neq0$，$d^{\perp}$ 在 $\mathbb R^4$ 中维数为 3，故 $\{A,B,C\}$ 是其基，原向量组也因此是同一空间的基。

16.8 $q_1=(1,1)/\sqrt2$，$u_2=(1/2,-1/2)$，故 $q_2=(1,-1)/\sqrt2$。于是 $Q=\frac{1}{\sqrt2}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$，$R=\begin{bmatrix}\sqrt2&1/\sqrt2\\0&1/\sqrt2\end{bmatrix}$。两列内积为 0、长度平方为 1；相乘得 $QR=A$。

16.9 $(QS)^{\mathsf T}(QS)=S^{\mathsf T}Q^{\mathsf T}QS=S^{\mathsf T}S=I$，所以 QS 正交。具体为 $QS=\begin{bmatrix}3/5&4/5\\4/5&-3/5\end{bmatrix}$。

16.10 $SQ=\begin{bmatrix}3/5&-4/5\\-4/5&-3/5\end{bmatrix}$，它也正交，因为 S、Q 都正交，乘积仍正交。两个乘积不等，例如 $QSe_1=(3/5,4/5)$，而 $SQe_1=(3/5,-4/5)$。

第 17 讲

行列式的性质

Properties of Determinants

对应材料：MIT 18.06SC Unit II Session 2.5（Strang 第 18 讲视频）课时摘要 *Properties of determinants*；习题 18.1、18.2 及解答（Strang《Introduction to Linear Algebra》5.1 节第 10、18 题）；参考 Strang 教材 5.1 节。

本讲目标

(1) 知道行列式是给每个方阵配的一个数，它最重要的用途是判断可逆性：A 可逆当且仅当 $\det A \neq 0$。

(2) 掌握定义行列式的三条基本性质：$\det I = 1$；交换两行变号；对每一行分别是线性的。

(3) 会从这三条出发，一步步推出性质 4–8，并理解每一步用到了哪条性质。

(4) 会用消元法计算行列式：$\det A = \pm$(主元之积)，并能正确处理行交换带来的符号。

(5) 理解乘积公式 $\det AB = (\det A)(\det B)$ 与转置公式 $\det A^{\mathsf{T}} = \det A$ 的证明，会用它们算 $\det A^{-1}$、$\det A^k$、$\det(cA)$。

(6) 明白「交换次数的奇偶性由置换本身决定」为什么是这套定义不自相矛盾的关键。

17.1 行列式要回答的问题

课程进行到一半，我们暂时告别长方形矩阵，专心研究方阵。接下来的两个大主题是行列式和特征值，而特征值（第 20 讲）正是通过行列式求出来的。

行列式（determinant）给每个 $n \times n$ 矩阵 A 配一个数，记作 $\det A$ 或 $|A|$。这一个数里压缩了矩阵的许多信息，其中最重要的一条是

$$\boxed{A \text{ 可逆} \iff \det A \neq 0.}$$

2×2 的情形可以直接计算。当 $ad - bc \neq 0$ 时，乘法验证给出

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

分母 $ad - bc$ 就是 2×2 行列式：

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

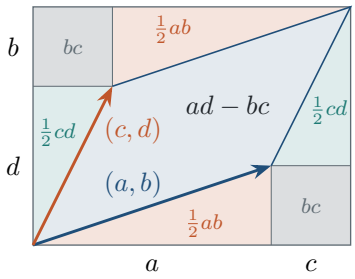
$ad-bc=0$ 时公式失效，此时两行线性相关（例如 $\left[\begin{smallmatrix}1&2\\2&4\end{smallmatrix}\right]$ 的第二行是第一行的 2 倍），矩阵确实不可逆。1×1 的情形更简单：$\det[a]=a$，a 可逆当且仅当 $a\neq 0$。

那么 3×3、$n\times n$ 呢？一般的行列式公式有 $n!$ 项，$n=10$ 时就有 3628800 项，拿它当出发点既难记又难用。Strang 的讲法是反过来走：**先不给公式，只列出三条性质**，然后证明这三条性质已经唯一地确定了行列式，并且直接给出一个高效的算法（消元）。大公式要到第 18 讲才作为副产品出现。

面积的符号　2×2 行列式有一个几何意义：以两行 (a,b)、(c,d) 为边的平行四边形的**有向面积**。图 17.1 取第一象限中 $ad>bc$ 的具体配置，此时外接矩形面积 $(a+c)(b+d)$ 减去四个直角三角形（合计 $ab+cd$）和两个小矩形（合计 $2bc$），剩下

$$(a+c)(b+d)-ab-cd-2bc=ad-bc.$$

符号记录方向：从第一行转到第二行是逆时针时为正，顺时针时为负。对 $n\times n$ 矩阵，$|\det A|$ 是以各行为棱的平行多面体的体积。这一事实第 19 讲再证明；本讲把它当作帮助理解三条性质的图像。



图中取 $(a,b)=(3,1)$，$(c,d)=(1,2)$：

$4\cdot 3-3-2-2=5$，

恰好等于 $ad-bc=3\cdot 2-1\cdot 1=5$。

图 17.1　2×2 行列式是平行四边形的有向面积：外接矩形减去四个三角形和两个小矩形。

记号　$|A|$ 中的竖线**不是绝对值**：$\begin{vmatrix}0&1\\1&0\end{vmatrix}=-1$。另外，行列式**只对方阵有定义**。$2\times 3$ 矩阵没有行列式，这一点在 §17.4 识别错误「证明」时会用到。

17.2　三条基本性质

下面三条性质就是行列式的定义。为了叙述方便，把 A 的各行记作 $\boldsymbol{r}_1,\dots,\boldsymbol{r}_n$，并把 $\det A$ 写成 $\det(\boldsymbol{r}_1,\dots,\boldsymbol{r}_n)$。

定义 17.1（规范化、交替性与逐行线性）　在固定阶数 $n \geq 1$ 下，行列式是定义在所有 $n \times n$ 实矩阵上的实值函数，要求：

(1) $\det I_n = 1$；

(2) 交换任意两个不同行，函数值乘以 -1；

(3) 固定其他行时，对剩下的一行是线性的，即

$$\det(\ldots, su + tv, \ldots) = s \det(\ldots, u, \ldots) + t \det(\ldots, v, \ldots).$$

第三条称为**多重线性**（multilinearity），强调每次只改变一个行变量。

这里 s, t 可以为零，也可以为负。对 1×1 矩阵，规范化与线性给 $\det[a] = a \det[1] = a$，第二条没有可交换的行，因而无需另加限制。对于 $n \geq 2$，交换与线性共同约束了不同的行，不能把它理解成矩阵全部元素的普通线性函数。

逐行线性不等于对整个矩阵线性　一般没有 $\det(A + B) = \det A + \det B$。取 $A = B = I_2$，左边为 $\det(2I_2) = 4$，右边为 2。逐行线性只允许在其他行固定时拆开一行；同时把两行都拆开，会出现四项而非两项。

■ **例 17.1　只拆一行，与同时拆两行**　用 2×2 公式验证

$$\det \begin{bmatrix} 3 & 5 \\ 2 & 1 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} + \det \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}.$$

再说明为何不能把两个矩阵整块相加后只取两个行列式。

解　三项分别为 $3 - 10 = -7$、$1 - 4 = -3$、$2 - 6 = -4$，确有 $-7 = -3 - 4$。第二行始终是 $(2, 1)$，所以这里正好在定义允许的范围内。若两个矩阵各行是 u_1, u_2 与 v_1, v_2，逐行展开给四项：$\det(u_1, u_2)$、$\det(u_1, v_2)$、$\det(v_1, u_2)$、$\det(v_1, v_2)$。中间两项没有理由消失；在 $A = B = I_2$ 的反例中它们分别等于 1，补齐了被漏掉的 2。 ♦

17.2.1　先从三条推四条

交换相同的两行看起来什么也没有做，第二条却要求变号。两个描述同时成立，就把原来的数限制为零。这个观察是后面所有消元规则的依据。

命题 17.2（重复行、零行与行替换）　满足定义 17.1 的函数具有下列性质：

(1) 两行相同，行列式为零；两行成比例，也为零。

(2) 一行全零，行列式为零。

(3) 对 $i \neq j$ 作 $R_i \leftarrow R_i + tR_j$，行列式不变。

(4) 一行乘 t，行列式乘 t；所有行同时乘 t，行列式乘 t^n。

证明　若第 i,j 行相同，交换后矩阵相同，却使值 d 变成 $-d$，于是 $d=-d$，在实数中得 $d=0$。若其中一行是另一行的 t 倍，先用该行的线性把 t 提出，剩下重复行，所得值仍为零。零行可以写成任意行向量的 0 倍；逐行线性直接给 0，包括阶数为 1 的情形。行替换则在被替换行上展开：

$$\det(\dots,R_i+tR_j,\dots,R_j,\dots)=\det(\dots,R_i,\dots,R_j,\dots)+t\det(\dots,R_j,\dots,R_j,\dots).$$

第二项含重复行，所以为零。最后一条的第一部分就是逐行线性；依次从 n 行中各提出一个 t，共得到 n 个因子，因此是 t^n。这里不需除以 t，所以 $t=0$ 也已包含。 ■

条件 $i\neq j$ 不能删掉：若同一行加上自身，就把该行乘了 2，行列式一般也乘 2，并非不变。区别由「另外一行」决定，不能只记住「加减一行不变」这句不完整的口令。

想一想　矩阵每个元素都乘 -1 时，行列式一定变号吗？例如 $-I_2$ 和 $-I_3$ 分别怎样？
答　不一定。每一行各提出 -1，得到 $\det(-A)=(-1)^n\det A$。偶数阶保持，奇数阶变号；因而 $\det(-I_2)=1$，$\det(-I_3)=-1$。变的是一行还是全部行，必须先辨明。

■ **例 17.2　相关行的零值证据**　不作三阶展开，求 $K=\begin{bmatrix}1&2&3\\2&4&6\\0&1&1\end{bmatrix}$ 的行列式，并写出一种保持值的操作把零值显示出来。

解　第二行是第一行的两倍，所以重复行规则与线性已经给 $\det K=0$。也可作 $R_2\leftarrow R_2-2R_1$，得到中间全零行，操作不改变行列式，零行规则同样给零。两种写法的依据相同；只是后一种把「相关」转成了一个可以直接检查的零行。 ♦

这说明只要某一行是其他行的线性组合，行列式就为零：将这个组合从该行减去，依次进行不改值的行替换，最终出现零行。没有要求组合系数都非零；若该行本来为零，也已由零行规则处理。

17.2.2　定义中的存在与唯一要分别回答

三条规则之间能够推导出许多结论，仍有两个不同的问题：是否真有函数满足它们；若有，是否只有一个。第 18 讲用排列的带符号公式完整构造这个函数。本讲的所有推导都对任何满足三条规则的函数成立，不把尚未构造的公式偷偷当作定义。

命题 17.3（三条规则至多确定一个函数）　固定 $n\geq 1$，满足规范化、交换变号与逐行线性的函数至多有一个。

证明　设 f,g 都满足三条规则，对任意 A 作同一串消元与必要的行交换，把它变为行阶梯形。两函数对每次交换都变号、对每次行替换都不变。若缺少主元，阶梯形中有零行，两个函数均取零；若有全部 n 个主元，就能继续作行替换，消去主元上方的项，得到非零对角矩阵 D。对 $D=\operatorname{diag}(d_1,\dots,d_n)$，逐行提出 d_i 后得到单位矩阵，故 $f(D)=d_1\cdots d_n=g(D)$。沿同一串变换倒推，

$f(A) = g(A)$。前一种情形的零行结论与后一种情形的对角结论都只用三条规则，所以没有预先引用唯一性。 ■

消元中的主元可以任选，但这个命题保证：只要每一步遵守规则，最后算出的值相同。它还没有证明「任选一串交换」的奇偶数本身不会矛盾；相容性要由下一讲的显式构造保证，原课摘要末尾留下的正是这一点。

17.3　三角矩阵与消元计算

上三角矩阵 $T = \begin{bmatrix} -2 & 3 & 1 \\ 0 & 4 & 5 \\ 0 & 0 & 3 \end{bmatrix}$ 含有三个非零的上方元素，但行列式最终只留下 $(-2)\cdot 4\cdot 3 = -24$。为什么上方元素完全不影响它？从最后一行开始拆开各行，可以看到所有额外项如何消失。

> **定理 17.4（三角矩阵的行列式）**　上三角或下三角 $n \times n$ 矩阵的行列式，等于主对角元之积；该结论允许一个或多个对角元为零。

思路　把每一行拆成坐标向量的线性组合。重复选择同一个坐标方向的项为零；三角形的零位置又强迫没有重复的选择恰好来自全部对角元。整个过程不除以对角元，因此零对角元无需另造极限论证。

证明　先设 T 上三角，标准行向量记为 $e_1^{\mathsf{T}}, \dots, e_n^{\mathsf{T}}$。第 i 行是 $\sum_{j=i}^{n} t_{ij} e_j^{\mathsf{T}}$。对每一行依次用线性展开，得到有限多项，每项的系数为某个 $t_{1j_1}\cdots t_{nj_n}$，其余部分是以 $e_{j_1}^{\mathsf{T}}, \dots, e_{j_n}^{\mathsf{T}}$ 为行的行列式，且每个 $j_i \ge i$。若有两个 j_i 相同，则两行相同，该项为零。若没有重复，$j_1, \dots, j_n$ 必恰为 $1, \dots, n$ 的一次排列，因此 $\sum_i j_i = \sum_i i$。由于每项 $j_i - i \ge 0$，它们的和为零，迫使每个 $j_i = i$。所以只剩对角选择，对应的坐标矩阵恰为 I_n，值为 1。最终 $\det T = t_{11}\cdots t_{nn}$。下三角时，每行只选 $j_i \le i$；没有重复的选择仍满足两个和相等，非负数 $i - j_i$ 的和为零，再次得到每个 $j_i = i$，结论相同。这一论证也包含 $n = 1$，并且没有假设系数非零。 ■

■ **例 17.3　零对角元不允许被约掉**　求 $T = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix}$ 的行列式。为什么不能为了套用「消成对角矩阵」而直接除以第一个对角元？

解　三角定理直接给 $\det T = 0\cdot 2\cdot 4 = 0$。此处第一列全零，所以不存在能换入左上角的非零主元；除以零没有意义。若确实要制造一个零行，可先把第二行减第三行的 $3/4$ 倍，得到 $(0, 2, 0)$，再把第一行减第三行的 $1/2$ 倍并减新第二行的 $1/2$ 倍，得到零行。这些除法的分母是 $4, 2$，均非零。三角公式的直接证明更短，也不会把零对角元遗漏。 ♦

定理 17.5（消元、奇异性与行列式）　对实方阵 A，若经过 s 次行交换与若干行替换得到上三角 U，则

$$\det A = (-1)^s \prod_{i=1}^{n} u_{ii}.$$

此外，$\det A \neq 0$ 当且仅当 A 可逆。

证明　每次行交换使值变号，每次行替换不改值，因而 $\det U = (-1)^s \det A$；由于 $(-1)^{2s} = 1$，乘回同一符号后便得所写公式。对可逆的 A，消元每一阶段都能在剩余部分找到足够的主元：若某列在所需位置及以下全零，第 2 讲已经构造出一个非零齐次解，与可逆性矛盾。因此可以通过必要的换行得到全部非零主元，三角公式给 $\det A \neq 0$。反过来，若 A 不可逆，逐列作完整的行阶梯消元，必要时跳过没有主元的列。若有 n 个主元，回代本应给每个右端唯一解，矩阵会可逆，矛盾。所以主元数少于 n，阶梯形至少有一条零行。行替换与交换只会保持值或变号，零行的行列式为零，于是 $\det A = 0$。两个方向合起来得到等价关系。 ■

公式中的 s 只数交换次数。如果计算时还把一行乘了非零数 t，就必须另外记录这个因子；直接把最终对角元相乘而不补偿它，会得到被放大或缩小后的矩阵的值。若使用归一化主元算法，把每个主元行都除以主元，则最终对角元全部为 1，原行列式的信息正藏在那些被除掉的数中。

■ **例 17.4　沿用第 2 讲的三元系统**　用消元求 $A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix}$ 的行列式，并比较「第一行乘 2」与「整个矩阵乘 2」。

解　按 $R_2 \leftarrow R_2 - 3R_1$、$R_3 \leftarrow R_3 - 2R_2$ 得

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix}.$$

没有交换，也没有缩放，故 $\det A = 1 \cdot 2 \cdot 5 = 10$。仅第一行乘 2 得 20；三行一起乘 2 得 $2^3 \cdot 10 = 80$。两个计算的差别是提出了一个因子还是三个因子，不能用「矩阵乘 2」的模糊描述掩盖它。 ♦

■ **例 17.5　需要一次交换的消元**　求 $S = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 0 & 3 \end{bmatrix}$ 的行列式。若把全部主元直接相乘，会错在哪里？

解　先交换前两行得到 $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$，再作 $R_3 \leftarrow R_3 - 2R_1$，第三行变成 $(0, -2, 3)$。接着 $R_3 \leftarrow R_3 + R_2$，得到上三角矩阵 $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}$。它的行列式为 8，原矩阵经过一次交换才到这里，所以

$\det S = -8$。主元之积给的是换行后的值，必须乘上交换记录的 (-1) 才能回到原矩阵。 ♦

一份随消元更新的记录　从因子 $q = 1$ 开始。交换一次就令 q 变号；一行乘 $t \neq 0$ 就把最终值除以 t；加上其他行倍数不改记录。结束后把当前对角元之积乘上所有补偿因子，才得到原矩阵的行列式。对大矩阵，记录操作比展开 $n!$ 项更实用。

图 17.2 将三种操作对解集和行列式的影响分开。三种都保持方程组的解，只要右端同步操作；只有行替换在一般情况下保持行列式数值。

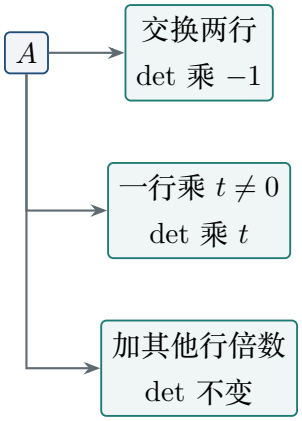


图 17.2　三种可撤销行操作都保持同步变换后的方程组解集，但对行列式具有三种不同的数值规则。

图里的 $t \neq 0$ 是为了保证缩放可撤销；行列式的缩放公式本身对 $t = 0$ 也正确。只是把一行乘零以后，原值无法从新值恢复，所以计算算法不能把它当作可逆步骤。

17.4　乘积公式：把第二个矩阵固定

令 $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$、$B = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$。直接相乘得到 $CB = \begin{bmatrix} 4 & 6 \\ 10 & 12 \end{bmatrix}$，三者的行列式依次为 $-2, 6, -12$。乘积的值确是前两者之积；证明要解释为什么所有同阶方阵都如此。

定理 17.6（乘积公式）　若 A, B 都是 $n \times n$ 实矩阵，则

$$\det(AB) = (\det A)(\det B).$$

思路　先把 B 固定。若 B 可逆，把函数 $A \mapsto \det(AB)$ 除以非零常数 $\det B$，它就重新满足三条定义性质，由唯一性只能等于 $\det A$。若 B 不可逆，两个乘积的零值直接从非零齐次解得到，不能除以零。

证明　若 $\det B \neq 0$，定义 $f(A) = \det(AB)/\det B$。当 $A = I_n$ 时，$f(I_n) = 1$。交换 A 的两行，AB 的同两行也交换，所以 f 变号。固定 A 的其他行，把一行写成 $su + tv$，乘 B 后该行成为 $s(uB) + t(vB)$，其余行不变。由行列式逐行线性，f 在这一行也线性。因此 f 满足全部三条规则，由命题 17.3，$f(A) = \det A$。乘回 $\det B$ 得结论。若 $\det B = 0$，奇异性判别说明 B 不可逆，因而存在

$v \neq 0$ 使 $Bv = 0$。于是 $(AB)v = A(Bv) = 0$，所以 AB 也不可逆，$\det(AB) = 0 = (\det A)(\det B)$。两种情况覆盖所有 B，没有假设 A 也可逆。 ■

这个证明只要求矩阵同阶，并未要求 $AB = BA$。事实上，将 A, B 对调还能得到 $\det(AB) = \det(BA)$，但这只是两个矩阵被映射成同一个数，不能反推两个矩阵相等。行列式只记录一部分信息；例如 I_2 与 $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ 行列式都为 1，矩阵本身不同。

长方形矩阵不能套同阶乘积式　取 $A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$、$B = \begin{bmatrix} 1 & 0 \end{bmatrix}$。$AB$ 是行列式为 0 的二阶矩阵，$BA = [1]$ 的行列式为 1。这里 A, B 各自没有行列式，不能写出 $\det A \det B$；「两种顺序的乘积都恰为方阵」不足以满足定理条件。

推论 17.7（逆、幂与整体系数）　若 A 可逆，则 $\det(A^{-1}) = 1/\det A$。对整数 $k \geq 1$ 有 $\det(A^k) = (\det A)^k$；若允许 $k < 0$，须再假设 A 可逆。任意实数 t 满足 $\det(tA) = t^n \det A$。

证明　$A^{-1}A = I_n$ 给 $\det(A^{-1}) \det A = 1$；由于 $\det A \neq 0$，可除得第一式。正整数幂对 k 归纳，每多乘一个 A 就多出一个因子 $\det A$。负整数幂是 A^{-1} 的正整数幂，结合第一式即得。tA 的每行都乘了 t，逐行提出 n 次得到最后的式子；$t = 0$ 不涉及除法，仍然成立。$A^0 = I_n$ 时行列式为 1，单独按单位阵理解即可。 ■

■ **例 17.6　从一个二阶数值算许多个**　对 $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$，求 $\det C^{-1}$、$\det C^2$ 与 $\det(2C)$。不用重新展开三个行列式。

解　先算 $\det C = -2 \neq 0$，所以逆存在；公式依次给 $-1/2, 4, 4(-2) = -8$。为了核对逆的符号，二阶逆公式给 $C^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$，其行列式为 $1 - 3/2 = -1/2$。平方的行列式变成正数，并不表示 C^2 的每个元素都正或它一定正定；这里得到的只是一个标量等式。 ♦

17.5　转置、列规则与置换奇偶性

交换 S 的行会变号，那么交换它的列呢？直接在二阶公式里可以看到相同规律，但要推广到所有阶数，还需要把「行」与「列」联系起来。转置正好交换这两个角色。

定理 17.8（转置不改变行列式）　每个实方阵均满足 $\det(A^{\mathsf{T}}) = \det A$。所以三条行规则及其推论都可逐字改成列规则。

思路　把消元的每一步撤销，写 $A = F_1 \cdots F_s U$。每个 F_i 只是交换矩阵或一步行替换的逆，它与转置的行列式相同；U 和 U^{T} 分别是上、下三角，主对角线相同。对转置乘积使用乘积公式，不

必假设 U 可逆。

证明　完整行阶梯消元使用若干行交换与行替换，得到 $E_s\cdots E_1A=U$，其中 U 上三角；即使主元不足，仍可跳过无主元列，所得阶梯形在对角线下方也全为零。每步可撤销，令 $F_i=E_i^{-1}$，得 $A=F_1\cdots F_sU$。如果 F_i 是行替换矩阵，它在单位阵的一个非对角位置放某个数，其本身与转置各为上三角或下三角矩阵，对角元全为 1，所以两者行列式都为 1。如果 F_i 是交换两行的矩阵，它与转置相同，且行列式为 -1。又因为 $U^{\mathsf T}$ 与 U 的对角元相同，三角公式给 $\det(U^{\mathsf T})=\det U$。于是

$$\det(A^{\mathsf T})=\det(U^{\mathsf T}F_s^{\mathsf T}\cdots F_1^{\mathsf T})=\det(U^{\mathsf T})\prod_{i=1}^{s}\det(F_i^{\mathsf T})$$

$$=\det U\prod_{i=1}^{s}\det F_i=\det A.$$

最后一步只交换实数因子的次序，没有声称矩阵因子可以交换。整个证明对奇异矩阵同样成立。 ■

例如 $S=\begin{bmatrix}0&2&1\\1&1&0\\2&0&3\end{bmatrix}$ 与 $S^{\mathsf T}=\begin{bmatrix}0&1&2\\2&1&0\\1&0&3\end{bmatrix}$ 都有行列式 -8；两矩阵并不相等。由转置规则，若一列全零，行列式为零；若两列相同，行列式为零；把某列加上另一列的倍数，行列式不变。用列操作时无需把整个矩阵真的抄成转置，只需知道这些规则已有证明。

17.5.1　为什么同一个置换不能既奇又偶

把 I_n 的行重新排列得到置换矩阵 P，若使用 s 次交换，规则给 $\det P=(-1)^s$。但如果同一个排列能用奇、偶两种次数实现，这两份计算将相互矛盾。不能仅用「行列式已定义」把这个疑问绕过去；原课摘要末尾也特意提出了它。先看不涉及行列式的多项式

$$\Delta(x_1,\dots,x_n)=\prod_{1\le i<j\le n}(x_j-x_i).$$

交换相邻两个变量 x_k,x_{k+1} 时，因子 $x_{k+1}-x_k$ 变号；它们与任何其他变量所形成的两个因子互换位置，乘积不变；其余因子不变。所以整式变号。交换位置 $i<j$ 的两个变量，可先把第 i 个依次移到第 j 个位置，再把原第 j 个向左移回第 i 个位置，共用 $(j-i)+(j-i-1)=2(j-i)-1$ 次相邻交换，仍为奇数次，所以任意单次交换都使 Δ 变号。若两串交换产生同一个排列，最终变量排列相同，得到的多项式也相同。取 $x_i=i$，此时 $\Delta=\prod_{i<j}(j-i)>0$，不为零。两串交换次数为 s,t 时必有 $(-1)^s\Delta=(-1)^t\Delta$，从而 $(-1)^s=(-1)^t$。因此交换次数的奇偶性只由排列决定；不同分解的次数可以不同，奇偶性不能不同。$n=1$ 时空乘积取 1，唯一排列也没有奇偶冲突。

想一想　把三行依次循环为「原第 2 行、原第 3 行、原第 1 行」，是一次循环，所以一定变号吗?

答　不能按「循环的次数」记符号。这个循环可先交换前两行，再交换后两行，共两次交换，因而保持符号。对应的置换矩阵为 $\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$，行列式为 1。一次任意置换与一次交换不是同一个操作。

奇偶性消除了交换规则的一个潜在矛盾。下一讲还会用它构造完整的排列公式，验证三条基本性质确实同时成立，从而把本讲的「至多一个」补成「恰有一个」。本讲计算不要求记住那个长公式，下一讲的公式却说明消元算法所计算的量有可靠的定义。

17.6　两道原题：行和与 Vandermonde 矩阵

如果一个三阶矩阵的每行元素和都为零，逐个做消元可能相当费时；但把列向量 $(1,1,1)$ 代进去，每行输出立刻变成零。把「各行共同满足的关系」转换为非零齐次解，往往比计算更能直接说明行列式为何为零。

■ **例 17.7**　**原习题 18.1：各行元素之和**　设 A 为 $n\times n$ 实矩阵，$n\ge 1$。若每行元素和为零，证明 $\det A=0$；若每行元素和为 1，证明 $\det(A-I_n)=0$。后一条件能否推出 $\det A=1$?

解　选取非零向量 $e=(1,\dots,1)^{\mathsf T}$。第一种情况下，$Ae$ 的第 i 个分量为 $\sum_j a_{ij}=0$，所以 $Ae=0$；A 具有非零齐次解，不可逆，行列式为零。这里 $n\ge 1$ 保证 e 不是空向量，也确实非零。第二种情况下 $Ae=e$，所以 $(A-I_n)e=0$，同样得到 $\det(A-I_n)=0$。但不能将行列式错误拆成 $\det A-\det I_n$。例如 $P=\begin{bmatrix}0&1\\1&0\end{bmatrix}$ 每行和为 1，行列式却是 -1。也可取 $\begin{bmatrix}1&0&0\\1/2&1/2&0\\0&1/4&3/4\end{bmatrix}$，各行和同为 1，行列式为 $3/8$。因此行和并不直接等于行列式。　♦

这个例题可以原样推广：若每行和都是实数 t，则 $(A-tI_n)e=0$，故 $\det(A-tI_n)=0$。当 $t=0$，直接判定 A 奇异；当 $t\neq 0$，判定的是 $A-tI_n$，并非 A。第 20 讲会给 $Ae=te$ 这样的关系一个统一名称；本讲只需要矩阵乘法与可逆性即可完成证明。

■ **例 17.8**　**原习题 18.2：三阶 Vandermonde 行列式**　对任意实数 a,b,c，证明

$$\det\begin{bmatrix}1&a&a^2\\1&b&b^2\\1&c&c^2\end{bmatrix}=(b-a)(c-a)(c-b).$$

还要处理其中两个参数相等的情形，不能在没有检查时直接除以 $b-a$。

解　先做行替换，再用乘法提取，而不作除法。第二、三行各减第一行，得到

$$\det\begin{bmatrix}1 & a & a^2\\ 0 & b-a & b^2-a^2\\ 0 & c-a & c^2-a^2\end{bmatrix}.$$

由于 $b^2-a^2=(b-a)(b+a)$、$c^2-a^2=(c-a)(c+a)$，逐行线性给出

$$(b-a)(c-a)\det\begin{bmatrix}1 & a & a^2\\ 0 & 1 & b+a\\ 0 & 1 & c+a\end{bmatrix}.$$

这个等式在 $b-a=0$ 或 $c-a=0$ 时也成立，因为它把一个行向量写成标量与向量的乘积，并未除以该标量。剩余矩阵的第三行减第二行，变为上三角，最后一个对角元是 $(c+a)-(b+a)=c-b$，其行列式为 $c-b$，所求公式随即成立。若任意两个参数相等，原矩阵有相同行，右边也有对应的零因子，两边均为零。例如 $(a,b,c)=(1,2,4)$ 时，结果为 $1\cdot 3\cdot 2=6$；交换 b,c 后，矩阵交换两行，右边三个差的乘积也变号。这个例子同时核对了数值、零值边界与符号，而不仅是展开一个三阶多项式。 ♦

注 17.1　原题解的一个中间矩阵在第一列误保留了 1；实施第三行减第一行以后，该位置必须为 0。抄写中间矩阵时，既要检查变化的那一列，也要检查整条被替换的行。上面的推导在每一次操作后都保留了完整行，因此后续三角公式可以直接使用。

■ **例 17.9　参数使矩阵何时失去可逆性**　对 $A_t=\begin{bmatrix}1 & 1 & 1\\ 1 & t & 1\\ 1 & 1 & t\end{bmatrix}$，求全部使它不可逆的实数 t。对可逆参数，写出其行列式的符号。

解　第二、三行分别减第一行，得到 $\begin{bmatrix}1 & 1 & 1\\ 0 & t-1 & 0\\ 0 & 0 & t-1\end{bmatrix}$，没有交换与缩放，所以 $\det A_t=(t-1)^2$。它恰在 $t=1$ 时为零，此时原矩阵三行全相同，确实不可逆；$t\neq 1$ 时行列式严格为正，所以可逆。即使 $t<1$ 两个后续主元都负，它们之积仍为正，不能由单个主元的符号判断总符号。 ♦

前后联系　本讲建立了行列式与消元、齐次方程、可逆性的联系；下一讲把逐行线性展开为排列公式与余子式展开。第 19 讲再用它求逆、写 Cramer 公式并证明体积含义。使用这些后续公式时，仍要先问矩阵是否为方阵、需要除的行列式是否非零，以及改变行或列后数值怎样变化。

本讲小结

- **三条规则**是 $\det I_n = 1$、交换两行变号、固定其他行后对一行线性；逐行线性不等于 $\det(A+B) = \det A + \det B$。
- **消元规则**：重复行、成比例行或零行导致零值；一行加其他行倍数不变，一行乘 t 乘 t，全部行乘 t 乘 t^n。
- **三角公式**是 $\det T = \prod_i t_{ii}$，允许对角元为零；消元时每次交换另乘 -1，缩放另作补偿。
- **可逆判别**是 A 可逆当且仅当 $\det A \neq 0$；找到非零 v 满足 $Av = 0$ 就可证明 $\det A = 0$。
- **乘积与逆**：同阶方阵满足 $\det(AB) = \det A \det B$；可逆时 $\det(A^{-1}) = 1/\det A$；正整数幂满足 $\det A^k = (\det A)^k$。
- **行列对称**：$\det(A^{\mathsf{T}}) = \det A$，所以交换、线性、零向量与行替换规则也适用于列。置换的符号取决于交换次数的奇偶性。
- **结构计算**：行和为 t 给 $\det(A - tI_n) = 0$；三阶 Vandermonde 值为 $(b-a)(c-a)(c-b)$，提因子无需除以参数差。

练习

17.1 ★ 判断真假并证明或给反例：(a) $\det(A+B) = \det A + \det B$；(b) 交换两列会变号；(c) 方阵有一个元素为零就不可逆；(d) 同阶方阵 AB 与 BA 的行列式相等，所以 $AB = BA$。

17.2 ★ 求 $T = \begin{bmatrix} 2 & 7 & -1 \\ 0 & -3 & 5 \\ 0 & 0 & 4 \end{bmatrix}$、$T^{\mathsf{T}}$、$-T$、$2T$ 的行列式。每一步说明使用哪条规则。

17.3 ★★ 只用行操作求 $\det \begin{bmatrix} 0 & 1 & 2 \\ 2 & 3 & 4 \\ 1 & 0 & 1 \end{bmatrix}$，记录交换次数，不许在第一步除以零。

17.4 ★ 已知 $C = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$。把第一行乘 2，再把第二行加第一行的 3 倍，最后交换两行，求最终行列式。第二步使用的是当前第一行还是原第一行？两种读取方式会不会改变本题最终的行列式？

17.5 ★★ 对 $K = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 3 & 3 & 3 \end{bmatrix}$，不用展开证明行列式为零。把你的理由推广到任意一行是其他行线性组合的方阵。

17.6 ★★ 设 A 每行元素和为零，证明 $\det A = 0$；若每行和为 1，证明 $\det(A - I) = 0$ 并给出 $\det A \neq 1$ 的例子。分别用非零齐次解与列相加两种方法证明第一问。

17.7 ★★ 设 A, B 是三阶可逆实矩阵，$\det A = -2$、$\det B = 2$。求 $\det(A^{-1}B^3A^{-1})$。若改为 $A^{-1}(A+B)A$，能否只由这两个已知行列式求出结果？给出理由。

17.8 ★★ 求 $\det\begin{bmatrix}1&-1&1\\1&2&4\\1&5&25\end{bmatrix}$，解释交换后两行怎样影响结果。再说明当 Vandermonde 题中 $a=b$ 时，「从第二行提出 $b-a$」为何仍合法。

17.9 ★★ 设 $P=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$。用两次行交换求 $\det P$，再求 $\det(P^{\mathsf T})$、$\det(P^{-1})$。用多项式 Δ 的变号性质解释为什么换一种交换分解不会改变答案。

17.10 ★★★ 对例 17.9 的 A_t，求不可逆的全部实参数；当 $t=1$ 时写出非零齐次解。若只知某对称方阵行列式为正，是否可推出对所有非零 x 都有 $x^{\mathsf T}Ax>0$（称为正定）？给出反例。

练习答案与提示

17.1 (a) 假，取 $A=\operatorname{diag}(1,0)$、$B=\operatorname{diag}(0,1)$，两个行列式都为零，和为单位阵，行列式为 1。(b) 真，对转置使用交换行规则，再转回去。(c) 假，I_2 含零元素却可逆；判别的是行列式，不是某个元素。(d) 前半真、推论假：取 $A=\begin{bmatrix}1&1\\0&1\end{bmatrix}$、$B=\operatorname{diag}(1,2)$，两个乘积分别为 $\begin{bmatrix}1&2\\0&2\end{bmatrix}$ 与 $\begin{bmatrix}1&1\\0&2\end{bmatrix}$，都取行列式 2，却不相等。

17.2 $\det T=2(-3)4=-24$；转置不改变值，所以 $\det(T^{\mathsf T})=-24$。三阶整体乘 -1 的因子为 $(-1)^3$，故 $\det(-T)=24$；整体乘 2 的因子为 2^3，故 $\det(2T)=-192$。这里对角元相乘只用了三角性质，上方的 $7,-1,5$ 不影响结论。

17.3 先交换第一、三行，得到 $\begin{bmatrix}1&0&1\\2&3&4\\0&1&2\end{bmatrix}$。第二行减第一行的两倍得 $(0,3,2)$；第三行减新第二行的 $1/3$ 倍得 $(0,0,4/3)$。上三角值为 $1\cdot3\cdot4/3=4$，有一次交换，原值为 -4。所有实际除法都使用非零的 $1,3$，未用原来的零主元。

17.4 原值 $\det C=5-6=-1$。第一步乘 2 后为 -2，第二步行替换不变，最后交换变成 2。按指定的顺序，第二步应使用已经乘 2 的当前第一行。若改用原第一行，它等于当前第一行的 $1/2$，因此这一步实际上是另一种行替换，矩阵元素会变，但行列式仍不变；本题最终值仍为 2。不能因为最终行列式相同，就说两种操作产生相同矩阵。

17.5 第三行等于前两行之和，作 $R_3\leftarrow R_3-R_1$、再作 $R_3\leftarrow R_3-R_2$，得到零行且不改值，所以为零。一般若 $R_i=\sum_{j\neq i}c_jR_j$，依次从第 i 行减去各个 c_jR_j，参照行均未改变，最终把该行化零，所以行列式为零。这里每步只改一行，不把多条同时更新的等式混作可逆操作。

17.6 令 $e=(1,\dots,1)^{\mathsf T}\neq0$。行和为零给 $Ae=0$，故 A 奇异。另一方法是把第 2 到第 n 列依次加到第 1 列，最终第一列正好是各行元素的和，所以全零，每步不改行列式，结果为零；$n=1$ 时本来就是零矩阵，直接成立。行和为 1 时 $(A-I)e=0$，推出的是 $\det(A-I)=0$。反例取二阶交换矩阵，它每行和为 1 而行列式为 -1；不能把行列式对矩阵的加减当作线性运算。

17.7 乘积与逆公式给 $(-2)^{-1}\cdot2^3\cdot(-2)^{-1}=2$。第二个表达式的行列式为 $\det(A+B)$，共轭两端的行列式互相抵消，但和的值并不由两个单独行列式决定。例如固定 $A=\operatorname{diag}(-2,1,1)$，分别取 $B_1=\operatorname{diag}(2,1,1)$ 与 $B_2=\operatorname{diag}(1,2,1)$，两者行列式同为 2；$A+B_1=\operatorname{diag}(0,2,2)$ 的行列式为零，而 $A+B_2=\operatorname{diag}(-1,3,2)$ 的行列式为 -6，故已知条件不足。

17.8 参数为 $a=-1,b=2,c=5$，结果为 $(2+1)(5+1)(5-2)=3\cdot6\cdot3=54$。交换后两行变成 -54。当 $a=b$，第二行减第一行后是零向量，它确实等于 $0\cdot(0,1,b+a)$；线性给出「左边等于 0 乘某个

行列式」，仍然成立。非法的是先除以 $b-a=0$，而非写出合法的乘法分解。

17.9 从 I_3 先交换第一、二行，再交换第二、三行就得 P，两次变号抵消，故 $\det P=1$。转置公式与逆公式分别给 $\det(P^{\mathsf{T}})=1$、$\det(P^{-1})=1$。任意交换分解都使 Δ 每次乘 -1；同一排列对非零的 $\Delta(1,2,3)$ 给同一最终值，故交换次数必同奇偶，不会把本题的 1 算成 -1。

17.10 行替换给 $\det A_t=(t-1)^2$，因此不可逆当且仅当 $t=1$。此时 $u=(1,-1,0)^{\mathsf{T}}\neq 0$ 满足 $A_1u=0$，是一个具体证据。正行列式不足以推出正定：对称矩阵 $-I_2$ 的行列式为 $1>0$，但任意非零 x 有 $x^{\mathsf{T}}(-I_2)x=-\|x\|^2<0$。行列式乘上了所有方向的因素，单独一个乘积的符号无法保证每个二次型方向都为正。

第 18 讲

行列式公式与代数余子式

Determinant Formulas and Cofactors

对应材料：MIT 18.06SC（2011 秋）Unit II，Session 2.6「Determinant Formulas and Cofactors」课时摘要，对应 Strang 第 19 讲视频；Session 2.6 习题 19.1、19.2 及解答；Strang《Introduction to Linear Algebra》第 5.1–5.2 节；ZoomNotes 第 5.1 节及第 5.2 节的余子式部分。

本讲目标

(1) 只用行列式的三条基本性质，推出 2×2、3×3 行列式的显式公式，看清 $n!$ 项是怎么来的，以及其余的项为什么都是 0。

(2) 理解排列的逆序数与奇偶性，写出一般的「大公式」$\det A=\sum_{\sigma}\operatorname{sgn}(\sigma)\,a_{1\sigma(1)}\cdots a_{n\sigma(n)}$，并用它证明 $\det A^{\mathsf T}=\det A$。

(3) 掌握余子式与代数余子式 $C_{ij}=(-1)^{i+j}M_{ij}$，会沿任意一行或一列做代数余子式展开，并会根据零元素的位置选择展开方向。

(4) 用代数余子式展开得到三对角矩阵行列式的递推式，算出全 1 三对角矩阵行列式以 6 为周期、$-1,2,-1$ 矩阵行列式等于 $n+1$。

(5) 会比较三种算法（主元乘积、大公式、代数余子式）的适用场合与计算量。

18.1 从三条性质出发：2×2 公式

第 17 讲用三条性质定义了行列式：

(1) $\det I=1$；

(2) 交换两行，行列式变号；

(3) 行列式对每一行分别是线性的（其余各行固定不动）。

由它们推出了一串推论：两行相同则 $\det=0$；从一行减去另一行的倍数不改变 $\det$；有零行则 $\det=0$；三角矩阵的行列式是对角元之积；A 奇异当且仅当 $\det A=0$；$\det AB=(\det A)(\det B)$；$\det A^{\mathsf T}=\det A$。

其中「三角矩阵的行列式是对角元之积」给出了第一种算法：消元化成上三角矩阵 U，再把主元乘起来（有行交换时添一个 $\pm$ 号）。这个方法快，但它不是一个公式：看着 A 的元素，我们还说不出 $\det A$ 是它们的什么函数。本讲回答这个问题，给出两个公式：

- 大公式：$\det A$ 是 $n!$ 个乘积的带符号和，每个乘积从每行每列各取一个元素；

- **代数余子式公式**：把 n 阶行列式写成若干个 $n-1$ 阶行列式的组合。

Strang 在摘要里说这些公式「相当乱」（rather messy）。它们确实不适合手算大矩阵，但它们揭示了结构：哪些项会出现、改动一个元素时行列式怎样变化、为什么 $\det A^{\mathsf{T}} = \det A$。这些是消元法回答不了的。

18.1.1 把每一行拆开

先看 2×2。由性质 3，第一行 $(a,b) = (a,0) + (0,b)$ 可以拆开，第二行不动：

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix}.$$

再对右边两个行列式的第二行做同样的拆分，$(c,d) = (c,0) + (0,d)$，一共得到四块：

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix}.$$

四块各自都很好算，每一步只用三条性质及其推论：

- $\begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} = 0$：两行都只在第 1 列有元素。由线性，从第一行提出 a、第二行提出 c，剩下 $\begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix}$，两行相同，等于 0。
- $\begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} = ad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = ad$。
- $\begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} = bc \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = bc \cdot (-1) = -bc$：剩下的是置换矩阵，交换一次两行就变成 I。
- $\begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix} = 0$：理由同第一块，两行都只在第 2 列有元素。

合起来就是熟悉的公式

$$\boxed{\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0 + ad + (-bc) + 0 = ad - bc.}$$

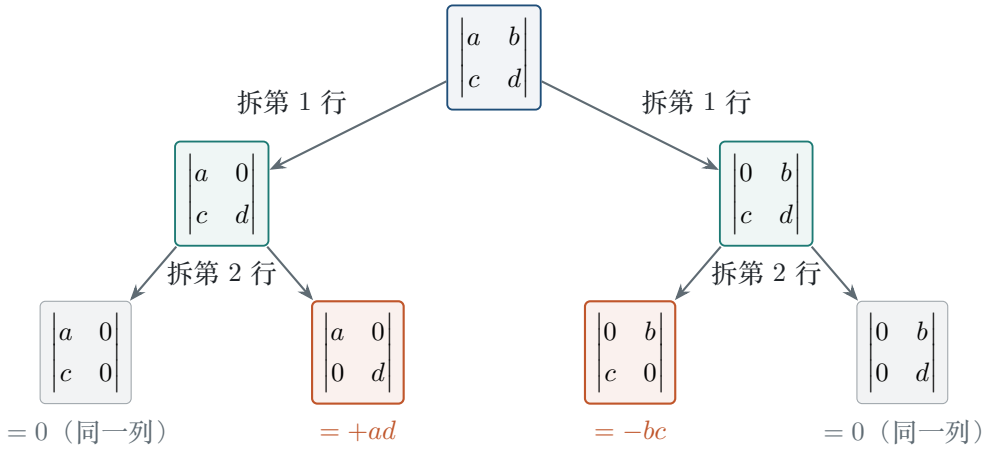


图 18.1　用性质 3 把 2×2 行列式逐行拆开。四块中有两块的两个元素落在同一列，等于 0；活下来的两块（橙色）每行每列各有一个元素。

图 18.1 记录了逐行拆分与去掉零项的过程。这个计算里最值得留意的是哪些块变成了 0：两个元素落在同一列的块。活下来的块恰好每行一个元素、每列一个元素，形状和置换矩阵一样。这就是一般公式的全部秘密。

线性是「每一行分别」线性　性质 3 允许一次只拆一行，其余各行保持不动。它不是说 $\det(A+B)=\det A+\det B$，后者一般不成立：$\det I_2+\det I_2=2$，而 $\det(I_2+I_2)=\det(2I_2)=4$。同理 $\det(2A)\neq 2\det A$：$2A$ 的每一行都乘了 2，每行提出一个 2，所以 $\det(2A)=2^n\det A$。

18.2　3×3：27 块里只剩 6 块

把每行拆成三个坐标方向，会有 $3^3=27$ 种选择。每项都从每行拿一个元素；若拿到的列号重复，对应坐标矩阵就有重复行，行列式为零。剩下的列号必须是 $1,2,3$ 的排列，只有 $3!=6$ 种。

列号排列	逆序数	对应项
$(1,2,3)$	0	$+a_{11}a_{22}a_{33}$
$(1,3,2)$	1	$-a_{11}a_{23}a_{32}$
$(2,1,3)$	1	$-a_{12}a_{21}a_{33}$
$(2,3,1)$	2	$+a_{12}a_{23}a_{31}$
$(3,1,2)$	2	$+a_{13}a_{21}a_{32}$
$(3,2,1)$	3	$-a_{13}a_{22}a_{31}$

这里的逆序数计算的是：前面的列号比后面的列号大的数对有几个。例如 $(3,1,2)$ 中，$3>1$、$3>2$，共两对；把 3 向右交换两次可得到递增排列，符号为正。$(3,2,1)$ 有三对逆序，符号为负。表里的每个符号都由具体排列决定，与矩阵元素是否正数无关。

■ **例 18.1**　**把六项逐一算出**　求 $A=\begin{bmatrix}2&1&3\\0&-1&4\\5&2&0\end{bmatrix}$ 的行列式，并用第一行分组检查结果。

解　按表中的顺序代入，不先省略为零的项：

$$\det A = 0 - 16 - 0 + 20 + 0 + 15 = 19.$$

例如第二项为 $-2\cdot 4\cdot 2 = -16$，最后一项为 $-3\cdot(-1)\cdot 5 = 15$。符号本身为负，乘上一个负元素后整项为正。按第一行重新分组，则

$$2\big((-1)0 - 4\cdot 2\big) - 1\big(0\cdot 0 - 4\cdot 5\big) + 3\big(0\cdot 2 - (-1)5\big) = -16 + 20 + 15.$$

两种组织方式的结果相同，后一个写法已经显露出下一节的代数余子式。 ♦

摘要中的一个下标笔误　课时摘要三阶公式的第二项印为 $-a_{11}a_{23}a_{33}$，重复选了第三列。正确项是 $-a_{11}a_{23}a_{32}$，每行、每列各取一次。用「列号是否重复」检查公式，比只背六项更可靠。

想一想　三阶公式常用三条向右下的乘积减去三条向左下的乘积来记。四阶能否也画四条斜线就得到公式？

答　不能。四阶要从每行每列各取一个，允许的排列有 $4! = 24$ 个；四条正向和四条反向斜线至多给八项，漏掉大量排列。三阶的图记法利用了六个排列恰好能这样排列的特殊性，不能推广成任意阶算法。

18.3　排列公式与定义的存在性

在 n 阶矩阵中，第一行有 n 个列号可选；第二行要避开已经用过的列，剩 $n-1$ 个；继续下去，总共有 $n!$ 个不重复选择。逐行展开虽产生 n^n 项，可能非零的只有这些排列项。即使选中的元素碰巧是零，该排列项仍合法，只是数值为零。

定义 18.1（排列的符号）　排列 σ 的**逆序数**（inversion number）是满足 $i<j$ 且 $\sigma(i)>\sigma(j)$ 的数对个数 $I(\sigma)$。定义 $\operatorname{sgn}(\sigma) = (-1)^{I(\sigma)}$；逆序数为偶数称偶排列，为奇数称奇排列。

第 17 讲已证明交换次数的奇偶性只由排列决定。交换相邻的逆序对使逆序数恰好减少 1；反复交换直到递增排列，需要恰好 $I(\sigma)$ 次。因此以 $e_{\sigma(i)}^{\mathsf{T}}$ 为第 i 行的置换矩阵，其行列式是 $\operatorname{sgn}(\sigma)$。例如 $(4,3,2,1)$ 有 6 个逆序，虽可通过两次远距离交换排好，其符号仍是正。

定理 18.2（排列公式）　对每个 $n\ge 1$ 和实 $n\times n$ 矩阵 A，

$$\det A = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} a_{i,\sigma(i)}.$$

右边满足规范化、交换两行变号、逐行线性，因此构造了第 17 讲所要求的唯一函数。

思路　先用逐行线性说明公式被三条规则强迫；再独立检查右端满足三条规则。两个方向分别回答唯一性和存在性，不能把「假设有行列式后推出公式」当作存在性证明。

证明　将第 i 行写成 $\sum_j a_{ij}e_j^{\mathsf{T}}$，逐行展开。列号重复时，坐标矩阵有重复行，该项为零；列号不重复时就是某个排列 σ，坐标矩阵值为 $\operatorname{sgn}(\sigma)$，得到公式。反过来把右边记为 $D(A)$。对 I_n，只有恒等排列项非零，值为 1。每项中第 i 行只出现一个因子，所以固定其他行后，每项对第 i 行线性；有限和仍线性。最后交换第 r,s 行。把新求和中选择列号 $\sigma(r),\sigma(s)$ 的项，与原矩阵中交换这两个列选择的项配对。乘积因子相同，列号排列差一次交换，符号相反。配对遍历全部排列，所以 D 变号。三条规则全部成立，结合第 17 讲的唯一性，D 就是行列式。 ■

直观理解　排列公式要求每行、每列各用一次：先为每行指定列号，再排除列号重复的选择。两个不同的合法选择即使都选到正元素，也可能因排列不同而带相反符号，所以正元素矩阵仍可能有零行列式。稀疏四阶例子将展示这种抵消。

核心思想　把某个元素作为共同因子分组时，剩下的排列不能再使用它所在的行与列。余子式展开正是这种分组后的公式；它保留全部排列项，只改变求和顺序。

交换不相邻的两个位置也会改变排列奇偶性。若位置为 $r<s$，可先向右移动第 r 个元素 $s-r$ 次，再把原第 s 个元素向左移 $s-r-1$ 次，总计 $2(s-r)-1$ 次相邻交换，为奇数。因此存在性证明中任意两行的交换都确实使对应排列项变号，不只限于相邻两行。符号与元素的正负要分开管理。例 18.1 的最后一项符号为负，但元素乘积为 -15，因此贡献 $+15$。排列公式没有要求元素非零，也没有要求矩阵可逆；即使许多项互相抵消，公式仍成立。

■ **例 18.2　原摘要的稀疏四阶矩阵**　求 $B=\begin{bmatrix}0&0&1&1\\0&1&1&0\\1&1&0&0\\1&0&0&1\end{bmatrix}$ 的全部非零排列项，再解释它为什么奇异。

解　若第一行选第 3 列，第二行只能选第 2 列，第三行必须选第 1 列，第四行选第 4 列，得到 $(3,2,1,4)$，逆序数为 3，贡献 -1。若第一行选第 4 列，第四行必须选第 1 列，第三行只能选第 2 列，第二行选第 3 列，得到 $(4,3,2,1)$，贡献 $+1$。两种情况已穷尽，故 $\det B=-1+1=0$。检查 $R_1-R_2+R_3-R_4=0$，存在非平凡行关系，与零值一致。 ♦

想一想　上例只有两项非零，能否说四阶排列公式只有两项?
答　不能。公式始终含 24 个排列位置；其中 22 项至少选了一个零元素。稀疏性减少实际计算量，没有改变一般公式的项数。

■ **例 18.3　循环移位，原习题 19.1**　求 $P=\begin{bmatrix}0&0&0&1\\1&0&0&0\\0&1&0&0\\0&0&1&0\end{bmatrix}$ 的行列式，推广到 n 阶同样的循环移位矩阵。

解　每行只有一个非零元素，唯一非零选择为 $(4,1,2,3)$。首项 4 与后面三个数各构成逆序，其余无逆序，故值为 $(-1)^3=-1$。一般排列为 $(n,1,2,\dots,n-1)$，逆序数为 $n-1$，故行列式为 $(-1)^{n-1}$。三阶循环保持符号，四阶循环改变符号，两者并不矛盾。 ♦

18.3.1　转置公式的另一种证明

原排列项选择位置 $(i,\sigma(i))$，转置后选择 $(\sigma(i),i)$。按转置矩阵的行号重排，就对应逆排列 σ^{-1}。排列若用 s 次交换排好，逆排列可把交换逆序执行，其符号相同；乘积因子只是换序。所以两边各项一一对应，求和得 $\det(A^{\mathsf T})=\det A$。这条证明不要求可逆，不涉及除法，奇异矩阵也无需额外分类。

18.4　代数余子式：把排列项按位置分组

例 18.1 按第一行分组后，每个第一行元素后面都跟着一个二阶行列式。固定 a_{ij} 后，排列项不能再用第 i 行或第 j 列，剩余选择恰好组成删掉这一行一列后的排列公式。

定义 18.3（余子式与代数余子式）　删去第 i 行、第 j 列所得矩阵记为 A_{ij}。其行列式 $M_{ij}=\det A_{ij}$ 称为**余子式**（minor）；$C_{ij}=(-1)^{i+j}M_{ij}$ 称为**代数余子式**（cofactor）。当 $n=1$ 时，约定空矩阵行列式为 1，从而 $C_{11}=1$。

余子式是数，删行删列后得到的是矩阵，两者应区分。符号按原矩阵的位置计算；删掉行列后重新编号，再拿新编号算符号，会产生错误。

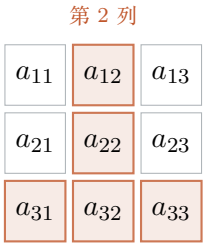
定理 18.4（按任意行或列展开）　固定一行 i 或一列 j，分别有

$$\det A=\sum_{k=1}^{n}a_{ik}C_{ik},\qquad \det A=\sum_{k=1}^{n}a_{kj}C_{kj}.$$

思路　先沿第一行，将排列公式按选中的列分组。移动该列到第一列，剩下的是低一阶公式；一般行再移到第一行，记录额外交换。列展开由转置得到。

证明　先固定第一行的 a_{1j}。将第 j 列向左交换 $j-1$ 次，使它到第一列；把包含此因子的排列项合并，剩余行列正好给 $\det A_{1j}$，符号为 $(-1)^{j-1}=(-1)^{1+j}$。按 j 求和，第一行展开成立。一般第 i 行先向上移动 $i-1$ 次。删掉该行后其他行顺序未变，因此总符号为 $(-1)^{i-1}(-1)^{j-1}=(-1)^{i+j}$。对全部列求和即得行展开。最后对 $A^{\mathsf T}$ 使用行展开，小矩阵转置不改行列式，得到列展开。$n=1$ 时为 $a_{11}\cdot 1=a_{11}$，也成立。 ■

符号的棋盘格记录把被选位置移到左上角所需的交换次数。图 18.2 用第三行第二列说明删行删列的过程。



删第 3 行、第 2 列。

$C_{32} = -(a_{11}a_{23} - a_{13}a_{21})$，符号来自 $(-1)^5$。

图 18.2　计算 C_{32} 时，符号由原位置决定，小矩阵保持剩余行与列的原顺序。

交点元素本身不进入 C_{32}。改变 a_{32} 会改变总行列式，却不会改变它自己的代数余子式；这个区别将在单元素扰动中直接使用。

例 18.4　计算一行的代数余子式　对例 18.1 的矩阵，计算第一行代数余子式，再沿第三列核对行列式。

解　删去对应位置的行列并带上符号，得到 $C_{11} = -8$、$C_{12} = 20$、$C_{13} = 5$，所以第一行给 $2(-8)+1(20)+3(5) = 19$。例如 $C_{12} = -(0\cdot 0 - 4\cdot 5) = 20$，负的余子式乘上负的位置符号，成为正数。第三列另有 $C_{23} = -(2\cdot 2 - 1\cdot 5) = 1$、$C_{33} = 2(-1) - 0 = -2$，因此沿第三列为 $3\cdot 5 + 4\cdot 1 + 0\cdot(-2) = 19$。 ♦

不要重复添加位置符号　沿第一行可以写 $\sum_j a_{1j}C_{1j}$，也可写 $a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13}$。若在 C_{12} 前再加负号，就把同一个位置符号用了两次。上例会从 19 错算成 -21。

18.5　三对角矩阵：展开产生低阶递推

令 T_n 的主对角元为 a，上对角元为 b，下对角元为 c，其他位置为零。若每次直接调用大公式，阶数增加一便使项数大增；但第一行始终只有两个可能非零的元素，代数余子式展开恰好保留这种结构。

定义 18.5（三对角矩阵）　仅允许在主对角线及相邻两条对角线上出现非零元素的方阵称为**三对角矩阵**（tridiagonal matrix）。设 $D_n = \det T_n$，并约定 $D_0 = 1$，则 $D_1 = a$。

例如 $T_3 = \begin{bmatrix} a & b & 0 \\ c & a & b \\ 0 & c & a \end{bmatrix}$。当 $b = c = 0$，它退化为对角矩阵；当 $a = 0$，它仍是三对角矩阵，行列式未必为零。三对角指的是允许非零的位置，不是说三条对角线上的元素必须全非零。

命题 18.6（常系数三对角递推）　对 $n \geq 2$，$D_n = aD_{n-1} - bcD_{n-2}$，其中 $D_0 = 1$、$D_1 = a$。该式对任意实数 a, b, c 成立，无需假设任何一个 D_k 非零。

证明　沿第一行展开，第一项为 aD_{n-1}。第二项是 $-b$ 乘删去第一行第二列的小矩阵行列式。这个小矩阵第一列只有顶部的 c 非零，再沿第一列展开得到 cD_{n-2}，所以第二项为 $-bcD_{n-2}$。其他

第一行元素全零，不贡献。$n=2$ 时直接为 $a^2-bc=aD_1-bcD_0$，说明空矩阵约定正好统一了公式。 ■

这里连续两次展开都利用了零的位置。第二次展开的小矩阵虽然不再是完整的三对角形，沿第一列只有一个元素，因而再次缩小后恢复同类矩阵；不能在第一次删行删列后直接宣布剩下的就是 T_{n-2}，中间还隔着一个因子 c。

■ **例 18.5　原摘要的全 1 三对角矩阵**　令 $a=b=c=1$。求 D_1 到 D_8，并证明序列以 6 为周期。哪些正整数阶数对应奇异矩阵？

解　递推变成 $D_n=D_{n-1}-D_{n-2}$。从 $(D_0,D_1)=(1,1)$ 出发，依次得到

$$D_0,D_1,\dots,D_8=1,1,0,-1,-1,0,1,1,0.$$

要证明周期，不能只凭这几个值猜测。连续计算三步给

$$D_{n+3}=(D_{n+1}-D_n)-D_{n+1}=-D_n,$$

故 $D_{n+6}=-D_{n+3}=D_n$。六个连续初值中只有下标 $2,5$ 的位置为零，所以奇异当且仅当 $n\equiv 2 \pmod 3$。例如 $n=2$ 时两行相同；$n=5$ 虽无重复行，仍因递推消去而奇异。 ♦

递推的状态是相邻两项。$(1,1)$ 再次出现后，后面的值将按相同规则重复；只看到某一项再次等于 1 不足以断言周期，因为它前一项可能不同。图 18.3 把这六个状态画成环。

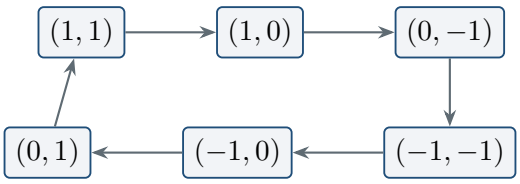


图 18.3　相邻行列式组成状态 (u,v)，下一步为 $(v,v-u)$；六步后回到原状态。

图中的状态同时记录当前值与下一值，所以每个节点只有唯一后继。这解释了为什么相邻两项的重复足以保证整个序列重复，而一个孤立值的重复不够。

■ **例 18.6　离散二阶差分矩阵**　令 K_n 的主对角元为 2，上下对角元为 -1。证明 $\det K_n=n+1$，并求四阶时的值。

解　取 $a=2,b=c=-1$，递推为 $E_n=2E_{n-1}-E_{n-2}$，$E_0=1$、$E_1=2$。若 $E_{n-1}=n$、$E_{n-2}=n-1$，则 $E_n=2n-(n-1)=n+1$。由归纳法，对全部 $n\ge 0$ 成立，故四阶值为 5，并且所有正整数阶矩阵都可逆。还可以把递推改写成 $E_n-E_{n-1}=E_{n-1}-E_{n-2}$，相邻差恒为 1；从 $E_0=1$ 出发，每增加一阶增加 1。这解释了线性增长的来源。 ♦

不要把递推除法用于零值　如果改用消元求 T_n 的主元，常会写出 $d_k = a - bc/d_{k-1}$。这要求前一主元非零。全 1 的例子在第二阶出现零值，后续这种除法无法继续；而 $D_n = aD_{n-1} - bcD_{n-2}$ 没有除法，始终合法。出现零主元时，应该换行或直接使用行列式递推，不能把未定义的商当作零。

■ **例 18.7　非对称三对角矩阵**　求主对角为 3、上对角为 2、下对角为 1 的四阶三对角矩阵行列式。

解　只需记录两个初值与系数乘积：$D_0 = 1, D_1 = 3$，$D_n = 3D_{n-1} - 2D_{n-2}$。于是 $D_2 = 7$、$D_3 = 15$、$D_4 = 31$。交换上下对角元后，乘积 bc 不变，因此值仍为 31，与转置不改变行列式相符。注意这并非说矩阵相同，只是这个标量相同。 ♦

18.6　改变一个元素：代数余子式给出变化率

例 18.1 中若只把右下角的 0 改为 t，是否要重新算六项？沿第三行展开，只有 $a_{33}C_{33}$ 中的元素发生变化，其系数 $C_{33} = -2$ 不含右下角。因此新的行列式为 $19 - 2t$。这个计算也说明：行列式对全体元素不是一次函数，对一个单独元素却总是一次或常数函数。

> **命题 18.7（单元素变化公式）**　设 E_{ij} 只有第 i 行第 j 列为 1，其他位置为 0。对任意实数 t，
>
> $$\det(A + tE_{ij}) = \det A + tC_{ij}(A).$$
>
> 这里 A 不必可逆，$C_{ij}(A)$ 允许为零。

证明　沿第 i 行展开。第 i 行的全部代数余子式都通过删除第 i 行形成，所以该行元素的变化不会改变任何一个展开系数。其余元素保持不变，第 j 项由 $a_{ij}C_{ij}$ 变成 $(a_{ij}+t)C_{ij}$，两式相减便得到结论。即使 $C_{ij} = 0$，上式仍成立，此时改变该元素不改变行列式。 ■

这一结论没有可逆假设。例如 $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ 奇异，仍有 $C_{11} = 4$，把左上角加上 t 后，行列式由 0 变为 $4t$。反过来，一个非零元素也可能有零代数余子式：$\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$ 的 $a_{11} = 1$，但 $C_{11} = 0$，改变它不影响行列式 -6。元素大小和变化系数是不同的信息。

■ **例 18.8　使矩阵奇异的唯一改动**　只改变矩阵 $A = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -1 & 4 \\ 5 & 2 & 0 \end{bmatrix}$ 的右下角，使新矩阵奇异。求改后的值，并判断是否唯一。

解　原右下角为 0，改成 t 等于加上 tE_{33}。因 $\det A = 19$、$C_{33} = -2$，奇异条件为 $19 - 2t = 0$，故 $t = 19/2$。系数非零使这个解唯一；代回行列式得到 $19 - 2(19/2) = 0$，所以条件充分。如果

相应代数余子式为零，改变该位置的值则可能始终奇异或始终可逆，不能保证存在唯一解。 ♦

■ 例 18.9　**原习题 19.2：Pascal 矩阵的一个改动**　设 $P_n=(p_{ij})_{0\le i,j<n}$，$p_{ij}=\binom{i+j}{i}$。已知每个 P_n 的行列式为 1。把右下角元素减去 1，证明所得矩阵奇异。

解　沿末行看右下角系数。删去末行末列留下 P_{n-1}，位置符号为正，故这个代数余子式为 1。取变化量 $t=-1$，新行列式为 $1+(-1)\cdot 1=0$。当 $n=1$，$P_1=[1]$，删行列后为空矩阵，行列式仍约定为 1，减去 1 后为 $[0]$，结论一样成立。四阶时原矩阵为

$$P_4=\begin{bmatrix}1&1&1&1\\1&2&3&4\\1&3&6&10\\1&4&10&20\end{bmatrix}.$$

把 20 改成 19 时，左上三阶行列式为 1，所以总值恰好减少 1。这里用到的是所有阶数的结论，只有 $\det P_4=1$ 这一个信息不足以确定右下角的变化系数。 ♦

一次改变两个元素，能否把两次变化直接相加?

答　一般不能。对 I_2 同时把两个对角元各加 t，行列式为 $(1+t)^2=1+2t+t^2$。若只拿原矩阵两个代数余子式相加，会得到 $1+2t$，漏掉交叉项。单元素公式可以逐次使用，但第二次必须重新计算已改矩阵的代数余子式。若两个改动在同一行，所有展开系数都不含该行，此时才可直接把两项变化相加。

18.6.1　三种算法如何选择

沿零最多的行或列展开，常能把工作量压到很小。例如

$$B=\begin{bmatrix}1&2&0\\3&4&0\\5&6&7\end{bmatrix},\qquad \det B=7\begin{vmatrix}1&2\\3&4\end{vmatrix}=7(4-6)=-14.$$

第三列只有一个非零元，符号为正；不需要计算含 5, 6 的任何小行列式。这也体现了分块三角矩阵的结构：左下角元素在本例中不影响结果。若没有大量零，重复沿行展开就不再节省工作。

方法	优点与适用结构	工作量与限制
排列大公式	直接描述元素组成，适合证明符号及多项式性质	$n!$ 项，每项 n 个因子；直接求值约 $O(n\,n!)$
代数余子式	零多的行列，三对角递推，单元素变化	稠密矩阵朴素递归近于阶乘；需利用零与重复子问题
消元与主元乘积	一般稠密矩阵的数值计算	约 $O(n^3)$；记交换符号，零主元要换行

复杂度比较针对朴素实现，不排除利用特定矩阵结构改进。例如三对角递推从 D_0,D_1 逐个算到 D_n，每步只做固定次数运算，总计 $O(n)$；若把递推的两个子问题每次重新从头算，就会浪费已

知结果。理论公式和实际算法的区别，在于是否保存并利用这些结构。另外，本讲讨论精确等式。实际浮点计算中，接近零的主元还涉及数值稳定性；不能仅凭计算器显示一个很小的行列式，就把数学上的奇异性当作已经证明。符号参数题应当保留精确表达式，再求它的零点。

本讲小结

(1) **逐行多线性**把 n 阶行列式拆成 n^n 个坐标选择；重复列号的项为零，只剩 $n!$ 个排列项。

(2) **排列公式**为 $\det A=\sum_{\sigma\in S_n}(-1)^{\mathrm{inv}(\sigma)}\prod_i a_{i\sigma(i)}$。符号由排列的逆序数决定，与元素自身的正负分别处理。

(3) **转置不变**来自排列与逆排列的一一对应，二者符号相同。因此行性质都有相应的列性质。

(4) **代数余子式**为 $C_{ij}=(-1)^{i+j}\det A_{ij}$；任意行或列都可展开，优先选择零最多的一行或一列。

(5) **三对角递推**为 $D_n=aD_{n-1}-bcD_{n-2}$，$D_0=1, D_1=a$；全 1 情形六步循环，差分矩阵情形 $D_n=n+1$。

(6) **单元素变化**精确满足 $\det(A+tE_{ij})=\det A+tC_{ij}(A)$，没有可逆假设；代数余子式为零时，该元素的任意变化均不影响行列式。

(7) **算法选择**依结构而定：排列公式用来理解结构，余子式用来利用稀疏性，消元适合一般稠密计算。

练习

18.1 ★ 写出排列 $(3,1,4,2)$ 的全部逆序数对，判断其符号；再判断 $(4,3,2,1)$ 的符号。说明为什么不能用排列首项的奇偶性代替逆序数。

18.2 ★ 用代数余子式展开求 $\det\begin{bmatrix}1&0&2\\0&3&0\\4&0&5\end{bmatrix}$。给出所选行列的理由，并用排列公式指出非零的两项。

18.3 ★ 对 $A=\begin{bmatrix}2&1&3\\0&-1&4\\5&2&0\end{bmatrix}$ 求第二行全部代数余子式，沿该行重新算出行列式。

18.4 ★★ 判断并说明理由：①若 $a_{ij}\neq 0$，则 $C_{ij}\neq 0$；②若 A 奇异，则全部 $C_{ij}=0$；③若一行全零，则行列式为零。假的给出具体反例。

18.5 ★★ 令 $T_3(t)=\begin{bmatrix}t&1&0\\1&t&1\\0&1&t\end{bmatrix}$。用递推求行列式，并找出全部使其奇异的实数 t。将每个候选值代回核对。

18.6 ★★ 全 1 三对角矩阵的阶数为 2026 时，行列式是多少？请从递推证明周期后作答，不能只列出几项作猜测。

18.7 ★★ 对上题矩阵，设其行列式为 D_n，证明 $D_n^2 - D_{n-1}D_{n+1} = 1$（$n \ge 1$）。解释为什么不应在证明中除以 D_n。

18.8 ★★ 改变正文矩阵 A 的第二行第一列，将原来的 0 改为 s。求全部使所得矩阵奇异的 s。再说明 C_{21} 是否会随 s 改变。

18.9 ★★★ 已知 Pascal 矩阵 P_n 对全部 $n \ge 1$ 都有行列式 1。若把右下角加 u，求新的行列式；再把题设改为「只知一个固定 P_n 的行列式为 1」，说明原证明缺少哪一步信息。

18.10 ★★★ 设 A 的第 i 行同时增加 $(t_1, \dots, t_n)$，证明新行列式为 $\det A + \sum_j t_j C_{ij}(A)$。说明当变化分布在不同行时，这个相加规则可能失败，并给出二阶反例。

练习答案与提示

18.1 $(3,1,4,2)$ 的逆序数对是位置 $(1,2),(1,4),(3,4)$，对应 $3>1, 3>2, 4>2$，共三对，所以符号为负。$(4,3,2,1)$ 每一对位置都构成逆序，共 $3+2+1=6$ 对，符号为正。两者首项分别为奇数与偶数，但符号由所有数对共同决定；例如 $(3,1,2,4)$ 首项也是 3，却只有两对逆序而为正。

18.2 第二行只有 3 非零，位置符号为正，故行列式为 $3(1\cdot 5 - 2\cdot 4) = -9$。大公式中只有列号 $(1,2,3)$ 与 $(3,2,1)$ 两项不为零，分别为 15 与 -24，相加同样为 -9。

18.3 $C_{21} = -(1\cdot 0 - 3\cdot 2) = 6$，$C_{22} = 2\cdot 0 - 3\cdot 5 = -15$，$C_{23} = -(2\cdot 2 - 1\cdot 5) = 1$。沿第二行给 $0\cdot 6 + (-1)(-15) + 4\cdot 1 = 19$。注意 $a_{21} = 0$ 使它在本次展开中不贡献，并不意味着 $C_{21} = 0$。

18.4 ①假。取 $A = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$，有 $a_{11} = 1$，但 $C_{11} = 0$。②假。取 $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$，其行列式为零，但 $C_{11} = 4$。③真。沿零行展开，每一项都含该行的零元素，所以总和为零。这里不需要对代数余子式本身作任何断言。

18.5 从 $D_0 = 1, D_1 = t$ 得 $D_2 = t^2 - 1$，$D_3 = t(t^2-1) - t = t(t^2-2)$。故奇异的实数恰为 $0, \sqrt{2}, -\sqrt{2}$。代入 0，乘积第一因子为零；代入 $\pm\sqrt{2}$，第二因子均为 $2 - 2 = 0$，三者都确实使行列式为零。非零实数只有平方等于 2 时才满足第二因子为零，因此没有遗漏。

18.6 递推为 $D_{n+2} = D_{n+1} - D_n$，所以 $D_{n+3} = -D_n$，再应用一次得 $D_{n+6} = D_n$。$2026 = 6\cdot 337 + 4$，而 $D_0, D_1, D_2, D_3, D_4 = 1, 1, 0, -1, -1$，故 $D_{2026} = -1$。下标应从 D_0 起算，不能把第一个 1 当作 D_1 后直接按表错移一位。

18.7 记 $Q_n = D_n^2 - D_{n-1}D_{n+1}$。用 $D_{n+1} = D_n - D_{n-1}$ 及 $D_{n+2} = D_{n+1} - D_n$，得到

$$Q_{n+1} - Q_n = D_{n+1}(D_{n+1} - D_n + D_{n-1}) = 0.$$

所以 $Q_n = Q_1 = 1^2 - 1\cdot 0 = 1$。证明只用加乘减；$D_2 = 0$，若除以 D_n 就无法覆盖所有正整数下标。这个恒等式还说明相邻两项不可能同时为零，否则左边会等于零。

18.8 变化量为 s，而 $C_{21} = 6$，所以新行列式是 $19+6s$。唯一零点为 $s = -19/6$，代回得到 $19+6(-19/6) = 0$。C_{21} 在形成小矩阵时删去第二行，因而不含 s，始终等于 6；这也是变化关系保持一次的原因。

18.9 删去末行末列得到 P_{n-1}，其行列式为 1，位置符号为正，故新值是 $1+u$。$n=1$ 时余子式是空行列式 1，结论仍成立。若只知道固定阶数的值，则缺少低一阶行列式等于 1 的信息，不能从 $\det P_n = 1$ 单独推出 $C_{nn} = 1$。例如一般矩阵 $\operatorname{diag}(2, 1/2)$ 的行列式为 1，但右下角代数余子式为 2；这个例子说明为何原证明必须使用矩阵族的低阶信息。

18.10 沿第 i 行展开。所有 C_{ij} 都删掉该行，故变化前后完全相同。新值为 $\sum_j (a_{ij} + t_j)C_{ij} = \det A + \sum_j t_j C_{ij}$。不同行同时变化则可能改变彼此的系数：对 I_2 的两个对角元各加 t，真值为 $(1+t)^2$，按原代数余子式简单相加却为 $1 + 2t$；取 $t = 1$ 时分别为 4 和 3，反例成立。

第 19 讲

Cramer 法则、逆矩阵与体积

Cramer's Rule, Inverse Matrix and Volume

对应材料：MIT 18.06SC Unit II Session 2.7 课时摘要（Strang 第 20 讲视频）；Session 2.7 习题 20.1、20.2 及解答；Strang《Introduction to Linear Algebra》§5.3。

本讲目标

(1) 掌握逆矩阵公式 $A^{-1}=\dfrac{1}{\det A}C^{\mathsf T}$，会用「对角元是行列式展开，非对角元是两行相同的行列式」证明 $AC^{\mathsf T}=(\det A)I$。

(2) 会用公式读出逆矩阵的结构：三角矩阵的逆仍是三角矩阵，整数矩阵在 $\det A=\pm1$ 时逆也是整数矩阵，A^{-1} 怎样随 A 的元素变化。

(3) 掌握 Cramer 法则 $x_j=\det B_j/\det A$ 的两种证明和几何意义，知道它为什么不适合做数值计算。

(4) 理解 $|\det A|$ 是以 A 的列为棱的平行多面体的体积，会用行列式求平行四边形、三角形、平行六面体和四面体的面积与体积。

(5) 会把叉积和混合积写成行列式，并理解线性变换把一切体积放大 $|\det A|$ 倍，由此认识 Jacobi 行列式。

19.1 逆矩阵的代数余子式公式

第 17 讲从三条性质出发定义了行列式，第 18 讲得到了两个计算公式：$n!$ 项的「大公式」和按行展开的代数余子式公式。本讲回答一个自然的问题：**行列式除了判断可逆，还能做什么？**答案有三个：它能写出 A^{-1} 的显式公式，能写出 $Ax=b$ 的解的显式公式，还能算体积。

19.1.1 从 2×2 的公式猜起

2×2 矩阵的逆早在第 3 讲就见过：

$$\begin{bmatrix}a&b\\c&d\end{bmatrix}^{-1}=\frac{1}{ad-bc}\begin{bmatrix}d&-b\\-c&a\end{bmatrix}.$$

分母 $ad-bc$ 是 $\det A$。分子里的四个数是什么？按第 18 讲的定义算 A 的四个代数余子式：划去第 1 行第 1 列剩下 d，所以 $C_{11}=+d$；划去第 1 行第 2 列剩下 c，所以 $C_{12}=-c$；同理 $C_{21}=-b$，$C_{22}=+a$。于是

$$C=\begin{bmatrix}C_{11}&C_{12}\\C_{21}&C_{22}\end{bmatrix}=\begin{bmatrix}d&-c\\-b&a\end{bmatrix},\qquad C^{\mathsf T}=\begin{bmatrix}d&-b\\-c&a\end{bmatrix}.$$

分子恰好是 C^{T}，**代数余子式矩阵的转置**。所以 2×2 的公式可以写成 $A^{-1}=C^{\mathsf{T}}/\det A$。我们猜它对 $n\times n$ 矩阵也成立。

这个猜测在「次数」上是合理的。$\det A$ 的每一项是 n 个元素的乘积，每个代数余子式是 $n-1$ 阶行列式，每一项是 $n-1$ 个元素的乘积。所以 $C^{\mathsf{T}}/\det A$ 的每个元素是「$(n-1)$ 次除以 n 次」，整体是 -1 次。这正是逆矩阵应有的样子：A 放大 2 倍，A^{-1} 应缩小为原来的 $\frac{1}{2}$。

定义 19.1（代数余子式矩阵与伴随矩阵）　设 A 是 $n\times n$ 矩阵。划去第 i 行第 j 列，剩下的 $n-1$ 阶行列式记为 M_{ij}，称为**余子式**（minor）；$C_{ij}=(-1)^{i+j}M_{ij}$ 称为**代数余子式**（cofactor）。以 C_{ij} 为 (i,j) 元的矩阵 C 称为**代数余子式矩阵**（cofactor matrix），它的转置 C^{T} 称为 A 的**伴随矩阵**（adjugate），记作 $\operatorname{adj}A$（国内教材常记作 A^{*}）。

定理 19.2（逆矩阵公式）　对任意 $n\times n$ 矩阵 A，都有

$$AC^{\mathsf{T}}=C^{\mathsf{T}}A=(\det A)\,I.$$

因此当 $\det A\neq 0$ 时，

$$\boxed{A^{-1}=\frac{1}{\det A}\,C^{\mathsf{T}},\qquad (A^{-1})_{ij}=\frac{C_{ji}}{\det A}.}$$

注意下标：A^{-1} 的 (i,j) 元用的是 C_{ji}，不是 C_{ij}。换一种说法：**A 第 1 行的代数余子式放进 A^{-1} 的第 1 列**。

思路　要证 $AC^{\mathsf{T}}=(\det A)I$，就逐个看乘积的元素。$AC^{\mathsf{T}}$ 的 (i,k) 元是「A 的第 i 行」与「C^{T} 的第 k 列」（即 C 的第 k 行）的内积 $\sum_j a_{ij}C_{kj}$。当 $k=i$ 时，这是按第 i 行展开 $\det A$；当 $k\neq i$ 时，它是另一个矩阵的按行展开，那个矩阵有两行相同，行列式为零。

证明　对角元。$(AC^{\mathsf{T}})_{ii}=\sum_{j=1}^{n}a_{ij}C_{ij}$。这正是第 18 讲按第 i 行展开的代数余子式公式，等于 $\det A$。

非对角元。取 $k\neq i$。把 A 的第 k 行换成第 i 行的一个副本，其余各行不动，得到矩阵 A'：

$$A=\begin{bmatrix}\vdots\\ \text{第}i\text{ 行}\\ \vdots\\ \text{第}k\text{ 行}\\ \vdots\end{bmatrix}\quad\longrightarrow\quad A'=\begin{bmatrix}\vdots\\ \text{第}i\text{ 行}\\ \vdots\\ \text{第}i\text{ 行}\\ \vdots\end{bmatrix}.$$

A' 有两行相同，由第 17 讲的性质 4，$\det A'=0$。另一方面，把 $\det A'$ 按第 k 行展开。第 k 行的元素是 $a_{i1},\dots,a_{in}$；而 A' 的 (k,j) 代数余子式是划去第 k 行第 j 列后的行列式，划掉的恰好是 A 与 A' 唯一不同的那一行，所以它与 A 的 C_{kj} 完全相同。于是

$$0=\det A'=\sum_{j=1}^{n}a_{ij}C_{kj}=(AC^{\mathsf{T}})_{ik}.$$

对角元都是 $\det A$，非对角元都是 0，所以 $AC^{\mathsf T} = (\det A)I$。

另一边。$A^{\mathsf T}$ 的余子式是 A 的余子式转置位置后的结果（划去 $A^{\mathsf T}$ 的第 i 行第 j 列，等于把 A 划去第 j 行第 i 列后再转置，行列式不变），所以 $A^{\mathsf T}$ 的代数余子式矩阵是 $C^{\mathsf T}$。对 $A^{\mathsf T}$ 用刚证的结论：$A^{\mathsf T}(C^{\mathsf T})^{\mathsf T} = (\det A^{\mathsf T})I$，即 $A^{\mathsf T}C = (\det A)I$。两边取转置得 $C^{\mathsf T}A = (\det A)I$。

最后，$\det A \neq 0$ 时两边除以 $\det A$，得 $A \cdot \dfrac{C^{\mathsf T}}{\det A} = I$，所以 $A^{-1} = C^{\mathsf T}/\det A$。 ■

证明的关键一步是：用一行的元素去乘另一行的代数余子式，得到的是两行相同的行列式，必为零。在 2×2 的情形，A 的第 1 行 (a,b) 乘 $C^{\mathsf T}$ 的第 2 列 $(-b,a)$，得 $a(-b)+ba=0$，这正是 $\det\begin{bmatrix} a & b \\ a & b \end{bmatrix}$。

■ **例 19.1　改编自习题 20.1**　设 $A = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 2 & 2 \\ 1 & 2 & 5 \end{bmatrix}$。求代数余子式矩阵 C，计算 $AC^{\mathsf T}$，由此读出 $\det A$ 和 A^{-1}。如果把 $a_{13}=4$ 改成 100，$\det A$ 为什么不变？

解　逐个划行划列。符号按棋盘格 $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$ 交替。

$$
\begin{aligned}
C_{11} &= 2\cdot 5 - 2\cdot 2 = 6, & C_{12} &= -(1\cdot 5 - 2\cdot 1) = -3,\\
C_{13} &= 1\cdot 2 - 2\cdot 1 = 0, & C_{21} &= -(1\cdot 5 - 4\cdot 2) = 3,\\
C_{22} &= 1\cdot 5 - 4\cdot 1 = 1, & C_{23} &= -(1\cdot 2 - 1\cdot 1) = -1,\\
C_{31} &= 1\cdot 2 - 4\cdot 2 = -6, & C_{32} &= -(1\cdot 2 - 4\cdot 1) = 2,\\
C_{33} &= 1\cdot 2 - 1\cdot 1 = 1.
\end{aligned}
$$

所以

$$
C = \begin{bmatrix} 6 & -3 & 0 \\ 3 & 1 & -1 \\ -6 & 2 & 1 \end{bmatrix}, \qquad AC^{\mathsf T} = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 2 & 2 \\ 1 & 2 & 5 \end{bmatrix}\begin{bmatrix} 6 & 3 & -6 \\ -3 & 1 & 2 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = 3I.
$$

以第一行为例核对：$(1,1)$ 元 $1\cdot 6+1\cdot(-3)+4\cdot 0=3$；$(1,2)$ 元 $1\cdot 3+1\cdot 1+4\cdot(-1)=0$；$(1,3)$ 元 $1\cdot(-6)+1\cdot 2+4\cdot 1=0$。由定理 19.2，$\det A = 3$，且

$$
A^{-1} = \frac{1}{3}\begin{bmatrix} 6 & 3 & -6 \\ -3 & 1 & 2 \\ 0 & -1 & 1 \end{bmatrix}.
$$

至于改动 a_{13}：按第 1 行展开，$\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = 6 - 3 + a_{13}\cdot 0$。$C_{13}$ 是划去第 1 行、第 3 列后的行列式，本身不含 a_{13}；它等于 0，因为剩下的 $\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ 两行相同。所以 a_{13} 取什么值，$\det A$ 都是 3。 ♦

别忘了转置　若直接用 C 而不转置，$AC=\begin{bmatrix}-15 & 6 & 3\\ 0 & 3 & 0\\ -18 & 9 & 3\end{bmatrix}$，根本不是 $3I$。记法：第 i 行的代数余子式进第 i 列。对称矩阵的 C 是对称的，转置不转置看不出差别，所以检验公式时要选非对称的例子。

19.2　从公式读出结构

若一个整数矩阵的行列式为 1，求逆后会不会出现分数？伴随矩阵公式让这个问题不用消元就有答案。每个代数余子式仍是整数，分母只有 1，所以逆矩阵的每个元素也都是整数。下面把这个观察的正反两面说清楚。

命题 19.3（整数矩阵的整数逆）　设 A 为整数方阵。A 有整数元素的逆矩阵，当且仅当 $\det A=1$ 或 -1。

证明　若 $\det A=\pm1$，则 $A^{-1}=\pm C^{\mathsf T}$ 为整数矩阵。反过来，若 A 和 A^{-1} 都是整数矩阵，二者的行列式都是整数，且由乘法公式 $\det A\det A^{-1}=1$。两个整数的乘积为 1，只能同为 1 或同为 -1。 ■

■ **例 19.2　逆矩阵没有分数**　对 $U=\begin{bmatrix}2 & 1\\ 3 & 2\end{bmatrix}$ 求逆，并比较 $V=2I_2$。

解　先算分母：$\det U=4-3=1$，故

$$U^{-1}=\begin{bmatrix}2 & -1\\ -3 & 2\end{bmatrix}.$$

验算 UU^{-1} 的两行分别为 $(4-3,-2+2)=(1,0)$ 和 $(6-6,-3+4)=(0,1)$。而 $\det V=4$，$V^{-1}=\frac12 I_2$ 有分数。整数矩阵可逆只要求行列式非零，要求逆仍为整数则进一步要求行列式是 ±1。 ♦

每个元素都变小了吗　把 A 换成 $2A$，为什么逆矩阵变成 $A^{-1}/2$？能否从余子式公式中看出来？
答　对 $n\ge2$，每个 $(n-1)$ 阶代数余子式放大 2^{n-1} 倍，行列式放大 2^n 倍，相除得到 $1/2$。$n=1$ 时空行列式为 1，同样成立。这说的是整体按同一系数缩放；单独放大一个元素时，逆矩阵的各元素通常不会按同一比例变化。

■ **例 19.3　上三角结构的保留**　求 $T=\begin{bmatrix}2 & 1 & 0\\ 0 & 3 & 1\\ 0 & 0 & 4\end{bmatrix}$ 的逆，并解释为什么左下方仍全为零。

解　思路是逐列解 $Tx=e_j$。先从第三个方程求 x_3，再向上回代，得到

$$T^{-1}=\begin{bmatrix}1/2 & -1/6 & 1/24\\ 0 & 1/3 & -1/12\\ 0 & 0 & 1/4\end{bmatrix}.$$

例如第三列：$4x_3=1$ 给 $x_3=1/4$，$3x_2+x_3=0$ 给 $x_2=-1/12$，$2x_1+x_2=0$ 给 $x_1=1/24$。求第 j 列时，下方齐次方程依次强迫 $x_n,\dots,x_{j+1}=0$，所以逆矩阵是上三角矩阵。前提是对角元都非零，否则回代未必能做。 ♦

单元素改变时还有一个精确的观察。把 a_{ij} 加 t，第 18 讲已经证明新行列式是 $\det A+tC_{ij}$，因为 C_{ij} 删去第 i 行而不含 a_{ij}。若 $C_{ij}\neq 0$，在 $t=-\det A/C_{ij}$ 处矩阵必奇异，逆矩阵公式的分母归零。若 $C_{ij}=0$，行列式不受这次改动影响，但其他余子式可能改变，因此不能推出逆矩阵也不变。

19.3　Cramer 法则：每个坐标是两个行列式之比

方程 $2x+y=5$、$x-y=1$ 的解是 $(2,1)$。消元很快，但若只想要一个坐标的表达式，有没有直接公式？把系数矩阵的第一列换成右端，行列式从 -3 变成 -6；换第二列则变成 -3。两个比值恰为 2 和 1。

定理 19.4（Cramer 法则）　设 $A=[a_1\ \cdots\ a_n]$ 为可逆方阵，$b\in\mathbb{R}^n$。令 B_j 是把 A 第 j 列替换为 b 所得矩阵，则 $Ax=b$ 的唯一解满足

$$\boxed{x_j=\frac{\det B_j}{\det A},\qquad j=1,\dots,n.}$$

直观理解　$b=x_1a_1+\cdots+x_na_n$。把 b 放到第 j 列时，其余 a_k 已经各占一列。只有 b 在 a_j 方向的那一份能贡献行列式；其他份都造成重复列。因此新行列式保留下来的恰是 $x_j\det A$。

证明　沿 B_j 第 j 列展开，余子式删掉了被换的列，故仍是 A 的 C_{ij}。因此

$$\det B_j=\sum_{i=1}^{n}b_iC_{ij}=(C^{\mathsf T}b)_j.$$

由逆矩阵公式，$x=A^{-1}b=C^{\mathsf T}b/\det A$，得到所求公式。

另一个证明直接用多重线性。把 $b=\sum_k x_ka_k$ 代入 B_j 的第 j 列，展开为 n 个行列式。$k\neq j$ 时有两列 a_k，所以那项为零；$k=j$ 时得到 $x_j\det A$。故 $\det B_j=x_j\det A$。可逆性保证分母非零，也保证任意 b 都能写成上述列组合。 ■

■ **例 19.4　二阶方程的逐项验证**　用 Cramer 法则解 $2x+y=5$、$x-y=1$。

解　系数矩阵 $B=\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix}$ 的行列式是 $-2-1=-3$，所以唯一解存在。两次换列给

$$\det B_1=\begin{vmatrix}5 & 1\\ 1 & -1\end{vmatrix}=-6,\qquad \det B_2=\begin{vmatrix}2 & 5\\ 1 & 1\end{vmatrix}=-3.$$

于是 $x=(-6)/(-3)=2$，$y=(-3)/(-3)=1$。回代：$2\cdot 2+1=5$，$2-1=1$。两式都满足，而可逆性排除了其他解。换列的位置由所求坐标决定，右端的分量顺序保持不动。 ♦

■ **例 19.5　只替换一列**　对例 19.1 的矩阵，解 $Ax=(-1,3,0)^{\mathsf T}$。

解　已知 $\det A=3$，先分别算三个分子：

$$\det\begin{bmatrix}-1 & 1 & 4\\ 3 & 2 & 2\\ 0 & 2 & 5\end{bmatrix}=3,\quad \det\begin{bmatrix}1 & -1 & 4\\ 1 & 3 & 2\\ 1 & 0 & 5\end{bmatrix}=6,\quad \det\begin{bmatrix}1 & 1 & -1\\ 1 & 2 & 3\\ 1 & 2 & 0\end{bmatrix}=-3.$$

因而 $x=(1,2,-1)^{\mathsf T}$。逐行验算给 $1+2-4=-1$、$1+4-2=3$、$1+4-5=0$。若只要求 x_2，只计算第二个分子就够了。若要大量不同右端的解，则一次分解 A 后反复回代通常更经济。 ♦

零分母不能用比值判解　$\det A=0$ 时，Cramer 比值没有定义。比如 $A=\begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}$，右端 $(1,2)^{\mathsf T}$ 给出无穷多解 $x+y=1$，右端 $(1,3)^{\mathsf T}$ 则无解。两者分母都为零，只有回到相容性与零空间才能区分。尤其 $0/0$ 不能被解释为「坐标等于零」。

为什么少有软件直接用 Cramer 法则　一个 n 元方程组用这条法则，要计算多少个行列式？
答　连分母在内是 $n+1$ 个。若每个都独立用消元算，一般要约 n^3 量级运算，总计约 n^4；而直接消元解方程只需约 n^3。若用排列公式，每个又含 $n!$ 项，代价更高。这些是一般稠密情形的量级比较，特殊稀疏结构可以更快。显式公式适合理论和小型手算，算法应利用可复用的分解。

19.4　面积：先分清符号与大小

边向量 $u=(3,1)^{\mathsf T}$、$v=(1,2)^{\mathsf T}$ 围成的平行四边形面积是多少？底边不是坐标轴，直接找高有点繁琐。行列式给出 $3\cdot 2-1\cdot 1=5$。但交换两条边后得到 -5，几何面积仍是 5。图 19.1 将同一个区域和两种列顺序区分开来。

也可以先做独立的底高核对：u 的长度为 $\sqrt{10}$，与它垂直的单位向量可取 $(-1,3)/\sqrt{10}$。v 在这个方向上的投影长为 $|-1+6|/\sqrt{10}=5/\sqrt{10}$，所以底乘高正好是 5。若 v 平行于 u，这个投影为零，区域就退化为一条线，面积与行列式都为零。

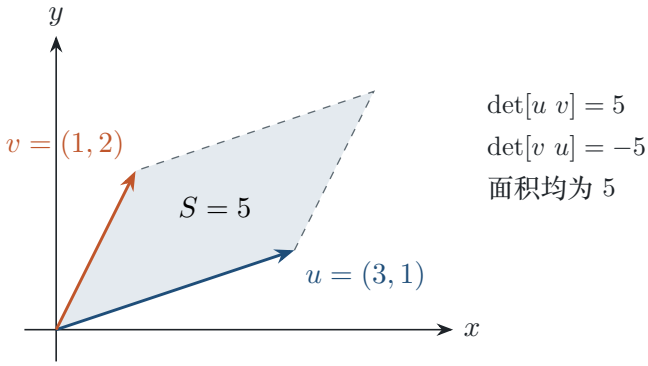


图 19.1　交换列改变定向，平行四边形本身及其面积都不改变。

从 u 逆时针转向 v 的较小夹角在 0 到 π 之间，所以此列顺序的行列式为正。图中的颜色只识别两条边，区域的大小由绝对值给出。

定理 19.5（二维面积公式）　设 $u=(a,b)^{\mathsf T}$、$v=(c,d)^{\mathsf T}$。以二者为邻边的平行四边形面积为 $|ad-bc|$，三角形面积为 $|ad-bc|/2$。当二者线性相关时，两种面积均为零。

证明　若 $u=0$，结论两边都是零。若 $u\neq 0$，取沿 u 的单位向量 $e=u/\|u\|$，再取与其垂直的单位向量 $f=(-b,a)^{\mathsf T}/\|u\|$。底长为 $\|u\|$，高为 $|f^{\mathsf T}v|=|ad-bc|/\|u\|$，乘积正好是 $|ad-bc|$。从平行四边形对角线分割得到两个全等三角形，所以面积减半。 ■

■ **例 19.6　从顶点坐标到面积**　求顶点 $P=(1,1)$、$Q=(4,2)$、$R=(2,5)$ 所成三角形的面积。若把三个点的坐标同时平移，答案是否改变？

解　先选一个共同起点。两条边为 $Q-P=(3,1)$、$R-P=(1,4)$，故面积是 $|3\cdot 4-1\cdot 1|/2=11/2$。也可直接写成

$$S=\frac{1}{2}\left|\begin{vmatrix}1&1&1\\4&2&1\\2&5&1\end{vmatrix}\right|=\frac{11}{2}.$$

后一公式的理由是从第二、三行分别减去第一行，沿最后一列展开，即得到边向量的二阶行列式。共同平移不改变坐标之差，所以面积不变。若直接把两个顶点当作边，则一般会算错；这里必须先减去同一个起点。 ♦

■ **例 19.7　剪切不改面积**　固定 $u=(3,1)^{\mathsf T}$，把另一条边从 $v=(1,2)^{\mathsf T}$ 换成 $v+tu$，其中 t 为任意实数。证明面积始终为 5。

解　用多重线性与重复列为零，

$$\det[u\ v+tu]=\det[u\ v]+t\det[u\ u]=5.$$

几何上，新边的端点沿平行于底边的直线移动，高保持不变，底也保持不变。t 可以是负数，也可以很大，结论都成立。这个变换叫剪切，给出了行列式「给一列加另一列的倍数不变」的面积解

释。 ♦

由坐标构造边　已知图形顶点时，先选同一个顶点为起点，将其余顶点减去它，再把边放进矩阵。二维平行四边形取行列式绝对值，三角形再除以 2；三维对应的除数是 1 与 6。顶点坐标、边向量与最后的单位必须分别核对。

19.5　体积与线性变换

长方体三条互相垂直的棱长为 $2,3,4$，体积是 24。若给第三条棱添一段沿第二条棱的分量，盒子倾斜了，底与高仍然不变。一般的平行六面体体积也有同样的计算规则。

定理 19.6（平行多面体体积）　设 A 为实 $n\times n$ 矩阵。以它的列为棱的平行多面体 $\{At : 0\le t_i\le 1\}$ 的 n 维体积为 $|\det A|$。线性变换 $x\mapsto Ax$ 将任意可测集合的体积乘以 $|\det A|$。

思路　先看初等行变换对应的几何动作：交换坐标不改体积、缩放一个坐标使体积乘缩放系数绝对值、一个坐标加另一个的倍数是剪切。消元把一般可逆矩阵分解成这些动作，从而把几何因子与行列式因子逐步对应。

证明　对交换坐标与缩放坐标，体积因子分别为 1 与 $|c|$。剪切例如 $(x_1,x_2,\dots)\mapsto(x_1+tx_2,x_2,\dots)$，在固定 $x_2,\dots$ 的每个截面上只作平移；截面的长度保持不变，逐层累积体积也保持不变。这里使用通常欧氏体积的截面计算性质。

可逆矩阵由有限个可逆初等矩阵相乘而成。连续实施上述动作，体积因子相乘。每步因子等于相应初等矩阵行列式的绝对值，由行列式乘法公式，总因子是 $|\det A|$。单位立方体体积为 1，其像的体积因而为 $|\det A|$。同一动作分解对任意可测集合成立，得到第二个结论。

若 A 奇异，所有像点位于一个维数低于 n 的线性子空间，n 维体积为零，而 $\det A=0$，结论仍成立。低维图形在自身所在平面中的面积可以非零，此处说的是环境空间的 n 维体积。 ■

三维中的几何形状可见图 19.2。图只把三维坐标投影到纸面，不能用纸面上的夹角或长度直接算三维体积；计算依据仍是图注中给出的真实向量。

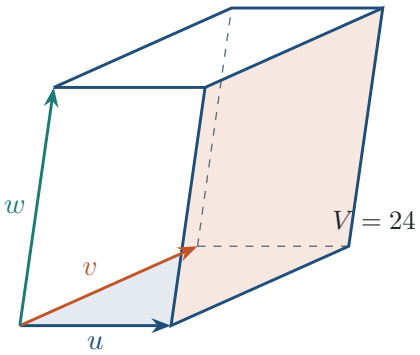


图 19.2　棱向量 $u=(2,0,0)$、$v=(1,3,0)$、$w=(0,1,4)$ 围成体积为 24 的平行六面体。

底面由 u,v 张成，位于 $z=0$ 平面，面积为 6；第三条棱的垂直高度是 4，所以体积 $6\cdot 4=24$。

w 的第二分量 1 只造成倾斜，不改变高度。图中的虚线表示被前面遮挡的棱，三条有色箭头都从同一顶点出发。

■ **例 19.8　平行六面体与四面体**　求图 19.2 的平行六面体体积，以及由 O, u, v, w 四个顶点围成的四面体体积。

解　按列放棱得到上三角矩阵

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{bmatrix}, \qquad |\det A| = 2 \cdot 3 \cdot 4 = 24.$$

标准四面体 $s_i \geq 0$、$s_1 + s_2 + s_3 \leq 1$ 的水平截面在高度 t 处是直角三角形，面积为 $(1-t)^2/2$，积分 $\int_0^1 (1-t)^2/2 \, \mathrm{d}t = 1/6$。线性变换 A 将它送到所求四面体，因此其体积为 $24/6 = 4$。三维除以 6 来自标准四面体的体积，不能沿用三角形的除以 2。 ♦

注 19.1　[叉积与混合积] 三维向量 u, v 的叉积 $u \times v$ 垂直于二者，长度为底面积；其坐标按二阶余子式排列。因此 $(u \times v)^{\mathsf{T}} w = \det[u \ v \ w]$，绝对值就是底面积乘垂直高度。对图中的边，$u \times v = (0, 0, 6)$，点乘 w 得 24。叉积只适用于此处的三维背景，而行列式体积公式适用于任意维数。

绝对值不能分配到各项　一般不能把 $|\det[u+v, w]|$ 写成 $|\det[u, w]| + |\det[v, w]|$。取 $u = (1, 0)$、$v = (-1, 0)$、$w = (0, 1)$，左边为 0，右边为 2。线性的是有向行列式；取绝对值后，不同方向的贡献可能相消。几何面积非负，与行列式的多重线性要分开使用。

19.6　Jacobi 行列式：局部的小盒子

注 19.2　[阅读准备] 本节用到多元函数的可微性与偏导数。尚未学过这些内容时，可暂缓本节及对应练习；前面的行列式、Cramer 法则和线性变换体积公式已构成完整的线性代数结论，后面的特征值课程不依赖本节。

极坐标中同样的 $\mathrm{d}r \, \mathrm{d}\theta$ 为什么在远处覆盖更大的面积？固定角宽时，圆弧长度随半径增大。非线性坐标变换在每一点附近近似为线性变换，这就把上一节的体积因子带进微积分。

定义 19.7（导数矩阵与 Jacobi 行列式）　设 $F : \mathbb{R}^n \to \mathbb{R}^n$ 在 q 可微。导数矩阵为 $DF(q) = (\partial F_i / \partial q_j)$，它的行列式称为该点的 Jacobi 行列式。小增量满足 $F(q+h) = F(q) + DF(q)h + o(\|h\|)$，局部的无符号体积因子为 $|\det DF(q)|$。

注意矩阵的第 j 列是只改变第 j 个参数时的位置变化率。参数单位也要带入：角变量用弧度，不能把角度数值直接代进以弧度求出的导数。严格的积分换元公式还要求相应的单射性与正则性；本讲只计算局部体积因子，不用行列式计算代替整个换元定理。

例 19.9　球坐标，原习题 20.2　球坐标约定为 $x=\rho\sin\phi\cos\theta$、$y=\rho\sin\phi\sin\theta$、$z=\rho\cos\phi$，其中 $\rho>0$、$0<\phi<\pi$。求按 (ρ,ϕ,θ) 排列的 Jacobi 行列式。

解　思路是先固定列的参数顺序，逐项求导，再提出列中的公因子。导数矩阵是

$$DF=\begin{bmatrix}\sin\phi\cos\theta & \rho\cos\phi\cos\theta & -\rho\sin\phi\sin\theta\\ \sin\phi\sin\theta & \rho\cos\phi\sin\theta & \rho\sin\phi\cos\theta\\ \cos\phi & -\rho\sin\phi & 0\end{bmatrix}.$$

沿第三行展开，先合并每个二阶行列式中的 $\cos^2\theta+\sin^2\theta$，得到

$$\begin{aligned}\det DF&=\cos\phi\,[\rho^2\cos\phi\sin\phi]+\rho\sin\phi\,[\rho\sin^2\phi]\\&=\rho^2\sin\phi(\cos^2\phi+\sin^2\phi)\\&=\rho^2\sin\phi.\end{aligned}$$

因此局部体积元为 $\rho^2\sin\phi\,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}\theta$。三个方向的局部长度分别为 $\mathrm{d}\rho$、$\rho\,\mathrm{d}\phi$ 与 $\rho\sin\phi\,\mathrm{d}\theta$，相乘也得到相同结果。在极点 $\phi=0,\pi$ 或原点 $\rho=0$，因子为零，表示此坐标描述退化，并不表示邻域本身没有三维体积。 ♦

球坐标的三个局部尺度可由图 19.3 读出。图中只画固定 θ 的子午截面；绕 z 轴改变 θ 时，运动半径是 $r=\rho\sin\phi$，这一长度需要从截面再作一次投影。

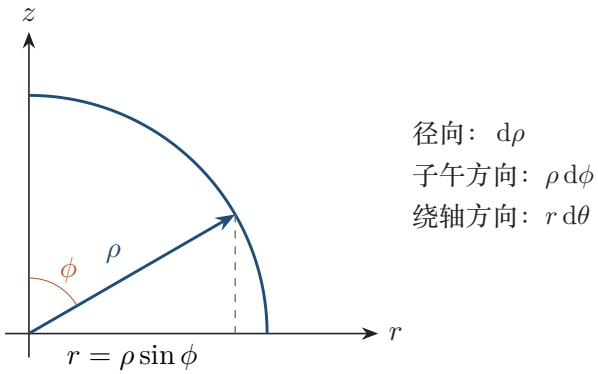


图 19.3　球坐标的角 ϕ 从正 z 轴量起，绕轴方向的尺度包含 $\sin\phi$。

图示点取 $\rho=3$、$\phi=\pi/3$，所以横坐标是 $3\sqrt{3}/2$，竖坐标是 $3/2$。子午弧在纸面可见，绕轴弧垂直于该截面；两者分别贡献一个 ρ 与一个 $\rho\sin\phi$。因而球坐标体积公式里的平方与正弦有各自明确的几何来源。

例 19.10　极坐标与顺序　对 $x=r\cos\theta$、$y=r\sin\theta$ 求按 (r,θ) 顺序的导数矩阵与面积因子。若改成 (θ,r)，会发生什么？

解　两列分别是径向与角向的变化率，故

$$DF=\begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{bmatrix},\qquad \det DF=r.$$

交换参数顺序会交换两列，使行列式变成 $-r$。在 $r>0$ 的区域，无符号面积因子两种顺序都为 r。若参数域允许负半径，则必须取 $|r|$，而且同一个点可能有多个参数表示，积分时还须处理重复覆盖。计算行列式只解决局部尺度问题，不能自动解决全局参数的唯一性。 ♦

本讲小结

(1) **伴随矩阵**为 $\operatorname{adj} A = C^{\mathsf{T}}$，满足 $A \operatorname{adj} A = \operatorname{adj} A\, A = (\det A) I$；只有 $\det A \neq 0$ 时才能除得逆矩阵。

(2) **整数逆**的条件为 $\det A = \pm 1$；整数元素与非零行列式本身不足以保证逆仍为整数。

(3) **Cramer 法则**给出 $x_j = \det B_j / \det A$，B_j 只替换第 j 列；批量求解宜复用消元分解。

(4) **面积与体积**取行列式绝对值；三角形再除以 2，三维四面体再除以 6，边向量先由顶点相减得到。

(5) **线性体积因子**为 $|\det A|$；剪切保持体积，奇异变换将像压到低维空间。

(6) **局部坐标因子**为 $|\det DF|$；球坐标按 (ρ,ϕ,θ) 排列时为 $\rho^2 \sin\phi$，极坐标为 r（$r>0$）。

练习

19.1 ★ 对 $A = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$，求代数余子式矩阵、伴随矩阵和逆矩阵，并分别验证左右乘积。

19.2 ★ 用 Cramer 法则解 $3x+2y=7$、$x+y=3$；必须写出两个换列矩阵并回代。

19.3 ★ 求顶点 $(0,0)$、$(2,1)$、$(1,3)$ 所成三角形的面积，再交换后两个顶点，比较行列式与面积。

19.4 ★ 一个平行六面体以 $(1,0,0)$、$(2,2,0)$、$(3,4,-3)$ 为棱。求体积；若仅将第三棱反向，体积和有向行列式分别怎样变化？

19.5 ★★ 判断真假并证明或给出反例：(a) 任何整数可逆矩阵的逆也为整数矩阵；(b) $\det A = 0$ 时 $\operatorname{adj} A = 0$；(c) 若方阵的某一列为零，则其行列式为零。

19.6 ★★ 设 $A = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$。求其逆，并解释为什么行列式虽不依赖 t，逆却依赖 t。从剪切图像解释其面积因子。

19.7 ★★ 对三个空间点 P,Q,R 和第四点 S，写出四面体体积公式。应用到 $P=(1,1,1)$、$Q=(3,1,1)$、$R=(2,4,1)$、$S=(1,2,5)$。

19.8 ★★ 设 A 可逆、$Ax=b$。不用求逆，证明把 A 第 j 列换成 b 后的行列式等于 $x_j \det A$。指出推导中哪些等式即使 A 奇异仍成立。

19.9 ★★★ 设 $n \geq 2$，证明 $\det(\operatorname{adj} A) = (\det A)^{n-1}$，先证可逆情形，再处理奇异情形。不要在奇异情形除以 $\det A$。

19.10 ★★★ 计算柱坐标 $x = r\cos\theta$、$y = r\sin\theta$、$z = h$ 按 (r,θ,h) 顺序的 Jacobi 行列式；交换前两个参数后重新计算。说明在 $r=0$ 退化的是哪一列，以及为什么这不代表空间的 z 轴

附近没有体积。

练习答案与提示

19.1 $\det A = 3-2 = 1$，代数余子式矩阵为 $C = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$，伴随矩阵与逆矩阵均为 $\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}$。右乘得到两行 $(3-2, -6+6) = (1,0)$、$(1-1, -2+3) = (0,1)$；左乘得到 $(3-2, 2-2) = (1,0)$、$(-3+3, -2+3) = (0,1)$。两个乘法顺序都给单位阵，非对称矩阵能检验转置位置是否写对。

19.2 分母是 $\begin{vmatrix} 3 & 2 \\ 1 & 1 \end{vmatrix} = 1$。换第一列得 $\begin{bmatrix} 7 & 2 \\ 3 & 1 \end{bmatrix}$，行列式为 $7-6=1$；换第二列得 $\begin{bmatrix} 3 & 7 \\ 1 & 3 \end{bmatrix}$，行列式为 $9-7=2$。故 $(x,y) = (1,2)$。代回 $3+4=7$、$1+2=3$。分母非零保证唯一性，所以已得到全部解。若把右端误放进一行，计算的将是另一个矩阵，其结果一般不会给出所求坐标。

19.3 两条边构成 $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$，行列式为 $6-1=5$，故三角形面积为 $5/2$。交换后两点，矩阵两列互换，有向行列式变为 -5，绝对值仍为 5，面积仍为 $5/2$。原点恰好是共同顶点，所以本题可以直接用其余两点；一般题目需先减去共同顶点。将所有点同时平移时差向量不变，因此面积也不变。

19.4 按列排成上三角矩阵 $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & -3 \end{bmatrix}$，行列式为 -6，体积为 6。第三棱乘 -1 会使行列式乘 -1 而变成 6，体积保持 6。棱的负向改变了定向，没有改变底面与高的绝对值。若三条棱都反向，行列式乘 $(-1)^3 = -1$；这也说明不能用体积的非负性推出行列式非负。

19.5 (a) 假，$2I_2$ 的逆为 $I_2/2$。正确条件为 $\det A = \pm 1$。(b) 假，$A = \operatorname{diag}(1,0)$ 的伴随矩阵为 $\operatorname{diag}(0,1)$。恒等式 $A \operatorname{adj} A = 0$ 只说明伴随矩阵各列属于零空间，不强迫每列为零。(c) 真，沿零列展开，每一项都含一个零元素，故和为零。注意这个证明并未断言该列代数余子式全为零；例如上述对角矩阵的第二列为零，但 $C_{22} = 1$。

19.6 $\det A = 1$，故 $A^{-1} = \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix}$，直接相乘的右上角是 $-t+t=0$。行列式只测面积因子，不能完整描述线性变换；不同剪切可有相同面积因子。几何上 $A(x,y) = (x+ty, y)$，每条水平线上的点都平移同一距离 ty，因此每条水平截面的长度与总面积保持不变。逆映射把该平移量减去，即把 t 换成 $-t$。

19.7 令 $u = Q-P$、$v = R-P$、$w = S-P$，体积为 $|\det[u\ v\ w]|/6$。本题三条边为 $(2,0,0)$、$(1,3,0)$、$(0,1,4)$，行列式为 24，体积为 4。这是图中平行六面体对应的四面体作共同平移，所得数值相同。公式在四点共面时也成立：三条差向量线性相关，行列式与三维体积均为零。共面三角形本身的二维面积可以非零，因此须明确题目求的是几维体积。

19.8 由 $b = \sum_k x_k a_k$，按第 j 列的线性展开。$k \neq j$ 时重复第 k 列，行列式为零；剩下 $k = j$ 项为 $x_j \det A$。只要已有等式 $Ax = b$，这段推导在奇异情形也成立。但奇异时不能对任意 b 保证有解，不能除以 $\det A$，也不能得到唯一性。因此恒等式的有效范围大于比值公式的有效范围。

19.9 可逆时 $\operatorname{adj} A = (\det A) A^{-1}$，所以其行列式为 $(\det A)^n / \det A = (\det A)^{n-1}$。奇异时从 $A \operatorname{adj} A = 0$ 看，伴随矩阵若可逆，右乘其逆便会推出 $A = 0$。但 $n \geq 2$ 且 $A = 0$ 时所有 $(n-1)$ 阶余子式均为零，伴随矩阵为零矩阵，仍不可逆，矛盾。因此奇异时 $\det(\operatorname{adj} A) = 0$，与右边一致。这里明确使用 $n \geq 2$；一阶零矩阵的伴随矩阵按空行列式约定是 $[1]$，不适用本段的零矩阵判断。

19.10 导数矩阵为 $\begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$，沿第三行展开得 r。交换前两个参数即交换前两列，行列式为 $-r$，无符号体积因子仍为 $|r|$。在 $r = 0$ 时第二列全零，因为改变角度不改变轴上点的位置；不同

角度代表同一个点。轴附近 $r > 0$ 的小柱壳仍有体积，退化只发生在参数表示上。

第 20 讲

特征值与特征向量

Eigenvalues and Eigenvectors

对应材料：MIT 18.06SC Session 2.8（Strang 第 21 讲视频 Eigenvalues and Eigenvectors）课时摘要；习题 21.1、21.2 及解答（Strang《Introduction to Linear Algebra》6.1 节第 19、29 题）；补充材料依 Strang 教材 6.1 节。

本讲目标

(1) 理解特征向量的含义：Ax 与 x 方向相同（或相反）的非零向量 x；会验证一个向量是不是特征向量，并知道 $\lambda=0$ 的特征向量就是零空间里的非零向量。

(2) 不做计算，直接从几何上读出投影矩阵、反射矩阵的特征值与特征向量。

(3) 掌握标准算法：解特征方程 $\det(A-\lambda I)=0$ 求 λ，再求 $A-\lambda I$ 的零空间得到特征向量。

(4) 会用「特征值之和 = 迹，特征值之积 = 行列式」检查答案；会由 A 的特征值求出 $A+cI$、A^k、A^{-1}、$p(A)$ 的特征值，并知道 $A+B$、AB 没有这样的规律。

(5) 认识两种「麻烦」：旋转矩阵的复特征值，三角矩阵中重复特征值导致的特征向量不足。

20.1 方向不变的向量

到目前为止，我们主要把矩阵看成方程组 $Ax=b$ 的系数表：给定 b，求 x。从这一讲开始，换一个角度：把 A 看成一台机器，输入向量 x，输出向量 Ax，就像函数输入 x 输出 $f(x)$ 一样。这门课余下的大部分内容都围绕下面这个问题展开：

有没有这样的输入 x，使输出 Ax 与 x 在同一条直线上?

一般的向量被 A 作用后会改变方向。少数特殊的向量方向不变，只是被拉长、缩短或反向。这些向量就是特征向量。先看一个例子。本讲行内把列向量写成 (x_1,x_2) 以节省篇幅。

■ 例 20.1 **哪些向量方向不变** 设 $A=\begin{bmatrix}4 & -2\\ 1 & 1\end{bmatrix}$。分别计算 A 作用在 $u=(1,1)$、$v=(2,1)$、$w=(1,0)$ 以及 $u+v=(3,2)$ 上的结果，判断哪些输出与输入平行。

解 逐个做矩阵乘向量（第 i 个分量是 A 的第 i 行与向量的内积）：

$$Au=\begin{bmatrix}4-2\\ 1+1\end{bmatrix}=\begin{bmatrix}2\\ 2\end{bmatrix}=2u,\qquad Av=\begin{bmatrix}8-2\\ 2+1\end{bmatrix}=\begin{bmatrix}6\\ 3\end{bmatrix}=3v,$$

$$Aw=\begin{bmatrix}4\\1\end{bmatrix},\qquad A(u+v)=\begin{bmatrix}12-4\\3+2\end{bmatrix}=\begin{bmatrix}8\\5\end{bmatrix}.$$

Au 是 u 的 2 倍，Av 是 v 的 3 倍，这两个输出都与输入平行。$Aw=(4,1)$ 不是 $(1,0)$ 的倍数（第二个分量不为 0）。$A(u+v)=(8,5)$ 也不是 $(3,2)$ 的倍数：若 $(8,5)=c(3,2)$，则第一个分量给出 $c=8/3$，第二个分量给出 $c=5/2$，矛盾。

所以 u 和 v 是「方向不变」的向量，放大倍数分别是 2 和 3；w 和 $u+v$ 都被 A 转了方向。特别注意最后一个：**两个特征向量的和一般不再是特征向量**。原因是 $A(u+v)=2u+3v$，两部分被放大的倍数不同，合起来就偏离了 $u+v$ 的方向。 ♦

图 20.1 把这几个向量画在一起。虚线是 u 和 v 所在的两条直线：A 把这两条直线各自映回自身，只是沿直线拉伸。其他方向（比如 w 的方向）都被转开了。

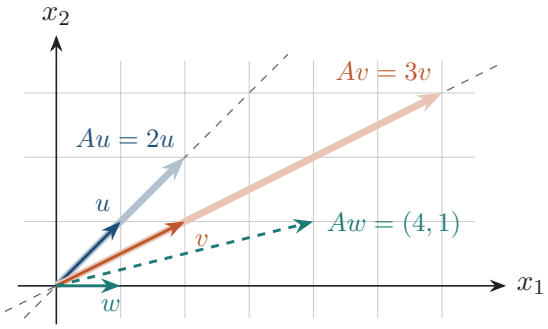


图 20.1　例 20.1 的矩阵 A 作用于三个向量（深色是输入，浅色或虚线是输出）。u、v 的方向不变，分别放大 2 倍和 3 倍；$w=(1,0)$ 被转到 $(4,1)$。

定义 20.1（特征值与特征向量）　设 A 是 $n\times n$ 矩阵。如果数 λ 和**非零**向量 $x\in\mathbb{R}^n$（或 $\mathbb{C}^n$）满足

$$Ax=\lambda x,$$

就称 λ 是 A 的一个**特征值**（eigenvalue），x 是 A 的**属于 λ 的特征向量**（eigenvector）。

在例 20.1 中，2 和 3 是 A 的特征值，$u=(1,1)$ 属于 2，$v=(2,1)$ 属于 3；w 和 $u+v$ 不是特征向量。关于这个定义有几点说明。

- **只对方阵讨论**。$Ax=\lambda x$ 要求 Ax 和 x 在同一个空间里才能比较，所以 A 必须是 $n\times n$ 的。
- λ 可以是正数、负数、0，甚至复数（§20.5）。$\lambda<0$ 表示方向反转，$|\lambda|<1$ 表示缩短。
- **特征向量的非零倍数仍是特征向量**，且属于同一个 λ：$A(cx)=c\,Ax=c\lambda x=\lambda(cx)$。所以特征向量真正代表的是一条过原点的直线（一个方向），而不是某一个向量。例 20.1 中 $(-3,-3)$、$(10,5)$ 同样是特征向量。

把 λx 写成 λIx，方程 $Ax=\lambda x$ 就等价于 $(A-\lambda I)x=0$。这说明属于同一个 λ 的特征向量，再添上零向量，恰好组成一个零空间。

定义 20.2（特征空间）　设 λ 是 A 的特征值。子空间

$$N(A-\lambda I)=\{\,x \mid Ax=\lambda x\,\}$$

称为 A 的属于 λ 的**特征空间**（eigenspace）。它的非零向量恰好是属于 λ 的全部特征向量。

想一想　定义要求 $x\neq 0$，却允许 $\lambda=0$。为什么这样规定？

答　若允许 $x=0$，则对任何数 λ 都有 $A0=\lambda\cdot 0$，每个数都成了特征值，定义失去意义。$\lambda=0$ 则是有内容的：它说存在非零 x 使 $Ax=0$，即 A 的零空间不只含零向量。下一小节专门讨论这种情形。

20.1.1　特征值 0 与零空间

$\lambda=0$ 时 $Ax=\lambda x$ 变成 $Ax=0$。所以**属于 0 的特征向量就是零空间 $N(A)$ 中的非零向量**，属于 0 的特征空间就是 $N(A)$ 本身。第 6、7 讲研究的零空间，现在有了一个新名字。

命题 20.3　方阵 A 以 0 为特征值，当且仅当 A 是奇异矩阵（不可逆）。等价地，A 可逆当且仅当 0 不是 A 的特征值。

证明　0 是特征值 $\iff$ 存在 $x\neq 0$ 使 $Ax=0$ $\iff$ $N(A)\neq\{0\}$ $\iff$ A 的列线性相关 $\iff$ A 不可逆（第 9 讲、第 3 讲）。 ■

■ **例 20.2　奇异矩阵**　求 $A=\begin{bmatrix}1&2\\2&4\end{bmatrix}$ 的两个特征值和对应的特征向量。

解　第二行是第一行的 2 倍，$\det A=4-4=0$，A 是奇异的，所以 0 是一个特征值。属于 0 的特征向量就是 $Ax=0$ 的非零解。由第一行 $x_1+2x_2=0$，取 $x_2=-1$ 得 $x_1=2$，即 $x=(2,-1)$。验证：$A(2,-1)=(2-2,\ 4-4)=(0,0)$。

另一个特征向量要从「输出在哪里」去找。A 的两列 $(1,2)$ 和 $(2,4)$ 都是 $(1,2)$ 的倍数，所以**任何输出 Ax 都落在直线 $\mathrm{span}\{(1,2)\}$ 上**。如果某个特征向量属于非零的 λ，那么 $x=\frac{1}{\lambda}Ax$ 也在这条直线上。于是试 $x=(1,2)$：

$$A\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}1+4\\2+8\end{bmatrix}=\begin{bmatrix}5\\10\end{bmatrix}=5\begin{bmatrix}1\\2\end{bmatrix}.$$

所以第二个特征值是 5，特征向量 $(1,2)$。两个特征值是 0 和 5。 ♦

这个例子里的两个特征向量分别来自零空间 $N(A)$ 和列空间 $C(A)$。第 10 讲的四个子空间在这里又出现了。

直观理解　特征向量是矩阵 A 「自带的坐标轴」。沿着这些方向，A 的作用退化成乘一个数：

$$Ax=\lambda x,\qquad A^2x=A(\lambda x)=\lambda^2 x,\qquad A^{100}x=\lambda^{100}x.$$

直接算 A^{100} 要做 99 次矩阵乘法；可是若把向量分解成特征向量的组合，每一部分只需乘 λ^{100}。n 维的问题拆成了 n 个一维问题。第 21 讲（对角化）和第 22 讲（微分方程）都建立在这个想法上。

20.2　从几何直接读出特征值：投影与反射

求特征值的一般方法要到 §20.3 才讲。在那之前，先看两类矩阵：它们的几何意义非常清楚，不必计算就能找到全部特征向量。

20.2.1　投影矩阵

设 P 是 $\mathbb{R}^3$ 中到某个过原点的平面上的投影矩阵（第 14、15 讲）。哪些向量 x 投影后与自己平行？

- **平面内的向量**：投影不改变它们，$Px=x$。所以平面内每个非零向量都是特征向量，特征值 $\lambda=1$。属于 1 的特征空间就是这个平面，是二维的。
- **垂直于平面的向量**：投影把它们压成零，$Px=0=0x$。它们是特征向量，特征值 $\lambda=0$。属于 0 的特征空间是平面的法线。
- **没有别的特征向量**：设 $Px=\lambda x$，$x\neq 0$。若 $\lambda\neq 0$，则 $x=\frac{1}{\lambda}Px$ 落在平面内，属于第一种；若 $\lambda=0$，则 $Px=0$，x 垂直于平面，属于第二种。

注意 P 的特征向量（平面内取两个无关向量，再加一个法向量）可以张成整个 $\mathbb{R}^3$。这一点并非对所有矩阵都成立，§20.6 会看到反例。

■ **例 20.3　正交投影的具体矩阵**　设 $P=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}$。不展开行列式，求全部特征向量。

解　这是到直线 $x_1=x_2$ 的正交投影，因为 $P(x_1,x_2)=((x_1+x_2)/2,(x_1+x_2)/2)$。对 $u=(1,1)$，$Pu=u$；对 $v=(1,-1)$，$Pv=0$。任意向量唯一写成 $au+bv$，若它是特征向量，就有 $au=\lambda(au+bv)$。比较 u,v 的系数：$a=\lambda a$、$0=\lambda b$。非零解只可能是 $b=0,\lambda=1$，或 $a=0,\lambda=0$。所以全部特征向量为 $t(1,1)$ 与 $t(1,-1)$（$t\neq 0$），没有第三种方向。 ♦

20.2.2　反射保留一个方向，反转另一个方向

交换两个坐标的矩阵 $B=\begin{bmatrix}0&1\\1&0\end{bmatrix}$ 将平面关于直线 $x_1=x_2$ 反射。直线上向量不动，垂直方向向量反向。图 20.2 将反射与投影放在相同坐标系里比较。

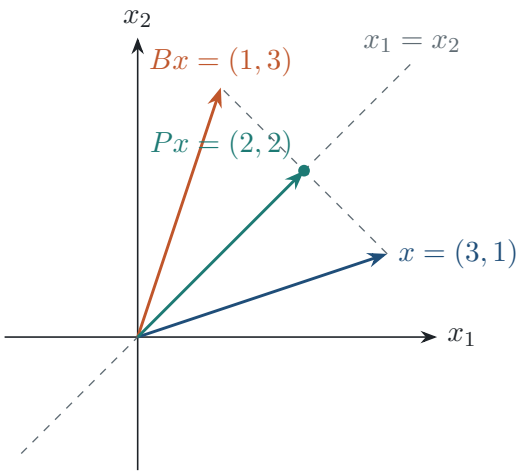


图 20.2　投影点是输入端点与反射端点的中点，反射保留平行分量并反转垂直分量。

图中的三个端点均可直接乘矩阵核对。反射端点与输入端点关于直线对称，投影点位于二者中间；坐标原点是三个箭头的共同起点。

■ **例 20.4　反射与投影的联系**　求 $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 的特征值与特征向量，并说明为什么 $P = (I + B)/2$。

解　直接验证 $B(1,1) = (1,1)$、$B(1,-1) = -(1,-1)$，故特征值为 $1, -1$，相应特征空间分别为两条直线。它们张成平面，比较分量即可排除其他方向。矩阵相加给 $(I+B)/2 = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = P$。在第一条直线上，该平均把倍数 1 变成 $(1+1)/2 = 1$；在第二条直线上，把 -1 变成 $(1-1)/2 = 0$。这解释了投影为何保留平行分量、消去垂直分量。 ♦

20.3　特征方程：先求倍数，再求方向

例 20.1 中两个特殊倍数为 $2, 3$。如果没有事先给出的向量，怎样发现这两个数？先把 $Ax = \lambda x$ 移到一边，得到 $(A - \lambda I)x = 0$。我们要的是非零解，因此必须让这个矩阵失去可逆性。

> **定理 20.4（特征方程）**　设 A 为 $n \times n$ 矩阵，标量取实数或复数。λ 是特征值，当且仅当 $\det(A - \lambda I) = 0$。求出每个根后，属于它的全部特征向量是 $N(A - \lambda I)$ 中的全部非零向量。

证明　λ 是特征值，当且仅当存在 $x \neq 0$ 满足 $(A - \lambda I)x = 0$；这等价于 $A - \lambda I$ 奇异，也等价于其行列式为零。固定 λ 后，等式正是该矩阵的齐次方程，因而全部解由零空间给出，排除零向量即可。 ■

■ **例 20.5　原课的对称矩阵**　求 $H = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ 的全部特征值与特征向量。

解　先求倍数：$\det(H - \lambda I) = (3-\lambda)^2 - 1 = (\lambda - 4)(\lambda - 2)$，所以根为 $4, 2$。对 $\lambda = 4$，方程是

$-x_1+x_2=0$，特征空间为 $\operatorname{span}\{(1,1)\}$；对 $\lambda=2$，方程是 $x_1+x_2=0$，特征空间为 $\operatorname{span}\{(1,-1)\}$。两条直线上的所有非零向量都是特征向量。直接核对 $H(1,1)=(4,4)$、$H(1,-1)=(2,-2)$。先解标量方程再解线性方程，避免把 x 和 λ 同时当未知数进行非线性消元。 ♦

求根并没有求完特征向量　特征方程只给出倍数，方向仍须求零空间。例如 H 的根为 4 时，$x=(1,0)$ 并不满足 $Hx=4x$。另外，不能把 $\det(A-\lambda I)$ 写成 $\det A-\lambda\det I$：行列式一般不对整个矩阵的加法保持线性。对 H，正确表达式是 $\lambda^2-6\lambda+8$，错误表达式却为 $8-\lambda$。

20.4　迹与行列式：两个快速检查

矩阵 H 的特征值是 $4,2$，和为 6，积为 8；恰与对角线之和及行列式一致。对例 20.1 的矩阵，特征值 $2,3$ 的和为 $5=4+1$，积为 $6=4\cdot1-(-2)\cdot1$。下面说明这两个检验来自特征多项式的系数。

定义 20.5（特征多项式与代数重数）　约定 $p_A(t)=\det(tI-A)$，它是首一的 n 次多项式。其根的重数称为特征值的代数重数。矩阵的迹为 $\operatorname{tr}A=\sum_i a_{ii}$。用 $\det(A-tI)$ 也能求相同的根，但整个多项式相差 $(-1)^n$。

例如 $A=3I_2$ 的特征多项式为 $(t-3)^2$，只有一个不同的特征值 3，代数重数为 2。说二阶矩阵有两个特征值，必须注明按代数重数计数，并允许复数根。实数范围内可能没有任何特征值，下一节的旋转会给出例子。

定理 20.6（和与积）　设 A 为 n 阶实或复矩阵，在复数范围内将其特征值按代数重数列为 $\lambda_1,\dots,\lambda_n$。则

$$\sum_{j=1}^{n}\lambda_j=\operatorname{tr}A,\qquad\prod_{j=1}^{n}\lambda_j=\det A.$$

证明　由代数基本定理，$p_A(t)=\prod_j(t-\lambda_j)$。右边 t^{n-1} 的系数为 $-\sum_j\lambda_j$。左边的排列公式中，恒等排列贡献 $\prod_i(t-a_{ii})$，其 t^{n-1} 系数为 $-\operatorname{tr}A$；非恒等排列至少移动两个位置，最多含 $n-2$ 个对角因子，无法贡献 t^{n-1}。比较该系数得到和公式。再令 $t=0$：左边为 $(-1)^n\det A$，右边为 $(-1)^n\prod_j\lambda_j$，消去符号得到积公式。证明不要求矩阵有足够多的特征向量。 ■

■ **例 20.6　二阶矩阵由和与积确定根**　已知某二阶矩阵的迹为 5、行列式为 6，求其特征值。若将矩阵阶数改为三阶，这两个数还够吗？

解　二阶时特征多项式为 $t^2-5t+6=(t-2)(t-3)$，因此根是 $2,3$。三阶时一般还缺 t 项系数。比如 $\operatorname{diag}(1,1,3)$ 的迹为 5、行列式为 3；$\operatorname{diag}(-1,3,3)$ 的迹也为 5，但积不同，说明迹单独不够。进一步固定积也仍有自由度：特征值可为 $1,1,3$，也可为 $-1,3+2\sqrt{3},3-2\sqrt{3}$；后一组和是 5，积是 3，却与前一组不同。因此在高阶题中，和与积适合核对，不能独自恢复整个谱。 ♦

相同的和与积能证明向量正确吗　若算出的两个数通过迹与行列式检查，是否还需要把特征向量代回？

答　需要。二阶时和与积可以验证两个根的集合，却完全不检验所配的向量。对 H，把 4 配给 $(1,-1)$、把 2 配给 $(1,1)$，两个根仍通过检查，但两组向量都配错了。每组必须分别代回 $Ax=\lambda x$。高阶时连根也可能只通过两个必要条件而仍错误。

20.5　旋转为什么需要复数

投影和反射都有实特征方向，但并非每个实矩阵都有这样的方向。最清楚的反例是逆时针旋转直角。它保持长度，却把每条实直线转到垂直方向，所以没有非零实向量能留在原直线上。零向量仍被映到零，但定义已经排除了它。

■ **例 20.7　直角旋转的两种观察方式**　求 $Q=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ 的特征值，并解释实平面图像与复数计算如何相容。

解　对任意实向量 $x=(a,b)$，$Qx=(-b,a)$，内积 $x\cdot Qx=-ab+ab=0$。当 $x\neq 0$ 时，Qx 非零且垂直于 x，因此不可能是 x 的实数倍。代数上 $\det(Q-\lambda I)=\lambda^2+1$，确实没有实根；允许复数后，根为 $i,-i$。当 $\lambda=i$，第一行给 $-ix_1-x_2=0$，取 $x_1=1$ 得 $x_2=-i$。直接验证

$$Q\begin{bmatrix}1\\-i\end{bmatrix}=\begin{bmatrix}i\\1\end{bmatrix}=i\begin{bmatrix}1\\-i\end{bmatrix}.$$

另一特征空间为 $\operatorname{span}_{\mathbb{C}}\{(1,i)\}$，对应 $-i$。这些向量属于 $\mathbb{C}^2$，没有作为实平面箭头的表示；不能把复向量的两个分量误当成一个复数的实部、虚部。 ♦

图 20.3 将实平面中的输入与输出分开画出，可以直接检查垂直关系与长度保持。

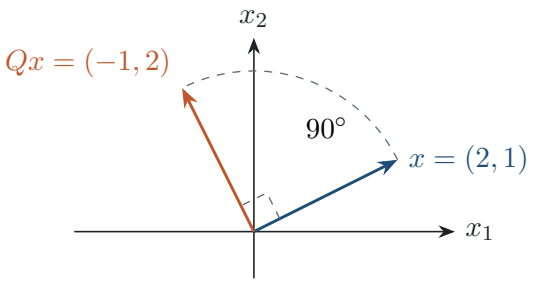


图 20.3　实平面中的旋转：输入与输出垂直，长度都为 $\sqrt{5}$。图只表示实向量，不试图把 $\mathbb{C}^2$ 的特征向量画成实平面箭头。

复特征值并没有改变实矩阵的作用，只是把可以用来研究它的标量范围扩大了。对任意实矩阵，如果 $Az=\lambda z$，逐项取共轭得 $A\bar z=\bar\lambda\bar z$；$z\neq 0$ 保证 $\bar z\neq 0$。因此非实根成共轭对出现，而且特征多项式为实系数多项式，配对根的代数重数相同。旋转例的和 $i+(-i)=0$、积 $i(-i)=1$，与迹和行列式一致。

定理 20.7（对称与反对称矩阵的根） 实对称矩阵的特征值全为实数；实反对称矩阵 $K^{\mathsf{T}}=-K$ 的特征值实部全为零，可以为 0。

证明 允许特征向量 z 为复数，记 $z^{*}=\bar{z}^{\mathsf{T}}$，则 $z^{*}z=\sum_j |z_j|^2>0$。若 $Az=\lambda z$，则 $z^{*}Az=\lambda z^{*}z$。实对称时 $A^{*}=A$，故 $\overline{z^{*}Az}=z^{*}A^{*}z=z^{*}Az$，该标量为实数，从而 λ 为实数。反对称时 $K^{*}=-K$，同理 $\overline{z^{*}Kz}=-z^{*}Kz$，故 $\bar{\lambda}=-\lambda$，即实部为零。这个证明只用内积，不依赖以后才讲的对角化定理。 ■

20.6　重根与三角矩阵

三角矩阵的行列式公式让特征方程立即分解。若 A 是上三角矩阵，则 $tI-A$ 也是上三角矩阵，行列式等于对角线元素之积：

$$p_A(t)=\det(tI-A)=\prod_{j=1}^{n}(t-a_{jj}).$$

因此对角线给出全部特征值，按重数计数。这个结论很快，但不能由此断定存在足够多的特征向量：重复根会提出另一个问题，零空间的维数是多少？

■ **例 20.8　同一重根，特征方向数不同**　比较 $J=\begin{bmatrix}3&1\\0&3\end{bmatrix}$ 与 $3I_2$。

解 两者的特征多项式都是 $(t-3)^2$，唯一特征值为 3，代数重数均为 2。对 J，$(J-3I)x=0$ 化为 $x_2=0$，故特征空间只有 $\operatorname{span}\{(1,0)\}$，维数为 1。对 $3I_2$，$(3I_2-3I)x=0$ 对任意 x 成立，特征空间为整个平面，维数为 2。所以重根不表示只有一个方向，也不表示必有两个独立方向，仍须解零空间。 ♦

代数重数不等于特征空间维数　J 的重根为二重，但只有一条特征直线。若把重数直接当成独立特征向量个数，就会误以为每个矩阵都可对角化。反过来，特征空间维数也不能超过该特征值的代数重数；本讲只需记住应分别由特征多项式和零空间计算。

■ **例 20.9　原课的三阶上三角矩阵**　求 $A=\begin{bmatrix}1&2&3\\0&4&5\\0&0&6\end{bmatrix}$ 的特征值，并给每个特征值找一个特征向量。

解 由上三角矩阵的行列式公式，$p_A(t)=(t-1)(t-4)(t-6)$，特征值为 $1,4,6$。分别解齐次方程，取向量 $v_1=(1,0,0)$、$v_4=(2,3,0)$、$v_6=(16,25,10)$。逐个乘回去：

$$Av_1=v_1,\quad Av_4=(8,12,0)=4v_4,\quad Av_6=(96,150,60)=6v_6.$$

三个根不同，所给向量的检验也直接成立。上三角矩阵只保证根来自对角线，不会自动告诉我们每个零空间的具体向量。 ♦

能由对角线读出什么　对角线出现 $2,2,5$，能否由此断定特征值 2 的特征空间是二维？

答　不能。对角线只能读出根和代数重数；还需解 $(A-2I)x=0$。例如 $\begin{bmatrix}2&1&0\\0&2&0\\0&0&5\end{bmatrix}$ 的前两行给 $x_2=0$，重根 2 只有一维特征空间。

20.7　从已知特征值推导新矩阵

若同一个向量 $x\neq 0$ 满足 $Ax=\lambda x$，则反复作用得到 $A^kx=\lambda^kx$。把这些等式按多项式的系数相加，即得 $p(A)x=p(\lambda)x$。常数项对应 $Ix=x$；因此 $A+cI$ 的同一方向乘数为 $\lambda+c$。若 A 可逆，从 $Ax=\lambda x$ 左乘 A^{-1} 得 $x=\lambda A^{-1}x$；此时 $\lambda\neq 0$，故 $A^{-1}x=\lambda^{-1}x$。这些结论都保持原来的特征向量。

■ 例 20.10　**同一方向上的移位与平方**　已知 $H=\begin{bmatrix}3&1\\1&3\end{bmatrix}$ 的特征值为 $4,2$。求 $H-3I$、H^2、$(H^2+I)^{-1}$ 的特征值。

解　在 $(1,1)$ 方向，三个乘数分别为 $1,16,1/17$；在 $(1,-1)$ 方向，分别为 $-1,4,1/5$。这两向量独立，张成整个平面，故已列出全部特征值，且 H^2+I 在这两个方向的乘数均非零，保证可逆。第一项正是先前的反射矩阵；沿两条垂直方向，减去 $3I$ 将伸缩倍数 $4,2$ 平移到 $1,-1$，并非在空间中把箭头端点平移固定距离。♦

一般矩阵的和与积不能这样逐个配对。令 $N=\begin{bmatrix}0&1\\0&0\end{bmatrix}$、$M=\begin{bmatrix}0&0\\1&0\end{bmatrix}$，两者都只有特征值 0，但 $N+M$ 的根为 $1,-1$，$NM=\operatorname{diag}(1,0)$ 的根为 $1,0$。前面的多项式结论依赖同一个特征向量。若确有 $Ax=\lambda x$、$Bx=\mu x$，则 $(A+B)x=(\lambda+\mu)x$、$ABx=\lambda\mu x$，这才是可用的共同方向条件。

■ 例 20.11　**原课习题 21.1：哪些量已经确定**　三阶实矩阵 B 的特征值为 $0,1,2$。能否确定其秩、$\det(B^{\mathsf T}B)$、$B^{\mathsf T}B$ 的特征值以及 $(B^2+I)^{-1}$ 的特征值？

解　零根使 B 奇异，所以秩至多为 2。取属于 $1,2$ 的非零特征向量 u,v；它们独立，否则同一非零方向同时有两个倍数，矛盾。又 $Bu=u$、$Bv=2v$，故两者都属于列空间，秩至少为 2。综上秩恰为 2。行列式乘法给 $\det(B^{\mathsf T}B)=(\det B)^2=0$。

但 Gram 矩阵的全部特征值无法确定。比较 $D=\operatorname{diag}(0,1,2)$ 和 $E=\begin{bmatrix}0&1&0\\0&1&0\\0&0&2\end{bmatrix}$：两者都是上三角矩阵，对角线同为 $0,1,2$。而 $D^{\mathsf T}D=\operatorname{diag}(0,1,4)$，$E^{\mathsf T}E=\operatorname{diag}(0,2,4)$，根的集合不同。原矩阵的特征值不能独自确定各方向上的长度变化。

最后，属于 $0,1,2$ 的三个特征向量线性独立：若 $ax_0+bx_1+cx_2=0$，作用 B 与 B^2 后相减给 $2cx_2=0$，再依次得 $b=a=0$。它们张成三维空间。B^2+I 在这三个基方向的乘数为 $1,2,5$，均非零，所以它可逆；逆的全部特征值为 $1,1/2,1/5$。♦

■ **例 20.12　原课习题 21.2：结构比展开更清楚**　求 $B=\begin{bmatrix}0&0&1\\0&2&0\\3&0&0\end{bmatrix}$ 与全为 2 的三阶矩阵 C 的特征值。

解　B 保持第二坐标轴，$Be_2=2e_2$。在 $x_2=0$ 的平面上，它把 $(a,0,c)$ 变成 $(c,0,3a)$，所以特征方程要求 $c=\lambda a$、$3a=\lambda c$，从而 $\lambda^2=3$。取 $a=1$，得到对应 $\sqrt{3},-\sqrt{3}$ 的向量 $(1,0,\sqrt{3})$、$(1,0,-\sqrt{3})$。也可沿第二行展开，得到 $\det(B-\lambda I)=(2-\lambda)(\lambda^2-3)$。三个根的和为 2，积为 -6，与迹和行列式一致。

对 C，若 $s=x_1+x_2+x_3$，则 $Cx=2s(1,1,1)$。故 $(1,1,1)$ 对应根 6，平面 $s=0$ 上所有非零向量对应根 0。这个平面由 $(1,-1,0)$、$(1,0,-1)$ 张成，连同 $(1,1,1)$ 构成三维基，所以全部根为 $6,0,0$。此法说明了重根来自一个二维零空间，比三阶行列式展开更直观。若使用 $\det(C-\lambda I)$，正确表达式为 $\lambda^2(6-\lambda)$；使用首一多项式 $\det(\lambda I-C)$ 才得到 $\lambda^2(\lambda-6)$。两者根相同，三次项符号相反。 ♦

本讲求特征对的流程可以整理成四步。先观察矩阵是否为投影、反射或三角矩阵，利用结构找到候选根；再用特征方程确认根与代数重数；随后逐根求齐次方程的零空间；最后分别检验每一组向量并检查迹、行列式。结构观察减少计算，代回核对排除配错方向，两者相互补充。下一讲会把独立特征向量排列成一组新基，研究何时可以把矩阵化成对角矩阵。

本讲小结

(1) 特征向量 x 必须非零，满足 $Ax=\lambda x$。特征值可为零，属于零根的特征空间是 $N(A)$。

(2) 投影的根为 $0,1$，反射的平行与垂直方向分别对应 $1,-1$。几何含义有时直接给出全部特征方向。

(3) 先解 $\det(A-\lambda I)=0$，再解 $(A-\lambda I)x=0$。零空间包括零向量，特征向量集合须排除零向量。

(4) 在复数范围内按代数重数计数，根之和为迹、根之积为行列式；两条检查一般不充分。

(5) 实旋转可以没有实特征方向。实矩阵的非实根共轭配对；实对称矩阵根为实数，实反对称矩阵根为纯虚数或零。

(6) 三角矩阵的根为对角线元素，重根的特征空间维数仍须计算。多项式 $p(A)$ 保持每个已有特征向量，并把倍数变成 $p(\lambda)$；可逆时逆矩阵把倍数变成倒数。

练习

20.1 ★　对 $A=\begin{bmatrix}4&-2\\1&1\end{bmatrix}$，判断 $(1,1)$、$(2,1)$、$(3,2)$ 是否为特征向量，并写出相应倍数或说明矛盾。

20.2 ★　求 $T=\begin{bmatrix}2&1\\0&5\end{bmatrix}$ 的特征值与每个特征空间；用迹、行列式及代回验证。

20.3 ★ 求到 x 轴的正交投影矩阵，以及关于 x 轴反射的矩阵；分别列出全部实特征向量。

20.4 ★ 已知二阶矩阵的一个特征值为 -2、迹为 3。求另一个特征值与行列式。能否因此求出特征向量？

20.5 ★★ 求 $R=\begin{bmatrix}0&-2\\2&0\end{bmatrix}$ 的复特征值与特征空间，并证明不存在实特征向量。画出 $(1,0)$ 与其像，标出长度和夹角。

20.6 ★★ 比较 $U=\begin{bmatrix}2&1&0\\0&2&0\\0&0&5\end{bmatrix}$ 与 $V=\operatorname{diag}(2,2,5)$：写出特征多项式、每个特征空间的基与维数。解释相同根的重数为何不能代替零空间计算。

20.7 ★★ 对 $H=\begin{bmatrix}3&1\\1&3\end{bmatrix}$，求 $H+I$、H^3、H^{-1} 的特征值，并写出共同特征向量。说明为什么这已包括全部特征值。

20.8 ★★ 用 $N=\begin{bmatrix}0&1\\0&0\end{bmatrix}$、$M=\begin{bmatrix}0&0\\1&0\end{bmatrix}$ 检验「矩阵和的根逐个相加、积的根逐个相乘」。写出四个矩阵的特征多项式，不能只给结论。

20.9 ★★★ 设 $A^2=A$。证明其特征值只能为 $0,1$。再证明任意向量可唯一写成 $x=u+v$，其中 $Au=u$、$Av=0$；说明零矩阵和单位阵是否必须同时有两个不同的根。

20.10 ★★★ 设实三阶矩阵的三个根为 $0,1,2$。证明秩恰为 2，求 $(A^2+I)^{-1}$ 的根；给出两个此类矩阵，使 $A^{\mathsf T}A$ 的根不同。证明所用三个特征向量确实独立。

练习答案与提示

20.1 前两者是，因 $A(1,1)=(2,2)=2(1,1)$、$A(2,1)=(6,3)=3(2,1)$。第三者不是，因为 $A(3,2)=(8,5)$；若为倍数，则同一倍数须同时为 $8/3$ 和 $5/2$，不相等。第三向量恰为前两者之和，但不同特征值使两个分量以不同倍数变化。这也展示了「特征向量的和还是特征向量」需要额外条件：若来自同一特征空间，和非零时才仍属于该空间的非零部分。

20.2 $\det(T-\lambda I)=(2-\lambda)(5-\lambda)$，故根为 $2,5$。根为 2 时 $(T-2I)x=0$ 给 $x_2=0$，空间为 $\operatorname{span}\{(1,0)\}$；根为 5 时第一行给 $-3x_1+x_2=0$，空间为 $\operatorname{span}\{(1,3)\}$。代回得 $T(1,0)=(2,0)$、$T(1,3)=(5,15)$。根之和 $7=\operatorname{tr}T$、根之积 $10=\det T$。全部特征向量分别是这两个基向量的非零数倍；注意不能只把所给基向量称为全部向量。

20.3 投影矩阵为 $P=\operatorname{diag}(1,0)$，因为 (a,b) 映到 $(a,0)$；反射为 $S=\operatorname{diag}(1,-1)$，因为映到 $(a,-b)$。对投影，x 轴上的非零向量对应 1，y 轴上的非零向量对应 0。对反射，两轴分别对应 $1,-1$。若 a,b 均非零，则投影的两个分量要求倍数同时等于 $1,0$，反射要求同时等于 $1,-1$，均矛盾，所以没有其他特征向量。投影消去垂直分量，反射把该分量反向。

20.4 设另一根为 μ，由 $-2+\mu=3$ 得 $\mu=5$，由乘积公式得行列式为 -10。这些信息没有确定特征向量。例如 $\operatorname{diag}(-2,5)$ 的特征轴是坐标轴，而 $\begin{bmatrix}-2&7\\0&5\end{bmatrix}$ 的 5 特征空间为 $\operatorname{span}\{(1,1)\}$，两者迹与根完全相同，但方向不同。因此可以恢复二阶根的集合，不能恢复矩阵的全部几何作用。要求向量仍须知道矩阵并解零空间。

20.5 $\det(R-\lambda I)=\lambda^2+4$，根为 $2i,-2i$。对 $2i$，第一行给 $x_2=-ix_1$，空间为 $\operatorname{span}_{\mathbb C}\{(1,-i)\}$；对 $-2i$，

空间为 $\mathrm{span}_{\mathbb{C}}\{(1,i)\}$。直接乘回即可检查。实向量 $x=(a,b)\neq 0$ 的像为 $(-2b,2a)$，与 x 垂直且非零，不能平行，所以没有实特征向量。图应把 $(1,0)$ 画为横向长 1 的箭头，把像 $(0,2)$ 画为向上长 2 的箭头，二者夹角为逆时针 90°。复特征值的模为 2，与统一的长度放大倍数一致。

20.6 两个矩阵都有 $p(t)=(t-2)^2(t-5)$。对 U，根 2 的方程为 $x_2=x_3=0$，基为 e_1，维数 1；根 5 的方程给 $x_1=x_2=0$，基为 e_3，维数 1。对 V，根 2 的方程只有 $x_3=0$，基为 e_1,e_2，维数 2；根 5 仍以 e_3 为基，维数 1。两者根的代数重数相同，但特征空间的维数不同，原因是 U 在前两坐标间还有一个非零耦合项。这个项不改变三角矩阵的行列式，却改变了齐次方程。

20.7 H 在 $u=(1,1)$、$v=(1,-1)$ 方向分别乘 $4,2$，所以 $H+I$ 的根为 $5,3$，H^3 为 $64,8$，H^{-1} 为 $1/4,1/2$。三个矩阵共同保留 u,v 两方向。因为 u,v 独立，它们构成平面的一组基，各矩阵相对于该基是这两个倍数构成的对角矩阵，所以已列出全部根。逆存在也可由 $\det H=8\neq 0$ 检查。这里的倒数运算只在矩阵可逆时使用；若有零根，则不能照样取倒数。

20.8 两个原矩阵都是三角矩阵，$p_N(t)=p_M(t)=t^2$。和为 $\begin{bmatrix}0&1\\1&0\end{bmatrix}$，其特征多项式 t^2-1，实际根为 $1,-1$；乘积 $NM=\mathrm{diag}(1,0)$，其多项式 $t(t-1)$，根为 $0,1$。把原来根逐个相加或相乘都只会给出两个零，因此两条说法均不成立。共同向量条件可使个别配对成立：如果同一 x 满足 $Nx=\lambda x$ 与 $Mx=\mu x$，才能推出相应和、积的等式；本反例的两个零特征空间是不同坐标轴，没有非零公共向量。

20.9 若 $Ax=\lambda x$ 且 $x\neq 0$，则 $A^2x=\lambda^2x=Ax=\lambda x$，故 $\lambda(\lambda-1)=0$。对任意 x，取 $u=Ax$、$v=x-Ax$，由幂等性得 $Au=A^2x=Ax=u$、$Av=Ax-A^2x=0$。若还有 $x=u'+v'$，作用 A 得 $Ax=u'$，所以 $u'=u$、$v'=v$，唯一性成立。这同时说明空间分为保留部分与消去部分。但某一部分可以为零子空间：零矩阵只有根 0，单位矩阵只有根 1；「只能属于某集合」不能写成「集合中每个数都一定出现」。

20.10 取三个非零特征向量 x_0,x_1,x_2。若 $ax_0+bx_1+cx_2=0$，作用 A、A^2 后分别得 $bx_1+2cx_2=0$、$bx_1+4cx_2=0$，相减给 $c=0$，再得 $b=a=0$，故独立。零根使秩至多为 2，两个非零根的独立向量均位于列空间，使秩至少为 2。在此特征基下，A^2+I 的倍数为 $1,2,5$，故其逆的根为 $1,1/2,1/5$。取 $D=\mathrm{diag}(0,1,2)$ 和 $E=\begin{bmatrix}0&1&0\\0&1&0\\0&0&2\end{bmatrix}$，两者上三角且根同为 $0,1,2$，但 $D^{\mathsf T}D=\mathrm{diag}(0,1,4)$，$E^{\mathsf T}E=\mathrm{diag}(0,2,4)$，所求两组根不同。两个 Gram 矩阵的行列式均为零，与原矩阵奇异一致。

第 21 讲

对角化与矩阵幂

Diagonalization and Powers of A

对应材料：MIT 18.06SC Session 2.9（Strang 第 22 讲）课时摘要、习题 22.1/22.2 及解答；补充二阶递推与坐标解释。

本讲目标

(1) 从列向量推导 $AS=S\Lambda$ 与 $A=S\Lambda S^{-1}$，理解新坐标、旧坐标各代表什么。

(2) 判断何时可以对角化，区分特征值重数与独立特征向量数。

(3) 用对角化计算 A^k，说明矩阵幂趋于零的条件以及初始向量造成的例外。

(4) 解差分方程 $u_{k+1}=Au_k$，把初始条件转换为特征方向上的系数。

(5) 把 Fibonacci 二阶递推改写为一阶矩阵系统，推导精确通项并解释增长率。

21.1 把特征方向选作坐标轴

上一讲已经找到了满足 $Ax=\lambda x$ 的特殊方向。若平面内有两个独立的这样的方向，我们能否沿它们建立坐标，让一次矩阵作用只需做两次标量乘法？这正是对角化的目的。新坐标通常不垂直，所以它不一定对应旋转纸面上的坐标轴；我们是在改用一组基描述同一个向量。

■ **例 21.1　先用两个特征方向描述向量**　设 $A=\begin{bmatrix}4&1\\2&3\end{bmatrix}$，$v_1=(1,1)$、$v_2=(1,-2)$。验证这两个方向，并用它们计算 $A(3,0)$。

解　直接相乘得 $Av_1=(5,5)=5v_1$、$Av_2=(2,-4)=2v_2$。两向量不共线，故构成平面的一组基。设 $(3,0)=c_1v_1+c_2v_2$，按分量比较得到 $c_1+c_2=3$、$c_1-2c_2=0$，解得 $c_1=2$、$c_2=1$。于是

$$A(3,0)=2Av_1+Av_2=10v_1+2v_2=(12,6).$$

同样，作用 k 次时两部分分别变为 $2\cdot5^kv_1$ 和 2^kv_2。这一次分解可以复用，无须每次重新算矩阵乘法。 ♦

把两个基向量排成列，令 $S=[v_1\ v_2]=\begin{bmatrix}1&1\\1&-2\end{bmatrix}$。若 $c=(c_1,c_2)$ 是新坐标，则原向量为 $u=Sc$。前例的新坐标是 $(2,1)$，原坐标是 $(3,0)$。矩阵 S 把系数变成向量，S^{-1} 把向量变回系数。不要将这两种坐标混写为同一组数。

定义 21.1（可对角化）　若存在可逆矩阵 S 与对角矩阵 Λ 满足 $S^{-1}AS=\Lambda$，称方阵 A 可对角化。等价写法为 $AS=S\Lambda$ 与 $A=S\Lambda S^{-1}$。在本讲默认实数范围内讨论；允许复数时会特别说明。

定理 21.2（对角化与特征向量基）　n 阶矩阵 A 可对角化，当且仅当它有 n 个线性独立的特征向量。若 $S=[v_1\ \cdots\ v_n]$，$Av_j=\lambda_j v_j$，则可取 $\Lambda=\operatorname{diag}(\lambda_1,\dots,\lambda_n)$，各列顺序必须与对角线顺序一致。

证明　按列计算 $AS=[Av_1\ \cdots\ Av_n]=[\lambda_1 v_1\ \cdots\ \lambda_n v_n]=S\Lambda$。独立性保证 S 可逆，左乘 S^{-1} 得所需等式。反过来，若 $S^{-1}AS=\Lambda$，则 $AS=S\Lambda$；比较第 j 列便得 $Av_j=\lambda_j v_j$，且可逆矩阵的列独立。这样两个方向都由同一个按列等式证明。 ■

对例 21.1，$\det S=-3$，因而

$$S^{-1}=\frac{1}{3}\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix},\qquad \Lambda=\begin{bmatrix}5 & 0\\ 0 & 2\end{bmatrix},\qquad A=S\Lambda S^{-1}.$$

具体核对可以先算 $S\Lambda=\begin{bmatrix}5 & 2\\ 5 & -4\end{bmatrix}$，再右乘 S^{-1}，得到 $\begin{bmatrix}4 & 1\\ 2 & 3\end{bmatrix}$。分解按从右到左的次序作用：先由 S^{-1} 找系数，再由 Λ 分别伸缩，最后由 S 组合回原坐标。图 21.1 将三个动作画成一条对应关系。

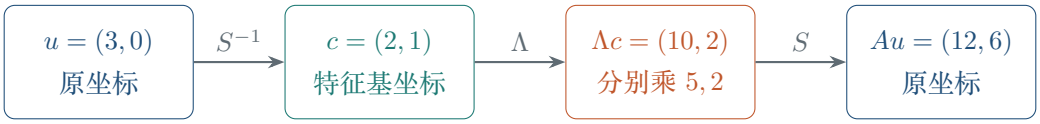


图 21.1　四组数描述同一次变换的输入、基系数、伸缩后系数与输出。S^{-1} 和 S 改变描述方式，Λ 执行各特征方向的伸缩。

把特征向量乘以常数会怎样　若把 v_1 改为 $7v_1$，矩阵 A 的对角化是否失效？

答　不会，非零数倍仍是相同特征值的特征向量，且不破坏独立性。新矩阵 S 的第一列变长，相应新坐标的第一分量变成原来的 $1/7$，组合出的原向量保持相同。S 与 S^{-1} 都随基的选择改变，A 本身不变。如果交换两列，就必须同时交换 Λ 的两项，否则按列等式 $AS=S\Lambda$ 不成立。

21.2　什么时候有足够多的特征方向

计算了特征方程，还需要检查所选向量能否构成一组基。重复写同一根的向量没有帮助。例如一个方向上取 v 与 $2v$，虽然两个都是特征向量，但排成 S 后两列相关，无法求逆。不同根提供了一条可靠的独立性保证。

定理 21.3（不同特征值的向量独立）　若 $Av_j=\lambda_j v_j$，$v_j\neq 0$，且 $\lambda_1,\dots,\lambda_m$ 两两不同，则 $v_1,\dots,v_m$ 线性独立。因此 n 阶矩阵若在所讨论数域中有 n 个不同的特征值，就可在该数域中对角化。

证明　对 m 归纳。$m=1$ 时非零向量独立。设 $\sum_{j=1}^{m} c_j v_j=0$，作用 $A-\lambda_m I$ 后得 $\sum_{j=1}^{m-1} c_j(\lambda_j-\lambda_m)v_j=0$。由归纳假设，所有 $c_j(\lambda_j-\lambda_m)=0$；根不同，故 $j<m$ 时 $c_j=0$。再代回原式得 $c_m v_m=0$，于是 $c_m=0$。归纳证明完成。取 $m=n$ 后，这些向量构成一组基，应用定理 21.2。 ■

■ **例 21.2　重复根不阻止对角化**　设 $B=\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}$。求一组特征向量基。

解　写成 $B=I+ee^{\mathsf T}$，其中 $e=(1,1,1)$。对 e，有 $Be=e+3e=4e$。对平面 $x_1+x_2+x_3=0$ 中的向量，$ee^{\mathsf T}x=0$，所以 $Bx=x$。取 $v_1=(1,-1,0)$、$v_2=(1,0,-1)$ 为这个平面的基，$v_3=e$ 为平面外方向。排成

$$S=\begin{bmatrix}1&1&1\\-1&0&1\\0&-1&1\end{bmatrix},\qquad \Lambda=\operatorname{diag}(1,1,4).$$

$\det S=3\neq 0$，所以 S 可逆，按列检查得 $BS=S\Lambda$，故可对角化。重根 1 的二维特征空间提供了两个独立方向。也可不用行列式验证独立性：前两向量张成和为零的平面，第三个向量各分量之和为 3，因此不在其中。 ♦

■ **例 21.3　相同特征值也可能缺少基向量**　比较 $J=\begin{bmatrix}2&1\\0&2\end{bmatrix}$ 与 $2I_2$ 的对角化情况。

解　两者的特征多项式均为 $(t-2)^2$。对 J，$(J-2I)x=0$ 化成 $x_2=0$，特征空间为 $\operatorname{span}\{(1,0)\}$，只有一个独立向量，不能组成二阶可逆矩阵 S。对 $2I_2$，每个非零向量都是特征向量，可取标准基 e_1,e_2，因此它已是对角矩阵。判断重根情形时必须计算零空间的维数；根重复本身既不保证可对角化，也不保证不可对角化。 ♦

A^2 可能有新的特征向量　若 $Ax=\lambda x$，则 $A^2x=\lambda^2 x$，所以 A 的特征向量仍是 A^2 的特征向量。但反向不成立：例如 $A=\operatorname{diag}(1,-1)$ 只有坐标轴上的非零特征向量，而 $A^2=I$ 的每个非零向量都是特征向量。因此要说两者「有共同的一组特征向量基」是正确的，说「全部特征向量集合相同」则错误。

21.3　对角化把矩阵幂变成标量幂

对角化中间的 $S^{-1}S$ 会相消，重复使用即可得到 $A^k=S\Lambda^k S^{-1}$。若要理解初始向量的演化，更直接的写法是 $u_0=Sc$，于是

$$u_k=A^k u_0=S\Lambda^k c=\sum_{j=1}^{n} c_j\lambda_j^k v_j.$$

这里 $c = S^{-1}u_0$；每个初始系数 c_j 只沿第 j 个特征方向乘以 λ_j^k。右端最终仍是原坐标中的向量，不能漏掉左侧的 S。

■ **例 21.4　逐个方向判断矩阵幂**　对例 21.1 的矩阵 A 求 A^k，并求其作用于 $u_0 = (3,0)$ 的结果。

解　已知 $A = S\,\mathrm{diag}(5,2)S^{-1}$，故 $A^k = S\,\mathrm{diag}(5^k, 2^k)S^{-1}$。代入 S, S^{-1} 得

$$A^k = \frac{1}{3}\begin{bmatrix} 2\cdot 5^k + 2^k & 5^k - 2^k \\ 2\cdot 5^k - 2\cdot 2^k & 5^k + 2\cdot 2^k \end{bmatrix}.$$

这个式子在 $k = 0$ 给单位阵，可用来检查幂次与系数。向量 $(3,0) = 2v_1 + v_2$，所以

$$A^k(3,0) = 2\cdot 5^k v_1 + 2^k v_2 = (2\cdot 5^k + 2^k,\ 2\cdot 5^k - 2\cdot 2^k).$$

$k = 1$ 得 $(12,6)$，与直接相乘相符；k 增大时，5^k 项相对 2^k 项增长更快，方向逐渐靠近 v_1。如果初始向量恰在 v_2 上，系数 $c_1 = 0$，则增长方向完全不同。 ♦

> **定理 21.4（可对角化情形的衰减判据）**　若 A 可对角化，则 $A^k \to 0$ 当且仅当每个特征值都满足 $|\lambda_j| < 1$。这里趋于零指矩阵每个元素都趋于零。

证明　若每个根的模小于 1，则对角矩阵 Λ^k 的每个元素趋于零，左右乘固定矩阵 S, S^{-1} 后，每个元素仍是有限个趋零项的线性组合，故 $A^k \to 0$。反过来，若 $A^k \to 0$，对任意非零特征向量 v_j，有 $A^k v_j = \lambda_j^k v_j \to 0$。选 v_j 的一个非零分量便得 $\lambda_j^k \to 0$，这要求 $|\lambda_j| < 1$。 ■

一个初值衰减能证明矩阵衰减吗　若 $A^k u_0 \to 0$，能否由此断定所有特征值的模都小于 1？

答　不能。取 $A = \mathrm{diag}(2, 1/2)$，$u_0 = (0,1)$，则 $A^k u_0 = (0, 2^{-k}) \to 0$，但根 2 的方向会增长。初值没有含该方向，所以看不到它。矩阵幂趋零是对所有初始向量的结论，单条轨迹趋零只约束该初值实际含有的方向。

■ **例 21.5　模小于、等于、大于一**　比较三个矩阵 $B_1 = \begin{bmatrix} .6 & .9 \\ .1 & .6 \end{bmatrix}$、$B_2 = \begin{bmatrix} .6 & .9 \\ .4 & .6 \end{bmatrix}$、$B_3 = \begin{bmatrix} .6 & .4 \\ .4 & .6 \end{bmatrix}$ 的长期行为。

解　B_1 的特征方向 $(3,1)$、$(3,-1)$ 分别对应 $.9, .3$，所以所有分量衰减，$B_1^k \to 0$。B_2 在 $(3,2)$、$(3,-2)$ 方向分别乘 $1.2, 0$；含第一方向的初值会增长，第二方向上的向量一步变为零。因此其矩阵幂不收敛。B_3 在 $(1,1)$、$(1,-1)$ 方向分别乘 $1, .2$，第一分量保持，第二分量衰减，故

$$B_3^k \longrightarrow \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

这个极限是到直线 $x_1 = x_2$ 的正交投影。根 1 不使矩阵幂趋零，却可以产生非零极限；根 -1 则使相应分量交替变号，一般不收敛。 ♦

21.4　差分方程：初值决定各方向的系数

离散时间中的线性演化写成 $u_{k+1}=Au_k$。从给定的 u_0 开始，依次有 $u_1=Au_0$、$u_2=A^2u_0$，归纳得 $u_k=A^ku_0$。若 A 可对角化，就把初值分解一次，后面的每个时刻都由标量幂描述。求解时先找方向，再求系数；只找根还不能得到符合给定初值的轨迹。

■ **例 21.6**　**两个变量逐渐达到共同值**　设 $u_{k+1}=Pu_k$，其中 $P=\begin{bmatrix}.8 & .2\\ .4 & .6\end{bmatrix}$，$u_0=(0,3)$。求全部 u_k 及其极限。

解　直接验证 $Pv_1=v_1$、$Pv_2=.4v_2$，其中 $v_1=(1,1)$、$v_2=(1,-2)$。初值分解为 $(0,3)=v_1-v_2$，所以

$$u_k=v_1-(.4)^kv_2=(1-(.4)^k,\ 1+2(.4)^k).$$

当 $k=0$，给出 $(0,3)$；$k=1$ 给出 $(.6,1.8)$，直接相乘也相同。因 $.4^k\to 0$，两分量均趋于 1，极限为 $(1,1)$。这不是初始两数的简单平均数 1.5：两个更新式采用的权重不同。事实上 $2u_{k,1}+u_{k,2}=3$ 对所有 k 保持，因为行向量 $(2,1)P=(2,1)$。极限又位于两分量相同的直线上，故必为 $(1,1)$。 ♦

图 21.2 画的是同一平面中的状态点，而非把时间当成横轴的函数图。每次状态都沿直线 $2x+y=3$ 靠近固定点，相对固定点的差向量每次缩短到 .4 倍；这个差向量沿 v_2，正是衰减特征方向。

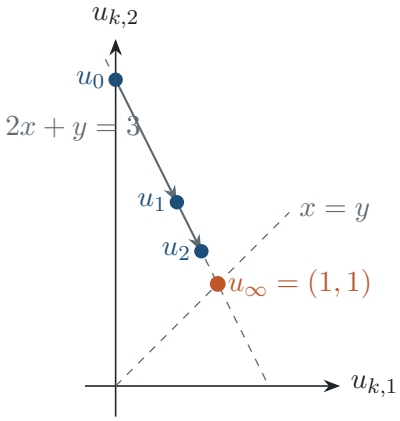


图 21.2　同一初值的状态轨迹：不衰减的特征分量决定终点，衰减的特征分量决定接近终点的方向与速度。

■ **例 21.7**　**原课习题 22.2：极限矩阵的两列**　对 $M=\begin{bmatrix}.6 & .9\\ .4 & .1\end{bmatrix}$ 求对角化、M^k 及其极限。

解　特征多项式为 $t^2-.7t-.3=(t-1)(t+.3)$。取 $v_1=(9,4)$、$v_2=(1,-1)$，乘回可验证对应根为 $1,-.3$。于是

$$S=\begin{bmatrix}9 & 1\\ 4 & -1\end{bmatrix},\quad S^{-1}=\frac{1}{13}\begin{bmatrix}1 & 1\\ 4 & -9\end{bmatrix},\quad \Lambda=\operatorname{diag}(1,-.3).$$

写 $r=(-.3)^k$，直接相乘给

$$M^k=\frac{1}{13}\begin{bmatrix}9+4r & 9-9r\\ 4-4r & 4+9r\end{bmatrix}\longrightarrow L=\frac{1}{13}\begin{bmatrix}9 & 9\\ 4 & 4\end{bmatrix}.$$

两列都是 $(9/13,4/13)$，分别表示从初值 e_1 或 e_2 演化得到同一个终态。更一般地，$Lu_0=(u_{0,1}+u_{0,2})(9/13,4/13)$。若初始两分量为非负且和为 1，它们可表示两个位置的概率，极限与起点无关。根为负数意味着偏差每一步反向，模 .3 使偏差逐渐变小。概率解释将在第 23 讲进一步展开，这里结论完全由已算出的幂公式得到。 ♦

21.5　把二阶递推改写成一阶系统

Fibonacci 数列从 $F_0=0,F_1=1$ 开始，满足 $F_{k+2}=F_{k+1}+F_k$。每个新值需要两个旧值，因此不能用一个标量作为完整状态；把两个相邻项装入一个向量，就能使下一状态仅依赖当前状态。这种增大状态维数的方法也适用于更高阶线性递推。

■ 例 21.8　**Fibonacci 的矩阵表示**　定义 $u_k=(F_{k+1},F_k)$，写出其状态矩阵，并用特征分解求通项。

解　第一行更新为 $F_{k+2}=F_{k+1}+F_k$，第二行复制原第一分量，因此

$$u_{k+1}=Au_k,\qquad A=\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix},\qquad u_0=(1,0).$$

特征方程为 $t^2-t-1=0$。记 $\phi=(1+\sqrt{5})/2$、$\psi=(1-\sqrt{5})/2$，则 $\phi+\psi=1$、$\phi\psi=-1$，且每个根满足 $t^2=t+1$。因此 $v_1=(\phi,1)$、$v_2=(\psi,1)$ 是对应特征向量：$A(t,1)=(t+1,t)=(t^2,t)=t(t,1)$。

由于 $\phi-\psi=\sqrt{5}\neq 0$，两个方向独立。设 $u_0=c_1v_1+c_2v_2$，第二分量给 $c_1+c_2=0$，第一分量给 $c_1\phi+c_2\psi=1$，解得 $c_1=1/\sqrt{5}$、$c_2=-1/\sqrt{5}$。所以

$$u_k=\frac{\phi^k}{\sqrt{5}}(\phi,1)-\frac{\psi^k}{\sqrt{5}}(\psi,1).$$

比较第二分量，得到 Binet 公式

$$\boxed{F_k=\frac{\phi^k-\psi^k}{\sqrt{5}}}.$$

比较第一分量同样给出 F_{k+1}，与状态定义相容。检验 $k=0$ 得 0，$k=1$ 得 1；两项各自满足递推，而初值又正确，因此公式给出整个数列。 ♦

直观理解　通项虽然含两个无理数幂，但两者的组合严格为整数，因为它是同一个整数递推的解。增长主要由 ϕ^k 决定；$|\psi|<1$ 使另一项逐步缩小。增长率由状态矩阵的根决定，前面的系数则由初始条件决定。这三个角色在差分方程、人口模型和概率更新中都会重复出现。

■ 例 21.9　**增长、取整与方向**　解释为什么 $F_k\sim\phi^k/\sqrt{5}$，为什么 $\phi^k/\sqrt{5}$ 的最近整数就是 F_k，以及 u_k 的方向趋向哪里。

解　从精确公式除以 $\phi^k/\sqrt{5}$，得到 $F_k/(\phi^k/\sqrt{5}) = 1-(\psi/\phi)^k \to 1$，故渐近等价成立。另一方面

$$\left|F_k - \frac{\phi^k}{\sqrt{5}}\right| = \frac{|\psi|^k}{\sqrt{5}} \le \frac{1}{\sqrt{5}} < \frac{1}{2} \qquad (k \ge 0).$$

因此最近整数唯一且等于 F_k，这是精确误差界，不能只由渐近等价推出。对方向而言，从状态式除以 ϕ^k 得 $u_k/\phi^k \to (\phi, 1)/\sqrt{5}$。也即 $F_{k+1}/F_k \to \phi$ $(k \ge 1)$；轨迹渐渐靠近直线 $y = x/\phi$。这里长度增长，但相对于主导项的方向误差衰减，两件事可以同时发生。 ♦

图 21.3 画出前几个状态及主导特征直线。$u_0 = (1, 0)$ 不在该直线上，它含有第二方向的分量；后者乘以负数 ψ，所以状态交替落在主导直线两侧，而偏离的绝对距离趋于零。

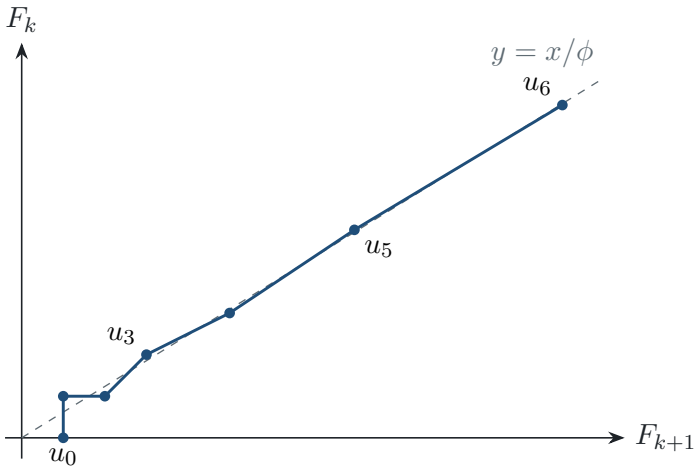


图 21.3　Fibonacci 状态点并非始终位于同一直线上；负的次要特征值使偏差交替变号，主导正特征值决定增长方向。

相邻项的比值在一开始就有定义吗　公式 $F_{k+1}/F_k \to \phi$ 是否允许直接从 $k = 0$ 开始列出比值？
答　不允许，因为 $F_0 = 0$。从 $k = 1$ 开始，分母均为正，才可以讨论相邻比值。矩阵状态 $u_0 = (1, 0)$ 则完全有定义。使用比值或逆矩阵时，都要检查分母非零；矩阵递推比单个比值表达式保留了更多初始时刻的信息。

同一个矩阵还可编码整个 Fibonacci 数列。对 $k \ge 1$，由直接归纳或特征分解可得

$$A^k = \begin{bmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{bmatrix}.$$

$k = 1$ 时等式就是 A；右乘一次 A，两列按 Fibonacci 加法更新，因此归纳成立。行列式也提供一个新恒等式：$\det(A^k) = (-1)^k$，所以

$$F_{k+1}F_{k-1} - F_k^2 = (-1)^k.$$

这是 Cassini 恒等式。它说明了一个递推关系如何通过矩阵乘法和行列式连接到另一个整式恒等式；证明不必展开 Binet 公式的根式。这里的条件 $k \ge 1$ 保证 F_{k-1} 已在数列定义范围内。

21.6　同一方法求解其他递推

如果递推变成 $a_{k+2}=p\,a_{k+1}+q\,a_k$，状态矩阵变为 $\begin{bmatrix}p & q\\ 1 & 0\end{bmatrix}$，特征方程为 $t^2-pt-q=0$。不同初值不会改变这两个根，却会改变对应方向的系数。不同参数可能造成根的绝对值小于、等于或大于 1，因此有衰减、保留、增长与交替等不同现象。下面的两个例子把算法完整执行一次，而不只写出候选通项。

■ **例 21.10**　**平均递推的非零极限**　设 $G_0=0, G_1=1$，$G_{k+2}=(G_{k+1}+G_k)/2$。求通项与极限。

解　状态矩阵为 $T=\begin{bmatrix}1/2 & 1/2\\ 1 & 0\end{bmatrix}$，其特征方程为 $t^2-t/2-1/2=(t-1)(t+1/2)$。对应向量可取 $v_1=(1,1)$、$v_2=(-1,2)$。初值 $(G_1,G_0)=(1,0)$ 满足 $c_1-c_2=1$、$c_1+2c_2=0$，故 $c_1=2/3$、$c_2=-1/3$。取状态第二分量得

$$G_k=\frac{2}{3}-\frac{2}{3}\left(-\frac{1}{2}\right)^k\longrightarrow\frac{2}{3}.$$

$G_0=0, G_1=1$ 均符合；$G_2=1/2, G_3=3/4$ 也与递推相符。极限不是初值的简单平均 $1/2$，因为该状态更新并非对两个时刻作对称的同时平均。实际上 $G_{k+1}+G_k/2=1$ 保持不变，若二者趋于同一个极限 ℓ，则 $3\ell/2=1$，再次得到 $2/3$。但这个不变量单独不能证明收敛，收敛仍来自次要根 $-1/2$ 的衰减。 ♦

■ **例 21.11**　**同一根，不同初值：Lucas 数列**　设 $L_0=2, L_1=1$，且 $L_{k+2}=L_{k+1}+L_k$。求通项，并比较它与 Fibonacci 数列的增长率。

解　矩阵仍是 $A=\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}$，特征根仍为 ϕ,ψ，方向仍为 $(\phi,1),(\psi,1)$。初值变成 $(L_1,L_0)=(1,2)$，系数方程是 $c_1+c_2=2$、$c_1\phi+c_2\psi=1$。由于 $\phi+\psi=1$，取 $c_1=c_2=1$ 正好满足，且基的独立性保证唯一。因此

$$L_k=\phi^k+\psi^k.$$

检验 $L_0=2$、$L_1=1$，并有 $L_2=3, L_3=4, L_4=7$。两数列都按 ϕ^k 的指数尺度增长，但前面系数不同：$L_k\sim\phi^k$，$F_k\sim\phi^k/\sqrt{5}$，因此 $L_k/F_k\to\sqrt{5}$。不能因为递推关系相同，就断定数列的通项或初值系数相同。相同矩阵确定可用的模式，初值决定这些模式分别出现多少。 ♦

■ **例 21.12**　**一个增长根与一个交替根**　设 $H_0=0, H_1=1$，$H_{k+2}=H_{k+1}+2H_k$。求通项，解释次要项为什么不会趋于零。

解　状态矩阵为 $T=\begin{bmatrix}1 & 2\\ 1 & 0\end{bmatrix}$，特征方程为 $t^2-t-2=(t-2)(t+1)$。对应向量取 $(2,1)$、$(-1,1)$。初值 $(1,0)=c_1(2,1)+c_2(-1,1)$，故 $2c_1-c_2=1$、$c_1+c_2=0$，解得 $c_1=1/3, c_2=-1/3$。比较第二分量得到

$$H_k=\frac{2^k-(-1)^k}{3}.$$

前几项为 $0,1,1,3,5,11$，逐项满足递推。主导项 $2^k/3$ 指数增长；余项 $-(-1)^k/3$ 的绝对值始终为 $1/3$，不趋于零。但它相对于 $2^k/3$ 的比例趋于零，所以仍有 $H_k \sim 2^k/3$。这再次区分绝对误差与相对误差：渐近等价只要求后者趋零。 ♦

21.7　检查对角化并处理特殊情形

对角化并不是要求某一个固定的矩阵 S，而是要求一组特征向量基。每个方向可乘非零常数，同一重根的特征空间中还可更换基；这些选择都改变 S，却不改变 A。这解释了为什么两份正确解答的 S 可以不同。检查时最方便的是 $AS = S\Lambda$：它只需核对各列，无须先做容易出错的两次矩阵乘法。但还必须检查 S 可逆，否则该等式单独不构成对角化。

■ **例 21.13**　**原课习题 22.1：找出全部对角化矩阵**　描述所有使 $S^{-1}BS$ 对角化的矩阵 S，其中 $B = \begin{bmatrix} 4 & 0 \\ 1 & 2 \end{bmatrix}$；再描述使 B^{-1} 对角化的矩阵。

解　由三角结构，根为 $4,2$。根 4 的零空间方程为 $x_1 - 2x_2 = 0$，所以 $E_4 = \mathrm{span}\{(2,1)\}$；根 2 的方程为 $2x_1 = 0$，所以 $E_2 = \mathrm{span}\{(0,1)\}$。每个特征空间一维，若要选择两个独立特征向量，必须各选一个。因此全部矩阵恰为

$$S = \begin{bmatrix} 2a & 0 \\ a & b \end{bmatrix} \quad \text{或} \quad S = \begin{bmatrix} 0 & 2a \\ b & a \end{bmatrix}, \qquad a, b \neq 0.$$

第一种对应 $\Lambda = \mathrm{diag}(4,2)$，第二种对应 $\mathrm{diag}(2,4)$。行列式分别为 $2ab, -2ab$，所以独立条件与参数条件完全吻合。注意第二方向是 $(0,1)$，代回 $B(0,1) = (0,2)$ 即可检查；若误写为 $(1,0)$，乘积 $(4,1)$ 并不平行。

B 可逆，且 B^{-1} 在这两个方向的特征值分别为 $1/4, 1/2$。因此以上同一批 S 都对角化 B^{-1}。反过来，若某个 S 对角化 B^{-1}，将等式取逆也得到 S 对角化 B，所以没有遗漏其他矩阵。这里既说明了「这些都可行」，也证明了「只有这些可行」，完成了求全部的要求。 ♦

本讲的计算可以浓缩为一条可复用流程：求特征值与零空间，挑出一组独立特征向量并排成 S，按列组成对应对角矩阵 Λ，先检查 $AS = S\Lambda$ 再使用 $A^k = S\Lambda^k S^{-1}$。有初始值时额外求 $c = S^{-1}u_0$，用 $\sum_j c_j \lambda_j^k v_j$ 描述每一时刻。若根重复，就不能跳过零空间维数检查；还须说明允许的标量范围：Fibonacci 矩阵在实数范围内可对角化，但所用的根为无理数，不能要求对角化的两个矩阵都只有有理数元素。

公式中的坐标不能省略　$A^k u_0 = S\Lambda^k S^{-1} u_0$。中间的 Λ^k 作用于特征基坐标，左右的 S^{-1} 与 S 分别负责换入和换回。若写成 $\Lambda^k S^{-1} u_0$，得到的是系数列而非原坐标中的状态；若 $u_0 = Sc$，则正确简写是 $S\Lambda^k c$。比较向量所在的坐标系，能迅速发现这类维数相同但含义不同的错误。

■ **例 21.14**　**不能对角化时仍可求幂**　对 $J = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$，不使用对角化求 J^k，并比较它与 $2^k I$。

解　令 $N = J - 2I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$，直接相乘得 $N^2 = 0$。由于 $2I$ 与 N 交换，二项式展开中含 N^2 及更高次的项全部为零，所以对 $k \geq 1$，

$$J^k = (2I + N)^k = 2^k I + k2^{k-1} N = \begin{bmatrix} 2^k & k2^{k-1} \\ 0 & 2^k \end{bmatrix}.$$

在 $k = 0$ 时单独取 $J^0 = I$。例如 $J^2 = \begin{bmatrix} 4 & 4 \\ 0 & 4 \end{bmatrix}$，可直接相乘检查。虽然唯一特征值为 2，矩阵幂还含多项式因子 k；只有对角化成功时，才能直接将矩阵幂写成固定基下的纯指数分量。这不影响上一节在可对角化条件下的论证，却解释了为何该条件需要明确保留。 ♦

实际计算时，对角化提供的收益也取决于目的。只求一次 Au 时，直接乘矩阵通常更快；要得到任意 k 的精确公式、多个初值的长期状态或参数化递推的增长率时，先求特征基才能反复复用。若某个特征值重复而特征向量不足，应在这一步停下对角化算法，不能把奇异的 S 当成有逆矩阵。上例通过幂零部分求幂，是特殊结构下的另一种计算方法；更系统的讨论会在相似矩阵与 Jordan 标准形一讲展开。

本讲小结

(1) $AS = S\Lambda$ 的第 j 列就是 $Av_j = \lambda_j v_j$；S 可逆等价于这些列构成一组基。对角化矩阵不唯一，列缩放、列换序需同步改变 Λ。

(2) 不同特征值的特征向量独立，因此 n 个不同根保证可对角化；重复根要实际求各特征空间，可能够用，也可能不够。

(3) 若 $A = S\Lambda S^{-1}$，则 $A^k = S\Lambda^k S^{-1}$。在可对角化条件下，$A^k \to 0$ 等价于所有特征值模小于 1。

(4) 差分系统满足 $u_k = A^k u_0 = \sum_j c_j \lambda_j^k v_j$，$c = S^{-1}u_0$。单条轨迹的极限只看该初值包含的方向。

(5) 高阶递推可用相邻项组成状态向量。Fibonacci 的两根为 ϕ, ψ，初值系数给出 $F_k = (\phi^k - \psi^k)/\sqrt{5}$；其他初值使用同一根和方向，但系数会变。

练习

21.1 ★　对 $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$，用 $v_1 = (1,1), v_2 = (1,-2)$ 对角化。写出 S, Λ, S^{-1} 并直接核对 $AS = S\Lambda$。

21.2 ★　对 $T = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ 求 T^k，其中 $k \geq 0$。请保留完整的 k 公式并代入 $k = 0, 1$ 检查。

21.3 ★　求 $N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ 的特征值与特征空间。它可对角化吗？N^2 的全部特征向量与 N 的是否相同？

21.4 ★ 对 $J=\begin{bmatrix}2&1\\0&2\end{bmatrix}$ 求特征多项式与特征空间，说明为什么不能以重根次数作为特征向量数。

21.5 ★★ 构造一个 2×2 矩阵，使特征值为 $2,5$ 且相应方向为 $(1,1),(1,2)$。求 A^k 并核对 $k=1$。

21.6 ★★ 比较 $C_1=\begin{bmatrix}.6&.9\\.1&.6\end{bmatrix}$、$C_2=\begin{bmatrix}.6&.9\\.4&.6\end{bmatrix}$、$C_3=\begin{bmatrix}.6&.4\\.4&.6\end{bmatrix}$ 的根，并说明哪一个矩阵幂趋于零、增长或收敛到非零矩阵。

21.7 ★★ 令 Lucas 数列 $L_0=2,L_1=1,L_{k+2}=L_{k+1}+L_k$。由 Fibonacci 状态矩阵推导 L_k 的通项，并求 L_5。

21.8 ★★ 对 $H_0=0,H_1=1,H_{k+2}=H_{k+1}+2H_k$，建立状态矩阵，求根、通项与渐近增长率。

21.9 ★★★ 设 A 可对角化，$A=S\Lambda S^{-1}$。证明 $p(A)=Sp(\Lambda)S^{-1}$，其中 p 为多项式；说明 $p(A)$ 何时可逆，并给出其逆的特征值。

21.10 ★★★ 给定 $A=\begin{bmatrix}.6&.9\\.4&.1\end{bmatrix}$，将其视为两状态概率更新矩阵。求极限 $\lim A^k u_0$，写出两列的意义，并检查总量守恒。

练习答案与提示

21.1 $Av_1=5v_1$、$Av_2=2v_2$，故 $S=\begin{bmatrix}1&1\\1&-2\end{bmatrix}$、$\Lambda=\operatorname{diag}(5,2)$、$S^{-1}=\frac{1}{3}\begin{bmatrix}2&1\\1&-1\end{bmatrix}$。$AS=\begin{bmatrix}5&2\\5&-4\end{bmatrix}=S\Lambda$。$\det S=-3\neq 0$，故 S 可逆并给出对角化。

21.2 T 的根为 $1,3$，方向可取 e_1,e_2，且 $Te_2=2e_1+3e_2$，所以仅这组不成特征基；解 $(T-\lambda I)x=0$ 得 $v_1=(1,0)$、$v_3=(1,1)$。令 $S=\begin{bmatrix}1&1\\0&1\end{bmatrix}$，$S^{-1}=\begin{bmatrix}1&-1\\0&1\end{bmatrix}$，故

$$T^k=S\operatorname{diag}(1,3^k)S^{-1}=\begin{bmatrix}1&3^k-1\\0&3^k\end{bmatrix}.$$

$k=0$ 为单位阵，$k=1$ 回到 T。把错误的标准基直接当作特征向量会漏掉非零上三角项；按零空间求向量是必要的。

21.3 根为 0（代数重数 2），零空间为 $\operatorname{span}\{(1,0)\}$，只有一条特征方向，故 N 不可对角化。但 $N^2=0$，所以 $N^2x=0x$ 对所有 x 成立；其全部非零向量都是特征向量，严格多于 N 的特征向量。平方可能合并不同根，也可能增加特征空间。

21.4 $p_J(t)=(t-2)^2$。$(J-2I)x=0$ 给 $x_2=0$，所以 $E_2=\operatorname{span}\{(1,0)\}$，维数只有 1。对角化需要两条独立特征方向；代数重数 2 并未提供第二条方向。

21.5 令 $S=\begin{bmatrix}1&1\\1&2\end{bmatrix}$，$\det S=1$，$S^{-1}=\begin{bmatrix}2&-1\\-1&1\end{bmatrix}$。则

$$A=S\operatorname{diag}(2,5)S^{-1}=\begin{bmatrix}-1&3\\-6&8\end{bmatrix},\quad A^k=S\operatorname{diag}(2^k,5^k)S^{-1}=\begin{bmatrix}2^{k+1}-5^k&-2^k+5^k\\2^{k+1}-2\cdot 5^k&-2^k+2\cdot 5^k\end{bmatrix}.$$

$k=1$ 得 A。该矩阵有负元素，虽然根是 $2,5$，它不代表概率转移矩阵；本题只要求线性变换的代数构造。

21.6 三者的根依次为 $.9,.3$；$1.2,0$；$1,.2$。C_1 可对角化且根模均小于 1，所以幂趋零。C_2 含根 1.2，沿该方向增长，矩阵幂不收敛。C_3 根为 $1,.2$，前者保留平均方向，后者衰减，所以 $C_3^k \to \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$。答案不能只用根的大小下结论，还应确认可对角化；这三者的根不同，确实有独立特征向量基。

21.7 状态矩阵特征方向为 $(\phi,1),(\psi,1)$。初值 $(L_1,L_0)=(1,2)$ 的分解系数均为 1，于是 $L_k=\phi^k+\psi^k$。由递推逐算为 $L_2=3,L_3=4,L_4=7,L_5=11$。次要根满足 $|\psi|<1$，所以 ψ^k 的绝对值趋于零；与主项 ϕ^k 相比，它的相对贡献也趋零。

21.8 状态矩阵 $T=\begin{bmatrix}1 & 2\\ 1 & 0\end{bmatrix}$，特征方程 $(t-2)(t+1)=0$。对应方向 $(2,1),(-1,1)$；初值 $(1,0)=\frac{1}{3}(2,1)-\frac{1}{3}(-1,1)$。故 $H_k=(2^k-(-1)^k)/3\sim 2^k/3$。前几项 $0,1,1,3,5$ 可回代。次要根模等于 1，交替项不绝对衰减，但相对主项消失。

21.9 若 $p(t)=\sum_j a_j t^j$，则 $p(A)=\sum_j a_j S\Lambda^j S^{-1}=S(\sum_j a_j\Lambda^j)S^{-1}=Sp(\Lambda)S^{-1}$。其对角项为 $p(\lambda_i)$，因此可逆当且仅当所有这些值都非零；逆为 $S\,\mathrm{diag}(1/p(\lambda_i))S^{-1}$，相应特征值为 $1/p(\lambda_i)$，重复值仍按各特征方向计数。注意本题关于特征值列举依赖 A 可对角化。

21.10 设 $u_0=(a,b)$，由原课公式得

$$A^k u_0 \to (a+b)(9/13,4/13).$$

极限矩阵两列均为稳态方向 $(9/13,4/13)$，分别是 $u_0=e_1,e_2$ 时的结果。总量行向量 $(1,1)$ 满足 $(1,1)A=(1,1)$，故每步分量和不变；极限两分量之和为 $(a+b)(9+4)/13=a+b$。若初值是概率向量，则极限仍为概率向量。负根 $-.3$ 使偏差交替变号并衰减。

第 22 讲

微分方程与 e^{At}

Differential Equations and exp(At)

对应材料：MIT 18.06SC Session 2.10（对应 Strang 第 23 讲视频 Differential Equations and exp(At)）课时摘要；Session 2.10 习题 23.1、23.2 及解答；Strang《Introduction to Linear Algebra》第 4 版 §6.3。

本讲目标

(1) 会用纯指数解 $e^{\lambda t}x$ 解常系数线性方程组 $\dfrac{du}{dt}=Au$：求特征值和特征向量，把 $u(0)$ 拆成特征向量的组合，每个分量乘上自己的 $e^{\lambda_i t}$。

(2) 理解矩阵指数 e^{At} 的级数定义和基本性质；会用 $e^{At}=Se^{\Lambda t}S^{-1}$ 计算它，并知道 A 不能对角化时 e^{At} 照样存在。

(3) 会从特征值判断解的长期行为：衰减到 0、趋于稳态还是爆炸；对 2×2 矩阵会用迹和行列式判断稳定性。

(4) 会把 $y''+by'+ky=0$ 这类高阶方程化成一阶方程组，并说明它的特征多项式为什么就是原方程的特征方程。

(5) 知道反对称矩阵（skew-symmetric matrix）$A^{\mathsf{T}}=-A$ 给出的运动保持长度不变，此时 e^{At} 是正交矩阵。

注 22.1 [阅读准备] 本讲用到一元导数、指数函数求导，以及幂级数的收敛和逐项求导。尚未学过微积分时，可暂缓本讲及对应练习，先读第 23 讲的 Markov 矩阵；后续谱定理、有限维 Fourier 变换与 SVD 的证明不依赖本讲的微分方程结论。补学这些知识后，再回到本讲研究矩阵指数与连续时间的应用。

第 21 讲用对角化 $A=S\Lambda S^{-1}$ 解了差分方程 $u_{k+1}=Au_k$：每一步乘一次 A，特征向量方向上的分量各自乘 λ_i。本讲把「每一步」换成「每一瞬间」，研究连续时间的方程 $\dfrac{du}{dt}=Au$。幂 λ^k 的角色将由指数 $e^{\lambda t}$ 接替，而对角化的思路完全不变。

22.1 从 $du/dt=au$ 到 $du/dt=Au$

22.1.1 一个未知函数的情形

先回忆只有一个未知函数的情形。方程

$$\frac{du}{dt}=au \qquad (a \text{ 是常数})$$

说的是：u 的变化率与 u 本身成正比。它的解是 $u(t)=e^{at}u(0)$，代入即可验证：$\frac{d}{dt}\big(e^{at}u(0)\big)=ae^{at}u(0)=au$。

你可能会问：还有没有别的解？设 $u(t)$ 是任意一个解，考察 $e^{-at}u(t)$ 的导数：

$$\frac{d}{dt}\big(e^{-at}u(t)\big)=-ae^{-at}u+e^{-at}u'=e^{-at}\left(u'-au\right)=0.$$

导数恒为 0，所以 $e^{-at}u(t)$ 是常数，等于它在 $t=0$ 的值 $u(0)$。于是 $u(t)=e^{at}u(0)$，解是唯一的。§22.3 证明矩阵情形的唯一性时，用的是同一个技巧。

常数 a 决定了解的长期行为：$a>0$ 时指数增长，$a<0$ 时指数衰减到 0，$a=0$ 时保持不动。若 $a=\alpha+i\beta$ 是复数，由 Euler 公式

$$e^{at}=e^{\alpha t}e^{i\beta t}=e^{\alpha t}(\cos\beta t+i\sin\beta t),\qquad |e^{at}|=e^{\alpha t}.$$

模只由实部 α 决定，虚部 β 只让 e^{at} 在复平面上绕圈。§22.4 判断稳定性时，看的正是特征值的实部。

22.1.2 相互耦合的两个量

现在设有两个容器，里面某种物质的量分别是 $u_1(t)$ 和 $u_2(t)$。物质在两个容器之间流动：容器 1 以速率 u_1 流向容器 2，容器 2 以速率 $2u_2$ 流向容器 1。于是

$$\frac{du_1}{dt}=-u_1+2u_2,\qquad \frac{du_2}{dt}=u_1-2u_2.$$

写成矩阵形式：

$$\frac{du}{dt}=Au,\qquad A=\begin{bmatrix}-1 & 2\\ 1 & -2\end{bmatrix},\qquad u(0)=\begin{bmatrix}1\\0\end{bmatrix}.$$

初始时物质全在容器 1 里。难点在于两个方程**耦合**在一起：u_1 的变化率依赖 u_2，u_2 的变化率又依赖 u_1，哪一个都不能单独解。从方程的样子可以猜到 u_1 会减少、u_2 会增加，但要知道减到多少、多快，需要线性代数。

22.1.3 纯指数解

一维时解是 e^{at} 乘一个常数。仿照它，我们猜一个**方向固定、只有大小按指数变化**的解：

$$u(t)=e^{\lambda t}x,\qquad x\text{ 是与}t\text{ 无关的非零向量.}$$

代入方程，左边 $\dfrac{du}{dt}=\lambda e^{\lambda t}x$，右边 $Au=e^{\lambda t}Ax$。消去永不为零的因子 $e^{\lambda t}$，得到

$$Ax=\lambda x.$$

这正是第 20 讲的特征值方程。

> **定理 22.1（纯指数解）**　设 A 是实方阵，$\lambda\in\mathbb{R}$、$x\in\mathbb{R}^n\setminus\{0\}$。$u(t)=e^{\lambda t}x$ 是 $\dfrac{du}{dt}=Au$ 的解，当且仅当 $Ax=\lambda x$，即 λ 是 A 的特征值、x 是对应的特征向量。

这样的解叫**纯指数解**（pure exponential solution）。它在 $\mathbb{R}^n$ 中沿着特征向量 x 所在的直线运动：$\lambda>0$ 时离原点越来越远，$\lambda<0$ 时向原点收缩，$\lambda=0$ 时一动不动。第 20 讲说「Ax 与 x

同向」，在这里变成了「速度 $\frac{du}{dt} = Au$ 与位置 u 同向」，所以轨迹是直线。非实特征值需要把共轭特征方向合成实解，后面的螺旋例子会展示这种情形。

方程是**线性**的：若 u 和 w 都是解，则 $\frac{d}{dt}(c_1u + c_2w) = c_1Au + c_2Aw = A(c_1u + c_2w)$，所以 $c_1u + c_2w$ 也是解。把各个纯指数解叠加起来，就得到一般解。

定理 22.2（一般解）　设实矩阵 A 有一组由 n 个实特征向量 $x_1, \dots, x_n$ 构成的基，对应实特征值 $\lambda_1, \dots, \lambda_n$。则

$$u(t) = c_1e^{\lambda_1 t}x_1 + c_2e^{\lambda_2 t}x_2 + \cdots + c_ne^{\lambda_n t}x_n$$

是 $\dfrac{du}{dt} = Au$ 的解，其中系数由初值决定：$u(0) = c_1x_1 + \cdots + c_nx_n$，即 $Sc = u(0)$，$S = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$。

因为 $x_1, \dots, x_n$ 线性无关，S 可逆，$c = S^{-1}u(0)$ 总能解出来，所以任何实初值都能匹配。定理 22.5 会证明这个解是唯一的，因此它就是全部答案。

每个特征向量各走各的路　解 $\dfrac{du}{dt} = Au$ 分三步：

(1) 把初值拆开：$u(0) = c_1x_1 + \cdots + c_nx_n$；

(2) 每一部分各自演化：c_ix_i 变成 $c_ie^{\lambda_i t}x_i$；

(3) 再加起来：$u(t) = \sum_i c_ie^{\lambda_i t}x_i$。

矩阵 A 把各分量耦合在一起；换到特征向量的坐标里，耦合就消失了，n 维问题变成 n 个一维问题。

这与第 21 讲解差分方程的步骤一模一样，只是把 λ_i^k 换成了 $e^{\lambda_i t}$。表 22.1 把两者并排列出。

	差分方程（第 21 讲）	微分方程（本讲）
方程	$u_{k+1} = Au_k$	$\dfrac{du}{dt} = Au$
纯解	$\lambda^k x$	$e^{\lambda t}x$
一般解	$\sum c_i\lambda_i^k x_i$	$\sum c_ie^{\lambda_i t}x_i$
矩阵形式	$A^k = S\Lambda^k S^{-1}$	$e^{At} = Se^{\Lambda t}S^{-1}$
衰减到 0 的条件	所有 $\lvert\lambda_i\rvert < 1$	所有 $\operatorname{Re}\lambda_i < 0$
稳态来自	$\lambda = 1$	$\lambda = 0$

表 22.1　离散时间与连续时间的对照。系数 c 都由 $Sc = u(0)$ 决定。

22.2　用特征向量求解：几个完整的例子

■ **例 22.1　两个容器之间的流动**　解本讲开头的方程组 $\dfrac{du}{dt}=Au$，$A=\begin{bmatrix}-1 & 2\\ 1 & -2\end{bmatrix}$，$u(0)=\begin{bmatrix}1\\0\end{bmatrix}$，并求 $t\to\infty$ 时的极限。

解　第一步：特征值。A 的第二列是第一列的 -2 倍，A 奇异，所以 $\lambda_1=0$ 是特征值。两个特征值之和等于迹 $-1+(-2)=-3$，所以 $\lambda_2=-3$。用特征多项式核对：

$$\det(A-\lambda I)=(-1-\lambda)(-2-\lambda)-2\cdot 1=\lambda^2+3\lambda=\lambda(\lambda+3).$$

第二步：特征向量。$\lambda_1=0$：解 $Ax=0$，第一行 $-x_1+2x_2=0$，取 $x_1=\begin{bmatrix}2\\1\end{bmatrix}$。$\lambda_2=-3$：解 $(A+3I)x=0$，

$$A+3I=\begin{bmatrix}2 & 2\\ 1 & 1\end{bmatrix},\qquad x_2=\begin{bmatrix}1\\-1\end{bmatrix}.$$

核对：$Ax_2=\begin{bmatrix}-1-2\\1+2\end{bmatrix}=\begin{bmatrix}-3\\3\end{bmatrix}=-3x_2$。

第三步：一般解。由定理 22.2，注意 $e^{0t}=1$：

$$u(t)=c_1e^{0t}x_1+c_2e^{-3t}x_2=c_1\begin{bmatrix}2\\1\end{bmatrix}+c_2e^{-3t}\begin{bmatrix}1\\-1\end{bmatrix}.$$

第四步：匹配初值。令 $t=0$，解 $Sc=u(0)$：

$$\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}c_1\\c_2\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}.$$

第二行给出 $c_1=c_2$，代入第一行得 $3c_1=1$，所以 $c_1=c_2=\frac{1}{3}$。于是

$$u(t)=\frac{1}{3}\begin{bmatrix}2\\1\end{bmatrix}+\frac{1}{3}e^{-3t}\begin{bmatrix}1\\-1\end{bmatrix},\qquad \text{即}\quad u_1=\frac{2}{3}+\frac{1}{3}e^{-3t},\quad u_2=\frac{1}{3}-\frac{1}{3}e^{-3t}.$$

第五步：验证。$u_1'=-e^{-3t}$，而 $-u_1+2u_2=-\frac{2}{3}-\frac{1}{3}e^{-3t}+\frac{2}{3}-\frac{2}{3}e^{-3t}=-e^{-3t}$，相符；$u_2'=e^{-3t}=-u_1'$，而 $u_1-2u_2=-(-u_1+2u_2)=e^{-3t}$，也相符。$t=0$ 时 $u=(1,0)$，初值也对。

第六步：极限。$t\to\infty$ 时 $e^{-3t}\to 0$，所以

$$u(\infty)=\frac{1}{3}x_1=\begin{bmatrix}2/3\\1/3\end{bmatrix}.$$

物质从容器 1 流向容器 2，最终容器 1 留下 $\frac{2}{3}$，容器 2 得到 $\frac{1}{3}$。 ♦

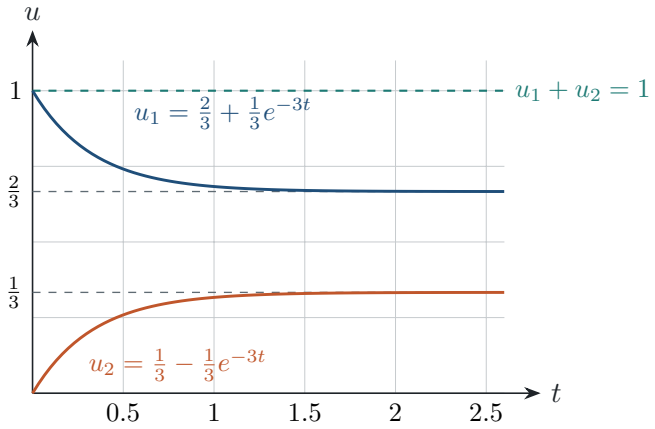


图 **22.1**　例 22.1 的解。$\lambda = -3$ 的分量迅速消失，$\lambda = 0$ 的分量 $\frac{1}{3}x_1 = (\frac{2}{3}, \frac{1}{3})$ 留下来成为稳态；总量 $u_1 + u_2$ 始终为 1。

这个例子展示了两类特征值的分工。$\lambda_1 = 0$ 的特征向量 $x_1 = (2, 1)$ 满足 $Ax_1 = 0$，代表稳态（steady state）：$u_1 = 2u_2$ 时，从容器 1 流出的速率 u_1 恰好等于流入的速率 $2u_2$，两边平衡。$\lambda_2 = -3$ 的特征向量代表偏离稳态的那部分，它以 e^{-3t} 的速度衰减。

想一想　图 22.1 中 $u_1 + u_2$ 始终等于 1。不解方程，能从矩阵 A 直接看出这一点吗？

答　能。$\frac{d}{dt}(u_1 + u_2) = (-u_1 + 2u_2) + (u_1 - 2u_2) = 0$。用矩阵的语言说：$A$ 每一列的元素之和都是 0，即 $\begin{bmatrix}1 & 1\end{bmatrix} A = \begin{bmatrix}0 & 0\end{bmatrix}$，所以 $\frac{d}{dt}\left(\begin{bmatrix}1 & 1\end{bmatrix} u\right) = \begin{bmatrix}1 & 1\end{bmatrix} Au = 0$。物质只在容器之间流动，不会凭空产生或消失。这也说明 A^{T} 有特征向量 $(1, 1)$、特征值 0，从而 A 奇异。

■ **例 22.2**　**一个增长、一个衰减**　解 $\dfrac{du}{dt} = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} u$，$u(0) = \begin{bmatrix}4\\ 2\end{bmatrix}$。即 $u_1' = u_2$，$u_2' = u_1$。

解　$\det(A - \lambda I) = \lambda^2 - 1$，特征值 $\lambda_1 = 1$，$\lambda_2 = -1$。A 交换两个分量，所以 $x_1 = (1, 1)$ 被保持（$Ax_1 = x_1$），$x_2 = (1, -1)$ 被反号（$Ax_2 = -x_2$）。

拆初值：$(4, 2) = c_1(1, 1) + c_2(1, -1)$，相加相减得 $c_1 = \frac{4+2}{2} = 3$，$c_2 = \frac{4-2}{2} = 1$。所以

$$u(t) = 3e^t \begin{bmatrix}1\\ 1\end{bmatrix} + e^{-t} \begin{bmatrix}1\\ -1\end{bmatrix} = \begin{bmatrix}3e^t + e^{-t}\\ 3e^t - e^{-t}\end{bmatrix}.$$

验证：$u_1' = 3e^t - e^{-t} = u_2$，$u_2' = 3e^t + e^{-t} = u_1$；$t = 0$ 时 $u = (4, 2)$。

长期行为：e^{-t} 那一项消失，e^t 那一项无限增长。t 很大时 $u(t) \approx 3e^t(1, 1)$，即解爆炸，并且方向越来越接近特征向量 x_1，$u_1/u_2 \to 1$。 ♦

例 22.2 说明：只要有一个特征值是正数，而初值在这个特征向量上的分量 $c_i \neq 0$，解就会爆炸，并且最终由增长最快的那个方向主导。这与第 21 讲中 $A^k u_0$ 被最大的 $|\lambda|$ 主导是同一个现象。

■ **例 22.3**　**三阶三角矩阵**　解 $\dfrac{du}{dt} = Au$，其中 $A = \begin{bmatrix}1 & 1 & 1\\ 0 & 2 & 1\\ 0 & 0 & 3\end{bmatrix}$，$u(0) = \begin{bmatrix}9\\ 7\\ 4\end{bmatrix}$。

解　三角矩阵的特征值就是对角元：$\lambda = 1, 2, 3$。

特征向量。$\lambda = 1$：$A - I = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$，由第三行 $x_3 = 0$，再由第一行 $x_2 = 0$，x_1 自由，取 $x_1 = (1, 0, 0)$。

$\lambda = 2$：$A - 2I = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$，第二行给出 $x_3 = 0$，第一行给出 $x_1 = x_2$，取 $(1, 1, 0)$。

$\lambda = 3$：$A - 3I = \begin{bmatrix} -2 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$，第二行给出 $x_2 = x_3$，第一行给出 $2x_1 = x_2 + x_3$，取 $(1, 1, 1)$。

核对：$A(1, 1, 0) = (2, 2, 0)$，$A(1, 1, 1) = (3, 3, 3)$。

拆初值。解 $Sc = u(0)$，S 是上三角矩阵，回代即可：

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ 4 \end{bmatrix} \implies c_3 = 4, \quad c_2 = 7 - 4 = 3, \quad c_1 = 9 - 3 - 4 = 2.$$

解。

$$u(t) = 2e^t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 3e^{2t} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 4e^{3t} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2e^t + 3e^{2t} + 4e^{3t} \\ 3e^{2t} + 4e^{3t} \\ 4e^{3t} \end{bmatrix}.$$

验证第一个方程 $u_1' = u_1 + u_2 + u_3$：左边 $2e^t + 6e^{2t} + 12e^{3t}$；右边 $(2e^t + 3e^{2t} + 4e^{3t}) + (3e^{2t} + 4e^{3t}) + 4e^{3t} = 2e^t + 6e^{2t} + 12e^{3t}$，相等。另外两个方程同样验证：$u_2' = 6e^{2t} + 12e^{3t} = 2u_2 + u_3$，$u_3' = 12e^{3t} = 3u_3$。 ♦

注 22.2　三角形的方程组也可以不用特征向量，从下往上逐个解：先由 $u_3' = 3u_3$ 得 $u_3 = 4e^{3t}$，再把它代入 $u_2' = 2u_2 + u_3$，解一个一维的非齐次方程，依此类推。特征向量方法的好处是对任何可对角化矩阵都适用，不要求三角形。

22.3　解耦与矩阵指数 e^{At}

22.3.1　换到特征向量坐标：解耦

上一节的三步可以写成矩阵运算。令 $u = Sv$，即用特征向量的坐标 v 来描述 u。代入 $\dfrac{du}{dt} = Au$：

$$S\frac{dv}{dt} = ASv \quad \implies \quad \frac{dv}{dt} = S^{-1}AS\,v = \Lambda v.$$

Λ 是对角矩阵，所以这是 n 个互不相干的一维方程 $\dfrac{dv_i}{dt} = \lambda_i v_i$，解为 $v_i(t) = e^{\lambda_i t} v_i(0)$。写成矩阵：

$$v(t) = e^{\Lambda t} v(0), \qquad e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix}.$$

再变回原来的坐标，$v(0)=S^{-1}u(0)$，$u(t)=Sv(t)$：

$$\boxed{u(t)=S\,e^{\Lambda t}S^{-1}u(0).}$$

这里 $v(0)=S^{-1}u(0)$ 正是定理 22.2 中的系数向量 c。所以这个公式只是把「拆开、各自演化、加起来」写成了三个矩阵的乘积（图 22.2）。

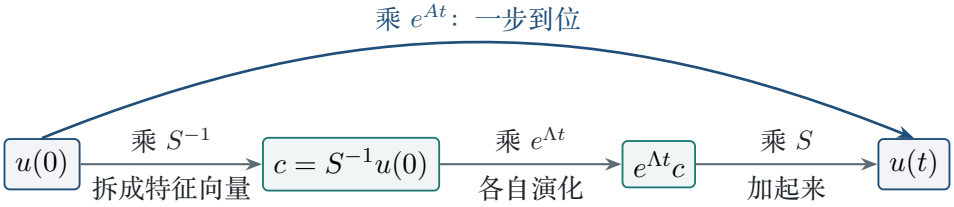


图 22.2　$e^{At}=Se^{\Lambda t}S^{-1}$ 的含义。下面一行绕道特征向量坐标，上面的弧线是直接作用的矩阵指数。

图中下方的路径把初值先变成特征坐标，再逐分量乘上指数因子，最后回到原坐标；上方弧线表示同一运算可合并为一次矩阵乘法。这个分解只要求 A 可对角化，矩阵指数的级数定义则不需要该条件。

22.3.2　矩阵指数的定义

上面的推导需要 n 个线性无关的特征向量。为了不依赖这一点，我们直接给「e 的矩阵次方」下定义。对实数，$e^{x}=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots$。把 x 换成方阵 At：

定义 22.3（矩阵指数）　对 $n\times n$ 矩阵 A 和实数 t，定义

$$e^{At}=\sum_{k=0}^{\infty}\frac{(At)^k}{k!}=I+At+\frac{(At)^2}{2!}+\frac{(At)^3}{3!}+\cdots.$$

这个级数对任何 A、任何 t 都收敛。理由如下：设 A 的每个元素的绝对值都不超过 m。对 $k\ge 1$ 归纳可得 A^k 的每个元素的绝对值不超过 $n^{k-1}m^k$（$A^{k+1}=A^kA$ 的一个元素是 n 个乘积之和，每个乘积不超过 $n^{k-1}m^k\cdot m$）。于是级数每个元素的各项绝对值不超过 $\frac{(nm|t|)^k}{k!}$，而 $\sum_k\frac{(nm|t|)^k}{k!}=e^{nm|t|}$ 收敛。

定理 22.4（矩阵指数的性质）　对任意方阵 A 和实数 s,t：

(1) $e^{A\cdot 0}=I$；

(2) $\dfrac{d}{dt}e^{At}=Ae^{At}=e^{At}A$；

(3) $e^{As}e^{At}=e^{A(s+t)}$；

(4) e^{At} 总是可逆的，$(e^{At})^{-1}=e^{-At}$；

(5) 若 $Ax=\lambda x$，则 $e^{At}x=e^{\lambda t}x$。

证明　(1) $t=0$ 时级数只剩第一项 I。

(2) 逐项求导（幂级数在收敛范围内可以逐项求导）：

$$\frac{d}{dt}\sum_{k=0}^{\infty}\frac{t^k}{k!}A^k=\sum_{k=1}^{\infty}\frac{t^{k-1}}{(k-1)!}A^k=A\sum_{j=0}^{\infty}\frac{t^j}{j!}A^j=Ae^{At}.$$

把 A 从右边提出来也可以，因为 A 与它自己的幂可交换，得到 $e^{At}A$。

(3) 两个级数相乘，把次数为 k 的项归在一起：

$$\begin{aligned}e^{As}e^{At}&=\sum_{k=0}^{\infty}\sum_{j=0}^{k}\frac{s^jA^j}{j!}\cdot\frac{t^{k-j}A^{k-j}}{(k-j)!}\\&=\sum_{k=0}^{\infty}\frac{A^k}{k!}\sum_{j=0}^{k}\binom{k}{j}s^jt^{k-j}=\sum_{k=0}^{\infty}\frac{A^k(s+t)^k}{k!}.\end{aligned}$$

最后一步是二项式定理。

(4) 在 (3) 中取 $s=-t$：$e^{-At}e^{At}=e^{A\cdot 0}=I$。

(5) $A^kx=\lambda^kx$，所以 $e^{At}x=\sum_k\frac{t^k}{k!}A^kx=\sum_k\frac{(\lambda t)^k}{k!}x=e^{\lambda t}x$。 ■

性质 (2) 和 (1) 合起来说：$U(t)=e^{At}$ 满足 $\dfrac{dU}{dt}=AU$，$U(0)=I$。它的每一列都是方程的一个解，第 j 列从初值 e_j（第 j 个坐标向量）出发。所以 e^{At} 也叫方程的基本解矩阵（fundamental matrix）。

定理 22.5（存在唯一性）　对任意方阵 A 和初值 $u(0)$，方程 $\dfrac{du}{dt}=Au$ 有且只有一个解：

$$\boxed{u(t)=e^{At}u(0).}$$

证明　*是解。*由性质 (2)，$\frac{d}{dt}\big(e^{At}u(0)\big)=Ae^{At}u(0)=Au$；由性质 (1)，$t=0$ 时它等于 $u(0)$。

*只有这一个。*设 $w(t)$ 也是解，$w(0)=u(0)$。考察 $e^{-At}w(t)$。由乘积求导法则和性质 (2)（把 A 换成 $-A$）：

$$\frac{d}{dt}\big(e^{-At}w\big)=-Ae^{-At}w+e^{-At}w'=-e^{-At}Aw+e^{-At}Aw=0.$$

这里用了 $Ae^{-At}=e^{-At}A$。所以 $e^{-At}w(t)$ 是常向量，等于 $t=0$ 时的值 $w(0)$。两边左乘 e^{At}，由性质 (4) 得 $w(t)=e^{At}w(0)=u(t)$。 ■

证明的关键一步和一维时完全一样：乘上 e^{-At}，导数变成 0。能这样做，靠的是 A 与 e^{-At} 可交换。

22.3.3 用对角化计算 e^{At}

无穷级数不适合直接手算。A 可对角化时有更好的办法。

定理 22.6　若 $A=S\Lambda S^{-1}$，则 $e^{At}=Se^{\Lambda t}S^{-1}$，其中 $e^{\Lambda t}=\operatorname{diag}(e^{\lambda_1 t},\dots,e^{\lambda_n t})$。

证明　由第 21 讲，$A^k = S\Lambda^k S^{-1}$。代入级数，每一项的 S 和 S^{-1} 都可以提到外面：

$$e^{At} = \sum_{k=0}^{\infty} \frac{t^k}{k!} S\Lambda^k S^{-1} = S\Big(\sum_{k=0}^{\infty} \frac{(\Lambda t)^k}{k!}\Big) S^{-1} = S e^{\Lambda t} S^{-1}.$$

对角矩阵的幂就是对角元的幂，所以中间括号里第 i 个对角元是 $\sum_k \frac{(\lambda_i t)^k}{k!} = e^{\lambda_i t}$，其余元素都是 0。 ■

于是 $e^{At}u(0) = Se^{\Lambda t}S^{-1}u(0)$，与 §22.3 开头解耦得到的公式完全一致。两条路殊途同归。

■ **例 22.4　改编自习题 23.2**　把 $A = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix}$ 写成 $S\Lambda S^{-1}$，由此求 e^{At}；再检验 e^{At} 在 $t = 0$ 时的值和导数。

解　A 是上三角矩阵，特征值 $\lambda_1 = 1$，$\lambda_2 = 3$。$\lambda_1 = 1$：$A - I = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$，取 $x_1 = (1, 0)$。$\lambda_2 = 3$：$A - 3I = \begin{bmatrix} -2 & 1 \\ 0 & 0 \end{bmatrix}$，取 $x_2 = (1, 2)$。于是

$$S = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, \qquad S^{-1} = \frac{1}{2}\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}.$$

核对 $S\Lambda S^{-1}$：$S\Lambda = \begin{bmatrix} 1 & 3 \\ 0 & 6 \end{bmatrix}$，再乘 S^{-1} 得 $\begin{bmatrix} 1 & -\frac{1}{2} + \frac{3}{2} \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} = A$。

计算 e^{At}：

$$Se^{\Lambda t} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} e^t & 0 \\ 0 & e^{3t} \end{bmatrix} = \begin{bmatrix} e^t & e^{3t} \\ 0 & 2e^{3t} \end{bmatrix},$$

$$e^{At} = \begin{bmatrix} e^t & e^{3t} \\ 0 & 2e^{3t} \end{bmatrix}\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} e^t & \frac{1}{2}e^{3t} - \frac{1}{2}e^t \\ 0 & e^{3t} \end{bmatrix}.$$

检验：$t = 0$ 时右上角是 $\frac{1}{2} - \frac{1}{2} = 0$，所以 $e^{A\cdot 0} = I$。求导：

$$\frac{d}{dt}e^{At} = \begin{bmatrix} e^t & \frac{3}{2}e^{3t} - \frac{1}{2}e^t \\ 0 & 3e^{3t} \end{bmatrix}, \qquad \frac{d}{dt}e^{At}\Big|_{t=0} = \begin{bmatrix} 1 & \frac{3}{2} - \frac{1}{2} \\ 0 & 3 \end{bmatrix} = A.$$

这与性质 (2) 在 $t = 0$ 时的结论 $Ae^{A\cdot 0} = A$ 一致。还可以用级数核对右上角：A^k 的右上角是 $\frac{3^k - 1}{2}$（对 k 归纳），所以右上角 $= \sum_k \frac{t^k}{k!} \cdot \frac{3^k - 1}{2} = \frac{1}{2}(e^{3t} - e^t)$。 ♦

e^{At} 不是逐个元素取指数　例 22.4 中 A 的右上角是 1，而 e^{At} 的右上角是 $\frac{1}{2}(e^{3t} - e^t)$，不是 e^t；左下角 0 对应的是 0，不是 $e^0 = 1$。只有对角矩阵才能逐个元素取指数（而且只对对角元取，非对角的 0 仍是 0）。

有时级数本身反而更快，尤其是 A 的幂有简单规律的时候。

■ **例 22.5　旋转矩阵**　设 $A=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$。直接用级数求 e^{At}，并解 $u'=Au$，$u(0)=(1,0)$。

解　先算幂：$A^2=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}=-I$，于是 $A^3=-A$，$A^4=I$，之后以 4 为周期循环。把级数按偶次项和奇次项分开：

$$e^{At}=\Big(1-\frac{t^2}{2!}+\frac{t^4}{4!}-\cdots\Big)I+\Big(t-\frac{t^3}{3!}+\frac{t^5}{5!}-\cdots\Big)A=(\cos t)I+(\sin t)A=\begin{bmatrix}\cos t & -\sin t\\ \sin t & \cos t\end{bmatrix}.$$

这正是转角为 t 的旋转矩阵。于是 $u(t)=e^{At}(1,0)=(\cos t,\sin t)$，点以单位角速度沿单位圆转圈，永远不衰减也不增长。

用特征值的语言：$\det(A-\lambda I)=\lambda^2+1$，$\lambda=\pm i$。$|e^{\pm it}|=1$，所以解既不衰减也不增长。实矩阵的复特征值成对出现，两个复指数 e^{it}、e^{-it} 组合起来，就得到实的 $\cos t$ 与 $\sin t$。 ♦

■ **例 22.6　不能对角化的情形**　设 $A=\begin{bmatrix}3 & 1\\ 0 & 3\end{bmatrix}=3I+N$，其中 $N=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$。求 e^{At}。

解　A 的特征值是二重的 3，但 $A-3I=N$ 的零空间只有一维（由 $(1,0)$ 张成），找不到两个无关的特征向量，定理 22.6 用不上。改用级数。

关键是 $N^2=0$，而 $3I$ 与 N 可交换，所以二项式定理成立：

$$A^k=(3I+N)^k=3^kI+k\,3^{k-1}N\quad(\text{含}N^2\text{及更高次的项都是}0).$$

代入级数，$k\ge 1$ 时 $\frac{k}{k!}=\frac{1}{(k-1)!}$：

$$e^{At}=\sum_{k=0}^{\infty}\frac{(3t)^k}{k!}I+\sum_{k=1}^{\infty}\frac{3^{k-1}t^k}{(k-1)!}N=e^{3t}I+te^{3t}N=e^{3t}\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}.$$

解里出现了 te^{3t} 这种纯指数之外的函数，这正是缺少特征向量的标志。一般的不可对角化矩阵要用若尔当标准形处理，留到第 27 讲。 ♦

e^{A+B} 一般不等于 e^Ae^B　数的指数满足 $e^{a+b}=e^ae^b$，矩阵却未必。取 $A=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$，$B=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$。$e^A=\mathrm{diag}(e,1)$；$B^2=0$ 给出 $e^B=I+B$；$(A+B)^2=A+B$ 给出 $e^{A+B}=I+(e-1)(A+B)$。于是

$$e^Ae^B=\begin{bmatrix}e & e\\ 0 & 1\end{bmatrix},\qquad e^Be^A=\begin{bmatrix}e & 1\\ 0 & 1\end{bmatrix},\qquad e^{A+B}=\begin{bmatrix}e & e-1\\ 0 & 1\end{bmatrix},$$

三者两两不同。原因是 $AB=B$ 而 $BA=0$，两者不可交换。$AB=BA$ 时等式成立（练习 22.6）；性质 $e^{As}e^{At}=e^{A(s+t)}$ 成立，正因为 As 与 At 可交换。

想一想　A 可以是奇异矩阵（例 22.1 中的 A 就是）。那么 e^{At} 会不会也是奇异的？

答　不会。由性质 (4)，e^{At} 总有逆矩阵 e^{-At}，对任何 A 和 t 都成立。从特征值看：A 的特征值 0 对应 e^{At} 的特征值 $e^0=1$，而 $e^{\lambda t}$ 永远不为 0。更精确地有 $\det e^{At}=e^{t\,\mathrm{tr}\,A}>0$（练习 22.4）。这和数的情形一样：$e^{at}$ 永远不等于 0，与 a 是否为 0 无关。

注 22.3　Strang 在课上把 e^{At} 的级数和几何级数 $(I-At)^{-1}=I+At+(At)^2+(At)^3+\cdots$ 对比。几何级数只在 At 的特征值都满足 $|\lambda t|<1$ 时收敛；指数级数因为有 $k!$ 作分母，对任何矩阵、任何 t 都收敛。代价是 e^{At} 很少有封闭形式，除非像本节这样先对角化，或者 A 的幂有规律。

历史与背景　一般认为，Peano 在 1888 年用逐次逼近求解线性微分方程组时，得到了今天所说的 e^{At} 级数。计算方面，Moler 与 Van Loan 在 1978 年发表综述《Nineteen Dubious Ways to Compute the Exponential of a Matrix》，比较了十九种算法并指出各自的数值隐患，2003 年又发表了「二十五年后」的修订版。MATLAB 的 `expm` 用的是其中推荐的「缩放与平方」配合 Padé 有理逼近。本讲的 $Se^{\Lambda t}S^{-1}$ 在理论上最清楚，但 S 接近奇异时数值上并不可靠。

22.4　稳定性

很多时候我们不需要解的精确表达式，只想知道 $t\to\infty$ 时解的去向：衰减到 0、停在某个稳态，还是无限增长。答案全写在特征值里。

22.4.1　只看特征值的实部

一般解 $u(t)=\sum_i c_i e^{\lambda_i t}x_i$ 中，每一项的大小由 $|e^{\lambda_i t}|$ 决定。写 $\lambda=\alpha+i\beta$，和一维情形的计算一样，

$$|e^{\lambda t}|=|e^{\alpha t}|\,|e^{i\beta t}|=e^{\alpha t}\cdot|\cos\beta t+i\sin\beta t|=e^{\alpha t}.$$

虚部 β 只贡献转圈（模恒为 1），大小完全由实部 $\alpha=\mathrm{Re}\,\lambda$ 控制。

定理 22.7（稳定性判据）　设 A 是实矩阵，并且在复数域上可对角化，特征值为 $\lambda_1,\dots,\lambda_n$，考察实初值的方程 $u'=Au$。

(1) **稳定**（stable）：对所有初值都有 $u(t)\to 0$，当且仅当所有 $\mathrm{Re}\,\lambda_i<0$。

(2) **稳态**（steady state）：若 $\lambda_1=0$，其余 $\mathrm{Re}\,\lambda_i<0$，则 $u(t)\to c_1x_1$。

(3) **爆炸**（blow-up）：若某个 $\mathrm{Re}\,\lambda_i>0$，则只要初值在 x_i 上的分量 $c_i\neq 0$，就有 $\|u(t)\|\to\infty$。

证明　把解写成 $u(t)=S\,\mathrm{diag}(e^{\lambda_1 t},\dots,e^{\lambda_n t})c$，其中 $c=S^{-1}u(0)$。由于 $|e^{\lambda t}|=e^{(\mathrm{Re}\,\lambda)t}$，所有模态都衰减到零，当且仅当所有实部严格为负；若有 $\mathrm{Re}\,\lambda\ge 0$，取相应特征方向，或在复特征值时取其实部、虚部，就能得到一条不趋于零的实解。这证明 (1)。

在 (2) 中，$\lambda_1=0$ 的模态恒为 c_1x_1，其余模态都衰减到零。对 (3)，$S^{-1}u(t)$ 的第 i 个坐标为 $c_ie^{\lambda_i t}$；若 $\mathrm{Re}\,\lambda_i>0$ 且 $c_i\neq 0$，该坐标模长趋于无穷。取与向量范数相容的矩阵范数，则 $\|u(t)\|\ge\|S^{-1}u(t)\|/\|S^{-1}\|$，所以 $\|u(t)\|\to\infty$。 ■

例 22.1 是情形 (2)：$\lambda = 0, -3$，解趋于 $c_1 x_1 = (\frac{2}{3}, \frac{1}{3})$。例 22.2 是情形 (3)：$\lambda = 1 > 0$。还剩一种边界情形：实部为 0 而 $\lambda \neq 0$，如例 22.5 的 $\pm i$，解既不衰减也不增长，一直在转圈。

注 22.4　A 不可对角化时，解里会出现 $te^{\lambda t}$、$t^2 e^{\lambda t}$ 等项（例 22.6）。$\operatorname{Re}\lambda < 0$ 时它们仍然趋于 0，因为指数衰减压倒任何多项式增长，所以判据 (1) 不变。但在实部为 0 的边界上会出问题：$te^{0t} = t$ 线性增长（练习 22.3）。

22.4.2　2×2 矩阵：迹与行列式

对 2×2 实矩阵，不用解特征方程也能判断稳定性。记 $T = \operatorname{tr} A = \lambda_1 + \lambda_2$，$D = \det A = \lambda_1 \lambda_2$，特征方程是 $\lambda^2 - T\lambda + D = 0$。

定理 22.8　2×2 实矩阵 A 满足 $\operatorname{Re}\lambda_1 < 0$ 且 $\operatorname{Re}\lambda_2 < 0$，当且仅当 $\operatorname{tr} A < 0$ 且 $\det A > 0$。

证明　分两种情况。*特征值是实数*：两个实数都是负数，当且仅当和 $T < 0$、积 $D > 0$（积为正说明同号，和为负说明同为负）。*特征值是共轭复数* $\alpha \pm i\beta$（$\beta \neq 0$）：这时 $D = \alpha^2 + \beta^2 > 0$ 自动成立，而 $T = 2\alpha$，所以实部为负当且仅当 $T < 0$。两种情况合起来，条件都是 $T < 0$ 且 $D > 0$。 ■

只看迹不够　$A = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$ 的迹 $T = -1 < 0$，但 $\lambda = 1$ 让解按 e^t 爆炸。负的迹只说明两个特征值之和为负，一个很负的特征值可以掩盖一个正的。行列式 $D = -2 < 0$ 揭示了问题：两个特征值异号。

特征值是 $\lambda = \frac{1}{2}\left(T \pm \sqrt{T^2 - 4D}\right)$。判别式 $T^2 - 4D > 0$ 时是两个实特征值；$T^2 - 4D < 0$ 时是一对共轭复特征值，解一边转圈一边伸缩。抛物线 $D = T^2/4$ 是两者的分界。以 T 为横轴、D 为纵轴，每个 2×2 矩阵对应平面上一点（图 22.3）。

■ **例 22.7　在迹-行列式平面上分类**　判断下列矩阵对应的方程 $u' = Au$ 在 $t \to \infty$ 时的行为：

$$\text{(a)} \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix} \quad \text{(b)} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{(c)} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \quad \text{(d)} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{(e)} \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$$

解　逐个算 T、D 和判别式，再解出特征值：

	T	D	$T^2 - 4D$	特征值	结论
(a)	-3	0	9	$0, -3$	趋于稳态（例 22.1）
(b)	0	-1	4	$1, -1$	爆炸（鞍点，例 22.2）
(c)	-2	2	-4	$-1 \pm i$	稳定，螺旋衰减
(d)	0	1	-4	$\pm i$	不衰减也不增长（中心，例 22.5）
(e)	-1	-2	9	$1, -2$	爆炸（鞍点）

以 (c) 为例：$\lambda^2+2\lambda+2=0$，$\lambda=-1\pm\sqrt{1-2}=-1\pm i$。只有 (c) 满足定理 22.8 的两个条件。(a) 落在边界 $D=0$ 上，有一个零特征值；(d) 落在边界 $T=0$ 上，特征值是纯虚数。这两个边界点的解都不会趋于 0，但也不会爆炸。 ♦

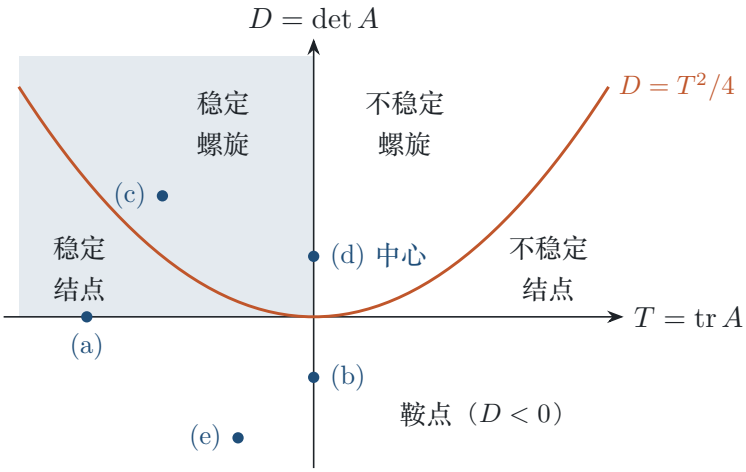


图 22.3　迹-行列式平面。浅蓝色区域 $T<0$，$D>0$ 是稳定区。橙色抛物线上方是复特征值（螺旋），下方是实特征值；$D<0$ 时两个特征值异号，是鞍点。点 (a)–(e) 对应例 22.7。

横坐标记录迹，纵坐标记录行列式。浅蓝区域恰好满足迹为负、行列式为正；抛物线把两个实特征值与一对共轭复特征值分开。图上 $D=0$ 与 $T=0$ 的边界也提醒我们，稳定性判据中的严格不等式不能放宽。

22.4.3　复特征值：螺旋

■ **例 22.8　衰减的螺旋**　解 $u'=Au$，$A=\begin{bmatrix}-1 & -4\\ 4 & -1\end{bmatrix}$，$u(0)=(1,0)$，并描述解的轨迹。

解　$T=-2$，$D=1+16=17$，特征方程 $\lambda^2+2\lambda+17=0$，$\lambda=-1\pm\sqrt{1-17}=-1\pm 4i$。$T<0$，$D>0$，所以稳定。

求 e^{At} 有一条捷径：$A=-I+4J$，其中 $J=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$ 是例 22.5 的矩阵。$-I$ 与 J 可交换，所以 $e^{At}=e^{-It}e^{4Jt}$（练习 22.6）。由例 22.5，e^{4Jt} 是转角为 $4t$ 的旋转矩阵：

$$e^{At}=e^{-t}\begin{bmatrix}\cos 4t & -\sin 4t\\ \sin 4t & \cos 4t\end{bmatrix},\qquad u(t)=e^{At}\begin{bmatrix}1\\0\end{bmatrix}=e^{-t}\begin{bmatrix}\cos 4t\\ \sin 4t\end{bmatrix}.$$

验证：$u_1'=e^{-t}(-\cos 4t-4\sin 4t)=-u_1-4u_2$，$u_2'=e^{-t}(-\sin 4t+4\cos 4t)=4u_1-u_2$，与 A 的两行一致。

轨迹到原点的距离是 $\|u(t)\|=e^{-t}$，转角是 $4t$。每转一圈用时 $\frac{\pi}{2}$，转一圈后半径缩为原来的 $e^{-\pi/2}\approx 0.208$；$t=\pi$ 时（两圈）只剩 $e^{-\pi}\approx 0.043$。**特征值的实部 -1 决定衰减快慢，虚部 4 决定转得多快**（图 22.4）。

也可以走特征向量的路：$\lambda=-1+4i$ 时 $(A-\lambda I)x=0$ 给出 $x=(1,-i)$，于是 $e^{\lambda t}x$ 是复

数解。A 是实矩阵，复数解的实部也是解，而 $\mathrm{Re}\big(e^{(-1+4i)t}(1,-i)\big) = e^{-t}(\cos 4t, \sin 4t)$，与上面相同。 ♦

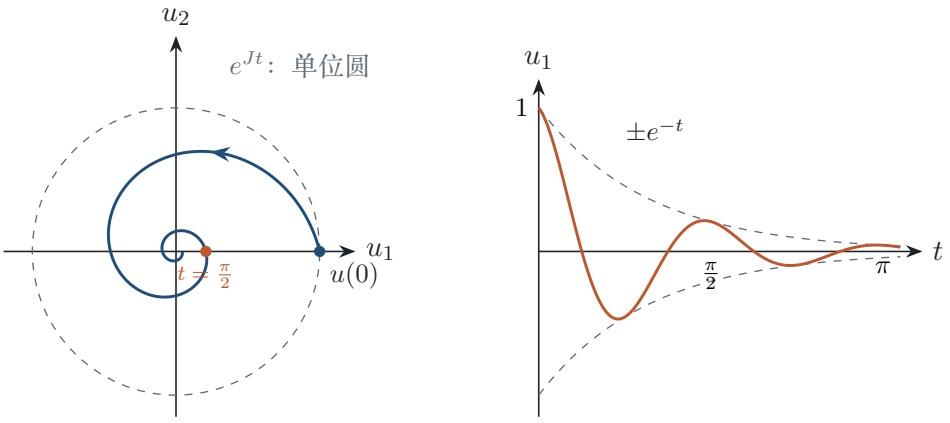


图 22.4　例 22.8。左：相平面上的轨迹 $u(t) = e^{-t}(\cos 4t, \sin 4t)$，箭头为运动方向；虚线是纯旋转 $e^{Jt}u(0)$ 的轨迹（例 22.5）。右：分量 $u_1 = e^{-t}\cos 4t$ 在包络线 $\pm e^{-t}$ 之间振荡。

左图每转一圈半径都缩小，右图则把同一运动拆成随时间振荡的一个坐标；两图中的衰减包络都是 e^{-t}。因此实部控制到原点的速度，虚部控制绕原点旋转的频率。

22.5　二阶方程与长度守恒

22.5.1　二阶方程改写成一阶系统

二阶方程的状态需要同时记录位置和速度。令

$$u(t) = \begin{bmatrix} y'(t) \\ y(t) \end{bmatrix}.$$

若 $y'' + by' + ky = 0$，则 $y'' = -by' - ky$，因此

$$u' = \begin{bmatrix} y'' \\ y' \end{bmatrix} = \begin{bmatrix} -b & -k \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y' \\ y \end{bmatrix} = Au.$$

这里的状态顺序是「速度在前、位置在后」。换成 $[y, y']^{\mathsf{T}}$ 也可以，但矩阵要随之改变。

m 阶方程 $y^{(m)} + a_{m-1}y^{(m-1)} + \cdots + a_0 y = 0$ 也可状态化：取 $u = (y^{(m-1)}, \ldots, y', y)^{\mathsf{T}}$，矩阵首行为 $(-a_{m-1}, \ldots, -a_0)$，次对角线全为 1，其余元素为 0；因此它的特征多项式是 $\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_0$，正好是原方程的特征多项式。

■ **例 22.9　把二阶方程化为特征值问题**　解 $y'' + 3y' + 2y = 0$，初值为 $y(0) = 1$、$y'(0) = 0$；并从状态矩阵看出特征多项式与原方程一致。

解　令 $u = (y', y)^{\mathsf{T}}$，则 $u' = Au$，其中 $A = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}$，初值 $u(0) = (0, 1)^{\mathsf{T}}$。特征多项式为

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda + 3 & 2 \\ -1 & \lambda \end{bmatrix} = \lambda^2 + 3\lambda + 2 = (\lambda + 1)(\lambda + 2),$$

正好是原方程的特征方程。对应特征向量可取 $x_1=(1,-1)^{\mathsf T}$、$x_2=(2,-1)^{\mathsf T}$，初值分解为 $(0,1)^{\mathsf T}=-2x_1+x_2$，故

$$u(t)=-2e^{-t}x_1+e^{-2t}x_2,\qquad y(t)=2e^{-t}-e^{-2t}.$$

求导得 $y'(t)=-2e^{-t}+2e^{-2t}$，所以两个初值都满足。把 y,y',y'' 代回原方程也为零。 ♦

图 22.5 把同一解画在时间图和相平面中：时间图读出位置、速度如何衰减，相平面则把 t 消去，显示状态如何回到原点。

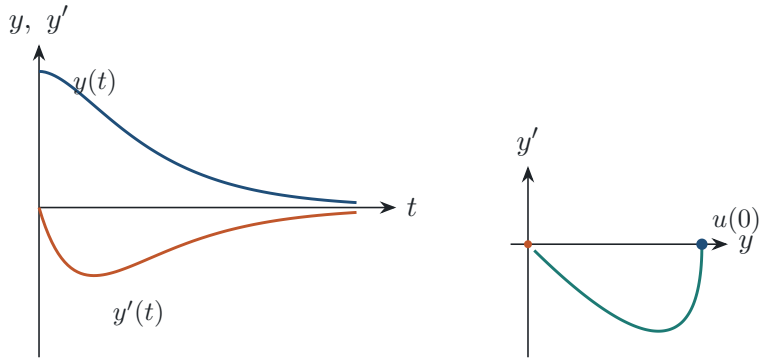


图 22.5　例 22.9 的状态轨迹：两个坐标都趋于零，右图中的曲线因此回到原点。

左图按时间显示 y,y'，右图显示状态 (y,y') 从 $(1,0)$ 回到原点。

22.5.2　反对称矩阵生成的长度守恒

旋转矩阵的生成矩阵 $J=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ 满足 $J^{\mathsf T}=-J$。这个条件推广到任意维后，仍能直接说明解的长度：

定理 22.9（反对称系统保持长度）　若实矩阵 $A^{\mathsf T}=-A$，且 $u'=Au$，则 $\|u(t)\|=\|u(0)\|$ 对所有 t 成立。

证明　按定义求平方长度的导数：

$$\frac{d}{dt}\|u(t)\|^2=\frac{d}{dt}(u^{\mathsf T}u)=(u')^{\mathsf T}u+u^{\mathsf T}u'=u^{\mathsf T}A^{\mathsf T}u+u^{\mathsf T}Au=u^{\mathsf T}(A^{\mathsf T}+A)u=0.$$

所以平方长度保持为常数。长度非负，故 $\|u(t)\|=\|u(0)\|$。 ■

因此 e^{At} 的每一列都保持内积关系，且 $(e^{At})^{\mathsf T}e^{At}=I$；矩阵指数本身是正交矩阵。若 $u(0)\neq 0$，轨迹始终落在以原点为球心、半径 $\|u(0)\|$ 的球面上。二维的例子正是例 22.5 的匀速圆周运动。

反对称条件不能去掉：若 $A=\begin{bmatrix}1&0\\0&0\end{bmatrix}$ 且 $u(0)=(1,0)^{\mathsf T}$，则 $u(t)=(e^t,0)^{\mathsf T}$，长度从 1 增长为 e^t。

本讲小结

- $Ax=\lambda x$ 时，$e^{\lambda t}x$ 是 $u'=Au$ 的纯指数解；若特征向量构成一组基，$u(t)=\sum_i c_i e^{\lambda_i t}x_i$，$c=S^{-1}u(0)$。
- 对任意方阵，$e^{At}=\sum_{k\ge 0}(At)^k/k!$ 收敛，满足 $(e^{At})'=Ae^{At}$、$e^{A(s+t)}=e^{As}e^{At}$，并有逆 e^{-At}。
- 若 $A=S\Lambda S^{-1}$，则 $e^{At}=Se^{\Lambda t}S^{-1}$；不可对角化时级数仍成立，幂零部分会带来 $te^{\lambda t}$ 等因子。
- 对角化的实系统全初值趋于 0 当且仅当所有特征值实部为负；2×2 实系统还可用 $\operatorname{tr}A<0$、$\det A>0$ 判断。
- $y''+by'+ky=0$ 可写为 $u'=Au$，$u=(y',y)^{\mathsf T}$、$A=\begin{bmatrix}-b & -k\\ 1 & 0\end{bmatrix}$，其特征多项式是 $\lambda^2+b\lambda+k$。
- $A^{\mathsf T}=-A$ 时 $\frac{d}{dt}\|u\|^2=0$，故 e^{At} 正交且解保持长度。

练习

22.1 ★　令 $A=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix}$。判断 $u_1(t)=e^{2t}(1,0)^{\mathsf T}$ 与 $u_2(t)=e^{-t}(1,1)^{\mathsf T}$ 哪些是 $u'=Au$ 的解，说明理由。

22.2 ★　解 $u'=Au$，其中 $A=\operatorname{diag}(-2,1)$、$u(0)=(2,-1)^{\mathsf T}$；判断解是否趋于零。

22.3 ★★　（原习题 23.1）设 $A=\begin{bmatrix}0 & c & -b\\ -c & 0 & a\\ b & -a & 0\end{bmatrix}$，$u'=Au$。从 $\|u(t)\|^2=u_1^2+u_2^2+u_3^2$ 出发求导，并说明 $u(t)$ 的长度及轨迹范围。

22.4 ★★　令 $N=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$。用级数求 e^{Nt}，再求 $u'=Nu$、$u(0)=(1,-2)^{\mathsf T}$ 的解。

22.5 ★★　对 $A=\begin{bmatrix}0 & 1\\ 0 & 2\end{bmatrix}$，求 e^{At} 并解 $u'=Au$、$u(0)=(1,1)^{\mathsf T}$。验证初值和微分方程。

22.6 ★★　设 A,B 是满足 $AB=BA$ 的方阵。用幂级数和二项式公式证明 $e^{A+B}=e^Ae^B$，并指出交换条件在本讲哪一个反例中失败。

22.7 ★★　设实矩阵 A 可在复数域上对角化。证明 $\det(e^{At})=e^{t\operatorname{tr}A}$，并说明此行列式的符号。

22.8 ★★　用迹、行列式或特征值判断下列系统的长期行为：$u'=\begin{bmatrix}-1 & -2\\ 2 & -1\end{bmatrix}u$。若 $u(0)=(1,0)^{\mathsf T}$，再写出解的显式形式。

22.9 ★★★　解 $y''+4y'+3y=0$，其中 $y(0)=2$、$y'(0)=-1$。写出状态矩阵，并验证特征多项式与原方程的特征方程相同。

22.10 ★★★ 对不可对角化矩阵 $J=\begin{bmatrix}-1 & 1\\ 0 & -1\end{bmatrix}$，求 e^{Jt}。取 $u(0)=(0,1)^{\mathsf T}$，解释解中出现的 te^{-t} 项为何仍趋于零。

练习答案与提示

22.1 $A(1,0)^{\mathsf T}=2(1,0)^{\mathsf T}$，所以 $u_1'=2e^{2t}(1,0)^{\mathsf T}=Au_1$，是解。对 $x=(1,1)^{\mathsf T}$，$Ax=(2,-1)^{\mathsf T}\neq -x$，故 $u_2'=-e^{-t}x$ 不等于 $Au_2=e^{-t}Ax$，不是解。

22.2 两个坐标满足 $u_1'=-2u_1$、$u_2'=u_2$，所以 $u(t)=(2e^{-2t},-e^{t})^{\mathsf T}$。第二分量无界，解不趋于零；全初值稳定要求每个特征值实部都为负。

22.3 方程给出 $u_1'=cu_2-bu_3$、$u_2'=au_3-cu_1$、$u_3'=bu_1-au_2$。逐项求导：

$$\frac{d}{dt}\|u\|^2=2u_1(cu_2-bu_3)+2u_2(au_3-cu_1)+2u_3(bu_1-au_2)=0.$$

因此 $\|u(t)\|^2=\|u(0)\|^2$，长度恒为 $\|u(0)\|$；非零初值的轨迹包含在半径为该长度的球面上，零初值则保持在原点。这正是来源题的长度守恒结论。

22.4 $N^2=0$，故 $e^{Nt}=I+tN=\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}$。因此 $u(t)=e^{Nt}u(0)=(1-2t,-2)^{\mathsf T}$；求导得 $u'=(-2,0)^{\mathsf T}=Nu$，且 $u(0)=(1,-2)^{\mathsf T}$。矩阵本身也满足 $(e^{Nt})'=N=Ne^{Nt}$，且 $e^{N0}=I$，因此它对任意初值都给出相应的解。

22.5 特征值为 $0,2$，对应向量为 $(1,0)^{\mathsf T},(1,2)^{\mathsf T}$。令 $S=\begin{bmatrix}1 & 1\\ 0 & 2\end{bmatrix}$，$S^{-1}=\begin{bmatrix}1 & -1/2\\ 0 & 1/2\end{bmatrix}$，得到

$$e^{At}=S\,\mathrm{diag}(1,e^{2t})S^{-1}=\begin{bmatrix}1 & (e^{2t}-1)/2\\ 0 & e^{2t}\end{bmatrix},\quad u(t)=\begin{bmatrix}(1+e^{2t})/2\\ e^{2t}\end{bmatrix}.$$

在 $t=0$ 时为 $(1,1)^{\mathsf T}$；导数 $(e^{2t},2e^{2t})^{\mathsf T}$ 等于 $Au(t)$。

22.6 由 $AB=BA$，二项式公式给出 $(A+B)^k=\sum_{j=0}^{k}\binom{k}{j}A^jB^{k-j}$。代入指数级数并重新按 j 与 $k-j$ 求和：

$$e^{A+B}=\sum_{j,\ell\ge 0}\frac{A^jB^\ell}{j!\ell!}=\left(\sum_{j\ge 0}\frac{A^j}{j!}\right)\left(\sum_{\ell\ge 0}\frac{B^\ell}{\ell!}\right)=e^Ae^B.$$

本讲的上三角反例中 $AB=B$、$BA=0$，所以交换条件不成立。重排时需要 A 与 B 可交换，才能把 $(A+B)^k$ 的每一项写成 A^jB^{k-j}；指数级数绝对收敛，故可以按 j 和 ℓ 重新分组。矩阵乘法的次序因此会直接影响结论。从二次项也能看出差别：e^{A+B} 中的交叉项是 $(AB+BA)/2$，而 e^Ae^B 中是 AB。这两种交叉项相等恰好要求 $AB=BA$；本讲的反例使它们不同。

22.7 写 $A=S\Lambda S^{-1}$，其中 $\Lambda=\mathrm{diag}(\lambda_1,\dots,\lambda_n)$。相似变换不改变迹，故 $\operatorname{tr}A=\sum_i\lambda_i$；又

$$\det(e^{At})=\det(S)\left(\prod_i e^{\lambda_i t}\right)\det(S^{-1})=e^{t\sum_i\lambda_i}=e^{t\operatorname{tr}A}.$$

实矩阵的迹为实数，所以该行列式严格为正。这也与 e^{At} 可逆相符。这里用到 $\det(S)\det(S^{-1})=1$。每个特征值贡献因子 $e^{\lambda_i t}$，这些因子的乘积正是迹对应的标量指数；即使特征值成共轭对，乘积仍为实数。

22.8 $T=-2$、$D=5$，特征方程为 $\lambda^2+2\lambda+5=0$，根为 $-1\pm 2i$，实部均负，故解趋于零。因 $A=-I+2J$ 且 $J^2=-I$，

$$u(t)=e^{-t}\begin{bmatrix}\cos 2t & -\sin 2t\\ \sin 2t & \cos 2t\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}=e^{-t}(\cos 2t,\sin 2t)^{\mathsf T}.$$

解的长度为 e^{-t}，每单位时间转过 2 弧度；转动不会改变半径，指数因子则使半径趋于零。因此轨迹是向原点收缩的螺旋。当 $t=0$ 时解为 $(1,0)^{\mathsf{T}}$，仍满足给定初值；任意时刻的模长为 e^{-t}，所以不会在有限时刻变成零，但当 $t\to\infty$ 时确实趋于零。

22.9 状态 $u=(y',y)^{\mathsf{T}}$ 满足 $u'=\begin{bmatrix}-4 & -3\\ 1 & 0\end{bmatrix}u, u(0)=(-1,2)^{\mathsf{T}}$。特征多项式为 $\det(\lambda I-A)=\lambda^2+4\lambda+3=(\lambda+1)(\lambda+3)$，与原方程一致。令 $y=C_1e^{-t}+C_2e^{-3t}$；初值给出 $C_1+C_2=2$、$-C_1-3C_2=-1$，解得 $C_1=5/2$、$C_2=-1/2$。故 $y(t)=\frac{5}{2}e^{-t}-\frac{1}{2}e^{-3t}$，直接代入可验两个初值及微分方程。具体地，第二个方程等价于 $C_1+3C_2=1$，与 $C_1+C_2=2$ 相减得 $2C_2=-1$。两根 $-1,-3$ 都使 $r^2+4r+3=0$，所以两个指数项分别满足微分方程；上面的系数再保证初值成立。相应的状态解为

$$u(t)=\begin{bmatrix}-\frac{5}{2}e^{-t}+\frac{3}{2}e^{-3t}\\ \frac{5}{2}e^{-t}-\frac{1}{2}e^{-3t}\end{bmatrix}.$$

其第一坐标是 y'，第二坐标是 y。代入状态矩阵后，第一行给出 $y''=-4y'-3y$，第二行给出 $y'=y'$，与状态方程的两行一致。

22.10 写 $J=-I+N$，$N^2=0$，且 $-I$ 与 N 可交换，所以 $e^{Jt}=e^{-t}e^{Nt}=e^{-t}(I+tN)=e^{-t}\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}$。于是 $u(t)=(te^{-t},e^{-t})^{\mathsf{T}}$，两个坐标都趋于零。直接求导得 $u'=Ju$，且 $u(0)=(0,1)^{\mathsf{T}}$。又因 $e^t\ge t^2/2$ $(t>0)$，有 $0\le te^{-t}\le 2/t\to 0$，故 $u(t)\to 0$。唯一特征值为 -1，特征空间只有一维，故 J 不可对角化；Jordan 块产生的 t 因子改变瞬态表达式，却不改变此例的长期极限。

第 23 讲

Markov 矩阵与 Fourier 级数

Markov matrices and Fourier series

对应材料：MIT 18.06SC Lecture 24; Problems 24.1–24.3

本讲目标

(1) 会判断实矩阵是否为列随机矩阵，并验证它的幂仍保持非负性与列和 1。

(2) 会证明列随机矩阵有特征值 1 且所有特征值满足 $|\lambda| \le 1$，并能区分特征值 1 与长时收敛。

(3) 会用特征向量分解求 Markov 链的有限步状态与稳态，并说明严格正矩阵为何收敛。

(4) 会用积分内积解释函数的正交投影，推导 2π 周期 Fourier 系数。

(5) 会计算三角多项式的 Fourier 系数，并区分有限阶投影与无限级数收敛问题。

23.1 列和为 1 意味着什么?

考虑两个地区的人口向量 $u_k = (u_{\text{加}}, u_{\text{麻}})^{\mathsf{T}}$。若每一步都有 90% 的加州居民留在加州、10% 搬去麻省，而麻省居民中 20% 搬去加州、80% 留在麻省，那么

$$u_{k+1} = A u_k, \qquad A = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}.$$

两列分别对应「从加州出发」和「从麻省出发」，每列的比例和都是 1。从 $u_0 = (0, 1000)^{\mathsf{T}}$ 开始，一步后 $u_1 = (200, 800)^{\mathsf{T}}$，总人口仍为 1000。这两个性质正是 Markov 矩阵的定义所要保留的。

更一般地，u_k 可以记录人数，也可以先除以总人数而记录概率。若转移表示个人随机迁移，u_k 的坐标可理解为期望人数，所以不必总是整数；非负性和总和不变仍有明确含义。列随机约定让这两类状态都能用同一个乘法 $u_{k+1} = A u_k$ 表示。

定义 23.1 （Markov 矩阵） 实方阵 $A = (a_{ij})$ 若满足 $a_{ij} \ge 0$，且每一列的元素之和为 1，就称为**Markov 矩阵**（Markov matrix），也称列随机矩阵。

例如上面的 A 是 Markov 矩阵，因为 $0.9 + 0.1 = 1$、$0.2 + 0.8 = 1$；交换矩阵 $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 也满足定义，虽然它有零元素。矩阵 $\begin{bmatrix} 6/5 & -1/5 \\ -1/5 & 6/5 \end{bmatrix}$ 的列和仍为 1，却因含负数而不是 Markov 矩

阵。

易错提醒　容易把「每列和为 1」错写成「每行和为 1」。例如 A 的第一行和是 1.1，但它仍是 Markov 矩阵；本讲采用列随机约定，因为状态向量是列向量，Au 的总和由 $\mathbf{1}^{\mathsf{T}}A=\mathbf{1}^{\mathsf{T}}$ 保持。

有些教材把概率分布写成行向量，并要求转移矩阵每行和为 1。转回本讲的列向量约定时，应将该转移矩阵转置。

列和条件会自然带出特征值 1，但推理需要经过转置。下面的命题解释了这一步。

命题 23.2（转置矩阵保留特征值）　任意实方阵 A 与 A^{T} 有相同的特征多项式，因此有相同的特征值（在复数域上按代数重数计算）。

证明　对任意数 λ，

$$\det(\lambda I-A^{\mathsf{T}})=\det\big((\lambda I-A^{\mathsf{T}})^{\mathsf{T}}\big)=\det(\lambda I-A).$$

第一步使用 $\det(B^{\mathsf{T}})=\det(B)$，第二步使用 $(A^{\mathsf{T}})^{\mathsf{T}}=A$。两边作为 λ 的多项式相等，故特征多项式相同。 ■

■ **例 23.1　转置前后的特征向量**　令 $D=\begin{bmatrix}2&1\\0&3\end{bmatrix}$。求 D 与 D^{T} 的特征值和一组对应特征向量，并比较两组特征向量。

解　先比较特征多项式，再分别解两个零空间。$\det(\lambda I-D)=(\lambda-2)(\lambda-3)$，所以特征值是 $2,3$。对 D，可取 $x_2=(1,0)^{\mathsf{T}}$、$x_3=(1,1)^{\mathsf{T}}$；对 D^{T}，可取 $y_2=(1,-1)^{\mathsf{T}}$、$y_3=(0,1)^{\mathsf{T}}$。直接相乘得 $Dx_j=\lambda_j x_j$、$D^{\mathsf{T}}y_j=\lambda_j y_j$。特征值相同，右特征向量一般不同；$D^{\mathsf{T}}$ 的右特征向量正是 D 的左特征向量。 ♦

图 23.1 把矩阵 A 的每一列解释为从一个出发地区到各个目的地的比例。

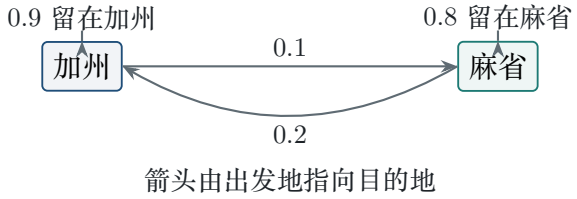


图 **23.1**　两个地区的人口转移中，每个出发地对应的留存和迁出比例合计为 1。

从加州出发的 0.9 与 0.1 是矩阵第一列；从麻省出发的 0.2 与 0.8 是第二列。矩阵乘列向量时，各列按出发地加权，这也解释了为何本讲检查列和。

想一想　为什么列和为 1 能推出特征值 1，而不一定有 $A\mathbf{1}=\mathbf{1}$?

答　列和条件给出 $A^{\mathsf{T}}\mathbf{1}=\mathbf{1}$，所以 A^{T} 有特征值 1；由命题 23.2，A 也有特征值 1。只有行和也为 1 时，才可直接写 $A\mathbf{1}=\mathbf{1}$。

■ **例 23.2**　**补全一个三州转移矩阵**　把

$$A=\begin{bmatrix}0.7 & 0.1 & 0.2\\ 0.1 & 0.6 & 0.3\\ ? & ? & ?\end{bmatrix}$$

补成 Markov 矩阵，并求一个稳态（steady state）向量。再说明为什么对称 Markov 矩阵的 $\mathbf{1}$ 一定是特征值 1 的特征向量。

解　逐列补足到 1：第一列缺 0.2，第二列缺 0.3，第三列缺 0.5。因此

$$A=\begin{bmatrix}0.7 & 0.1 & 0.2\\ 0.1 & 0.6 & 0.3\\ 0.2 & 0.3 & 0.5\end{bmatrix}.$$

各项非负且三列和都是 1。矩阵还满足 $A^{\mathsf T}=A$，所以行和也都是 1；于是 $A\mathbf{1}=\mathbf{1}$，一个稳态向量是 $(1,1,1)^{\mathsf T}$，对应总质量为 1 的概率稳态是 $(1/3,1/3,1/3)^{\mathsf T}$。一般地，对称性把列和条件转成行和条件，因此对称 Markov 矩阵都满足 $A\mathbf{1}=\mathbf{1}$。 ♦

■ **例 23.3**　**课时摘要中的三状态矩阵**　对

$$G=\begin{bmatrix}0.1 & 0.01 & 0.3\\ 0.2 & 0.99 & 0.3\\ 0.7 & 0 & 0.4\end{bmatrix}$$

求一个稳态向量并将它归一化为概率向量。

解　三列之和分别是 $0.1+0.2+0.7=1$、$0.01+0.99+0=1$、$0.3+0.3+0.4=1$，所以 G 是 Markov 矩阵。令稳态为 $(x,y,z)^{\mathsf T}$，方程 $G(x,y,z)^{\mathsf T}=(x,y,z)^{\mathsf T}$ 的前两行为

$$0.9x=0.01y+0.3z,\qquad 0.01y=0.2x+0.3z.$$

代入第二式，得 $0.9x=0.2x+0.6z$，即 $7x=6z$；再由第二式得 $y=55x$。取 $x=6$，便有 $(x,y,z)=(6,330,7)$。乘回验证 $G(6,330,7)^{\mathsf T}=(6,330,7)^{\mathsf T}$；坐标和为 343，所以概率稳态为 $(6/343,330/343,7/343)^{\mathsf T}$。 ♦

23.2　稳态何时吸引所有初值?

对矩阵 A，特征向量 $(2,1)^{\mathsf T}$ 对应特征值 1，而 $(1,-1)^{\mathsf T}$ 对应特征值 0.7。总人口沿第一个方向不变，两个地区的人口差异沿第二个方向每一步乘 0.7。这提示：谱半径控制长期行为。下面的定理把这个观察推广到 Markov 矩阵，并说明严格正性何时带来收敛。

定理 23.3（Markov 矩阵的谱界与稳态）　设 A 是 $n\times n$ 的 Markov 矩阵。则 $\mathbf{1}^{\mathsf T}A=\mathbf{1}^{\mathsf T}$，1 是 A 的特征值，且每个特征值都满足 $|\lambda|\le 1$。若进一步 $a_{ij}>0$ 对所有 i,j 成立，则每个 $\lambda\neq 1$ 都满足 $|\lambda|<1$，并且存在唯一满足 $p_*\ge 0$、$\mathbf{1}^{\mathsf T}p_*=1$、$Ap_*=p_*$ 的概率向量，使每个概率初值都满足 $A^kp_0\to p_*$。

直观理解　每列记录从一个出发状态流出的全部比例，所以总质量不会凭空产生或消失。若每个目的地都能从每个出发地得到正比例，两个初始分布之间的差异每一步都会缩小；只有「列和为 1」而没有严格正性时，状态仍可能周期振荡。

思路　先用列和与非负性证明质量守恒和 1 范围谱界。严格正时，取最小元素 δ，把 A 写成 $\delta\mathbf{1}\mathbf{1}^{\mathsf{T}}+B$；$B$ 在零和向量上给出严格收缩，再用相邻迭代差的几何级数证明收敛与唯一性。

证明　**列和与谱界**。列和条件等价于 $\mathbf{1}^{\mathsf{T}}A=\mathbf{1}^{\mathsf{T}}$，所以 $A^{\mathsf{T}}\mathbf{1}=\mathbf{1}$。命题 23.2 随即给出 $1\in\sigma(A)$。若 $z\neq 0$ 是复特征向量，$Az=\lambda z$，则

$$|\lambda|\|z\|_1=\|Az\|_1\le\sum_{i,j}a_{ij}|z_j|=\sum_j|z_j|\sum_i a_{ij}=\|z\|_1.$$

这里用到 $a_{ij}\ge 0$ 和每列和为 1，故 $|\lambda|\le 1$。同一列和等式也说明 $\mathbf{1}^{\mathsf{T}}Au=\mathbf{1}^{\mathsf{T}}u$；非负向量经 A 作用仍非负，因此概率单纯形（probability simplex）映到自身，A^k 仍是 Markov 矩阵。

严格正时的收缩。设 $\delta=\min_{i,j}a_{ij}>0$，令 $B=A-\delta\mathbf{1}\mathbf{1}^{\mathsf{T}}$、$\rho=1-n\delta$。$B$ 非负，且每列和为 ρ；由每列有 n 个至少为 δ 的元素，$0\le\rho<1$。若 $\mathbf{1}^{\mathsf{T}}z=0$，则 $\mathbf{1}\mathbf{1}^{\mathsf{T}}z=0$，所以 $Az=Bz$，并有

$$\|Az\|_1=\|Bz\|_1\le\sum_j|z_j|\sum_i B_{ij}=\rho\|z\|_1.$$

因此，对任意两个概率初值 p_0,q_0，它们的差坐标和为零，且

$$\|A^kp_0-A^kq_0\|_1\le\rho^k\|p_0-q_0\|_1.$$

这先说明不同初值的影响会消失；后面的 Cauchy 论证再证明迭代确实有一个概率向量作为共同极限。若 $Az=\lambda z$ 且 $\lambda\neq 1$，由 $\mathbf{1}^{\mathsf{T}}A=\mathbf{1}^{\mathsf{T}}$ 得 $(\lambda-1)\mathbf{1}^{\mathsf{T}}z=0$，所以 $\mathbf{1}^{\mathsf{T}}z=0$。上面的收缩估计随即给出 $|\lambda|\le\rho<1$。

存在与唯一。从任意概率向量 p_0 出发，令 $p_{k+1}=Ap_k$。差向量 p_1-p_0 的坐标和为 0，于是

$$\|p_{k+1}-p_k\|_1=\|A^k(p_1-p_0)\|_1\le\rho^k\|p_1-p_0\|_1.$$

右侧可求和，故 (p_k) 是 Cauchy 列。它的极限 p_* 仍非负且坐标和为 1；由 $p_{k+1}=Ap_k$ 取极限得 $Ap_*=p_*$。若 q_* 也是概率稳态，则两者之差坐标和为 0，对任意 k 有 $\|p_*-q_*\|_1\le\rho^k\|p_*-q_*\|_1$，令 $k\to\infty$ 得 $p_*=q_*$。∎

定理中的三个条件分别控制不同结论。去掉列和条件，$2I$ 虽非负却没有特征值 1；去掉非负性，矩阵 $\begin{bmatrix}6/5&-1/5\\-1/5&6/5\end{bmatrix}$ 列和为 1，但有特征值 $7/5>1$；去掉严格正性，交换矩阵仍是 Markov 矩阵，却会让状态在 $(1,0)^{\mathsf{T}}$ 与 $(0,1)^{\mathsf{T}}$ 间往返。

严格正性是保证所有初值收敛的方便充分条件，并非必要条件。例如 $R=\begin{bmatrix}1&1/3\\0&2/3\end{bmatrix}$ 含零元素，却有特征值 $1,2/3$；从 $(0,1)^{\mathsf{T}}$ 出发，$R^k(0,1)^{\mathsf{T}}=(1-(2/3)^k,(2/3)^k)^{\mathsf{T}}$ 仍收敛到 $(1,0)^{\mathsf{T}}$。能否收敛要看非稳态方向是否衰减，不能只数矩阵中的零元素。

■ **例 23.4　两州人口的稳态**　对 $A=\begin{bmatrix}0.9 & 0.2\\ 0.1 & 0.8\end{bmatrix}$ 和 $u_0=(0,1000)^{\mathsf T}$，求 $u_k=A^k u_0$ 的显式公式与极限。

解　沿两个特征向量分解初始人口。因为 $A(2,1)^{\mathsf T}=(2,1)^{\mathsf T}$、$A(1,-1)^{\mathsf T}=0.7(1,-1)^{\mathsf T}$，且

$$u_0=\frac{1000}{3}\begin{bmatrix}2\\1\end{bmatrix}-\frac{2000}{3}\begin{bmatrix}1\\-1\end{bmatrix},$$

所以

$$u_k=\frac{1000}{3}\begin{bmatrix}2\\1\end{bmatrix}-\frac{2000}{3}(0.7)^k\begin{bmatrix}1\\-1\end{bmatrix}\longrightarrow\begin{bmatrix}2000/3\\1000/3\end{bmatrix}.$$

首步为 $(200,800)^{\mathsf T}$，第二步为 $(340,660)^{\mathsf T}$，每一步总和都是 1000。长期比例为加州 2/3、麻省 1/3；0.7 模态决定接近稳态的速度。 ♦

把两个特征向量作为 S 的列，就得到

$$S=\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix},\qquad \Lambda=\begin{bmatrix}1 & 0\\ 0 & 0.7\end{bmatrix},\qquad S^{-1}=\begin{bmatrix}1/3 & 1/3\\ 1/3 & -2/3\end{bmatrix}.$$

直接相乘可得 $A=S\Lambda S^{-1}$，而 $S^{-1}u_0=(1000/3,-2000/3)^{\mathsf T}$ 正是初始状态在两种特征模态上的坐标。矩阵对角化把「逐方向乘以特征值」写成一条矩阵公式。

表 23.1 将前几步并列，便于同时检查总人口守恒和比例变化。

k	加州人数	麻省人数
0	0	1000
1	200	800
2	340	660
3	438	562
4	506.6	493.4

表 23.1　总人口始终为 1000，两个州的人口逐步接近稳态比例 2 : 1。

稳态也可以直接由方程求出。若人口为 $(c,m)^{\mathsf T}$ 且一步后不变，第一行给出

$$c=0.9c+0.2m,\qquad 0.1c=0.2m,\qquad c=2m.$$

再用总人口 $c+m=1000$，得到 $m=1000/3$、$c=2000/3$。这与特征向量 $(2,1)^{\mathsf T}$ 给出的比例相同；线性方程求比例，初始向量分解则说明从哪里开始都如何靠近这个比例。

把人数除以总人口，令 p_k 是加州所占比例，则麻省比例为 $1-p_k$。从转移矩阵第一行得到

$$p_{k+1}=0.9p_k+0.2(1-p_k)=0.2+0.7p_k.$$

固定点满足 $p_*=0.2+0.7p_*$，所以 $p_*=2/3$；相减得 $p_{k+1}-p_*=0.7(p_k-p_*)$。反复递推便有 $p_k-p_*=0.7^k(p_0-p_*)$。这个标量式和矩阵特征向量分解是同一条衰减规律的两种写法。

方法：分解稳态模态与衰减模态　要算 $A^k u_0$，先把 u_0 写成特征向量的线性组合，再分别把各分量乘以对应的 λ^k。$\lambda = 1$ 的分量保留，$|\lambda| < 1$ 的分量消失，$|\lambda| = 1$ 的非稳态分量可能留下振荡。

图 23.2 将人口向量归一化为加州比例 p_k。每个点都由递推 $p_{k+1} = 0.2 + 0.7p_k$ 得到。

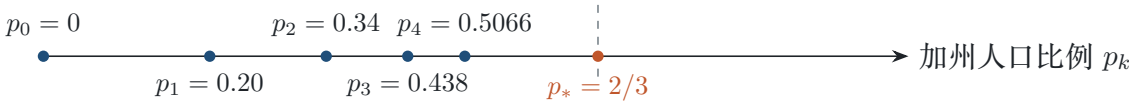


图 23.2　从全在麻省开始，加州人口比例逐步接近唯一稳态比例 $2/3$。

稳态是 $p_* = 2/3$，误差满足 $p_* - p_k = (2/3)(0.7)^k$。图中各点从左侧靠近橙色稳态点；误差每一步缩为原来的 0.7，这与第二个特征值完全一致。

■ 例 23.5　Markov 矩阵也可能有不同的矩阵类别　判断

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad E = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

是否可逆、正交、投影、置换、可对角化、Markov；并判断它们是否有 LU、QR、$S\Lambda S^{-1}$、$Q\Lambda Q^{\mathsf{T}}$ 分解。

解　先看 P：它交换第一与第三坐标，每行每列恰有一个 1，且 $P^{\mathsf{T}} = P$、$P^2 = I$。因此 P 是置换矩阵、正交矩阵、对称矩阵、可逆矩阵和 Markov 矩阵；它不是投影，因为 $P^2 \neq P$。向量 $(1,0,-1)^{\mathsf{T}}$ 对应特征值 -1，而 $(1,0,1)^{\mathsf{T}}$、$(0,1,0)^{\mathsf{T}}$ 对应特征值 1，所以其特征值为 $-1,1,1$。对称性给出正交对角化 $Q\Lambda Q^{\mathsf{T}}$，也给出 $S\Lambda S^{-1}$；可逆性保证 QR 分解。无行交换的 LU 不存在：首元 $P_{11} = 0$，但 $P_{31} = 1$，消元必须交换行。

E 满足 $E^{\mathsf{T}} = E$、$E^2 = E$，所以是投影到 $\operatorname{span}\{(1,1,1)^{\mathsf{T}}\}$ 上的对称 Markov 矩阵。具体地，$Ev = \frac{v_1+v_2+v_3}{3}(1,1,1)^{\mathsf{T}}$：它保留向量在均匀方向上的分量，消掉坐标和为零的部分。其特征值为 $1,0,0$，故可作两种特征分解；它秩为 1，所以不是正交或置换矩阵，但有薄 QR：$Q = \frac{1}{\sqrt{3}}(1,1,1)^{\mathsf{T}}$、$R = \frac{1}{\sqrt{3}}(1,1,1)$，满足 $Q^{\mathsf{T}}Q = 1$ 且 $QR = E$。无行交换的 LU 可行，但只有一个非零主元。这题说明 Markov、对称、投影等类别可以同时出现，也不能互相混为一谈。 ♦

想一想　「1 是 Markov 矩阵的特征值」是否意味着每个初值都收敛到稳态？

答　不意味着。交换矩阵 $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 的特征值是 $1,-1$；从 $e_1 = (1,0)^{\mathsf{T}}$ 出发，状态在 $e_1, e_2, e_1, e_2, \ldots$ 间交替。特征值 -1 的模等于 1，所以该模态不会衰减。严格正性在定理 23.3 中排除了这种情形。

23.3　从正交坐标到函数坐标

在平面中，单位正交基 q_1, q_2 给出坐标

$$v = (q_1^{\mathsf{T}} v) q_1 + (q_2^{\mathsf{T}} v) q_2 .$$

例如 $q_1 = (3/5, 4/5)^{\mathsf{T}}$、$q_2 = (-4/5, 3/5)^{\mathsf{T}}$ 是正交标准基。对 $v = (2,1)^{\mathsf{T}}$，两个坐标是 $q_1^{\mathsf{T}} v = 2$、$q_2^{\mathsf{T}} v = -1$，所以 $v = 2q_1 - q_2$。坐标的求法只用到正交性：与 q_i 点乘会消掉其他方向。

命题 23.4（正交标准基的坐标公式）　若 $q_1, \dots, q_n$ 是 $\mathbb{R}^n$ 的正交标准基，则任意 $v \in \mathbb{R}^n$ 都有

$$v = \sum_{j=1}^{n} (q_j^{\mathsf{T}} v) q_j .$$

若 $Q = [q_1 \ \cdots \ q_n]$，则 $Q^{\mathsf{T}} Q = I$、$Q^{-1} = Q^{\mathsf{T}}$，坐标列向量为 $Q^{\mathsf{T}} v$。

证明　因为 q_j 是基，存在系数 c_j 使 $v = \sum_j c_j q_j$。左乘 q_i^{T}，得到

$$q_i^{\mathsf{T}} v = \sum_j c_j q_i^{\mathsf{T}} q_j = c_i ,$$

其中 $q_i^{\mathsf{T}} q_j = 0$（$i \neq j$）且 $q_i^{\mathsf{T}} q_i = 1$。代回即得公式。把所有列合在一起就是 $Q(Q^{\mathsf{T}} v) = v$，所以 $Q^{-1} = Q^{\mathsf{T}}$。 ■

■ **例 23.6　正交但未单位化的坐标**　令 $q_1 = (2,0)^{\mathsf{T}}$、$q_2 = (0,3)^{\mathsf{T}}$、$v = (6,3)^{\mathsf{T}}$。求 v 在这组正交基下的坐标。

解　由于基向量长度不是 1，坐标要除以长度平方：

$$c_1 = \frac{v^{\mathsf{T}} q_1}{q_1^{\mathsf{T}} q_1} = \frac{12}{4} = 3, \qquad c_2 = \frac{v^{\mathsf{T}} q_2}{q_2^{\mathsf{T}} q_2} = \frac{9}{9} = 1 .$$

于是 $v = 3q_1 + q_2 = (6,3)^{\mathsf{T}}$。这说明正交只会让坐标分离；只有再单位化，坐标才简化成单纯的点积。 ♦

想一想　函数正交基若没有单位长度，坐标公式要怎样改？
答　与向量情形一样，$c_j = \langle f, q_j \rangle / \langle q_j, q_j \rangle$。三角函数的范数平方分别是 2π 与 π，所以常数项和其他 Fourier 系数的分母不同。

23.4　三角函数的正交性与 Fourier 系数

注 23.1　[阅读准备] 本节用到定积分、三角函数的积分和分部积分。尚未学过积分时，可暂缓本节及对应练习，保留前面的 Markov 矩阵和有限维正交坐标公式；第 25 讲的离散 Fourier 变换与 FFT 使用有限和，不依赖这里的积分推导。

在 $\mathbb{R}^n$ 中，向量点积是离散求和。这里先取 $[0,2\pi]$ 上的连续实函数组成的向量空间 $C([0,2\pi])$，定义

$$\langle f,g\rangle=\int_0^{2\pi} f(x)g(x)\,\mathrm{d}x$$

称为**函数内积**（function inner product）。连续性保证积分存在；若 f 在某点非零，则它在一段正长度区间上的绝对值有正下界，因此 $\langle f,f\rangle>0$。这说明该公式在上述函数空间中满足内积的正定性。函数 $1,\cos x,\sin x,\cos 2x,\sin 2x,\dots$ 就像一组候选正交方向。选择区间 $[0,2\pi]$ 是因为这些三角函数在这里走过整数个周期。三角多项式和 $f(x)=x$ 都属于这个连续函数空间，后面的例题可以直接使用该内积。

注 23.2　[更一般的函数空间] 在 $L^2([0,2\pi])$ 中，使用平方可积函数的等价类：仅在零测集上不同的函数视为同一个元素，即按几乎处处相等来识别。这样 $\int f^2=0$ 才等价于该元素为零。以下例题只用连续函数，无需先学习这项推广。

定理 23.5（三角函数正交关系）　对正整数 m,n，有

$$\int_0^{2\pi}\cos(mx)\cos(nx)\,\mathrm{d}x=\begin{cases}\pi, & m=n,\\ 0, & m\neq n,\end{cases}$$

$$\int_0^{2\pi}\sin(mx)\sin(nx)\,\mathrm{d}x=\begin{cases}\pi, & m=n,\\ 0, & m\neq n,\end{cases}\qquad \int_0^{2\pi}\sin(mx)\cos(nx)\,\mathrm{d}x=0,$$

且 $\int_0^{2\pi}\cos(nx)\,\mathrm{d}x=\int_0^{2\pi}\sin(nx)\,\mathrm{d}x=0$。

思路　把乘积化成和：两个同类乘积用积化和公式，混合乘积用 $\sin u\cos v=[\sin(u+v)+\sin(u-v)]/2$。非零整数频率的正弦、余弦在整周期上的积分都是 0；只有频率为 0 的常数项留下。

证明　对任意非零整数 k，

$$\int_0^{2\pi}\cos(kx)\,\mathrm{d}x=\Big[\frac{\sin(kx)}{k}\Big]_0^{2\pi}=0,\qquad \int_0^{2\pi}\sin(kx)\,\mathrm{d}x=\Big[-\frac{\cos(kx)}{k}\Big]_0^{2\pi}=0.$$

这两条计算是下面积化和论证的基础。

例如

$$\cos(mx)\cos(nx)=\frac{1}{2}\big(\cos((m-n)x)+\cos((m+n)x)\big).$$

当 $m\neq n$ 时，两个频率都非零，积分为 0；当 $m=n$ 时，第一项是 1，第二项频率为 $2m$，积分为 $\frac{1}{2}(2\pi)=\pi$。正弦乘积公式给出 $\frac{1}{2}(\cos((m-n)x)-\cos((m+n)x))$，同样只有 $m=n$ 时留下常数，积分为 π。混合乘积是两个正弦的和，积分为 0。最后，直接积分 $\sin(nx)$、$\cos(nx)$ 即得它们与常数 1 正交。 ■

如果希望把系数写成「与单位向量点乘」，可把正交函数单位化：

$$q_0(x)=\frac{1}{\sqrt{2\pi}},\qquad q_n^c(x)=\frac{\cos(nx)}{\sqrt{\pi}},\qquad q_n^s(x)=\frac{\sin(nx)}{\sqrt{\pi}}.$$

这些函数范数都是 1，且彼此正交，所以单位化坐标是 $\langle f, q\rangle$。原函数组的范数平方分别为 2π 和 π，因此公式 23.6 有两种分母。把内积整体乘以常数不会改变投影。

例如 $\int_0^{2\pi} \sin x \cos x \, \mathrm{d}x = 0$，而 $\int_0^{2\pi} \sin^2 x \, \mathrm{d}x = \pi$。正交不等于「函数图像互不相交」；它表示乘积的带符号面积相消。

图 23.3 画出 $\sin x \cos x$ 在一个周期中的正负面积；它们恰好抵消。

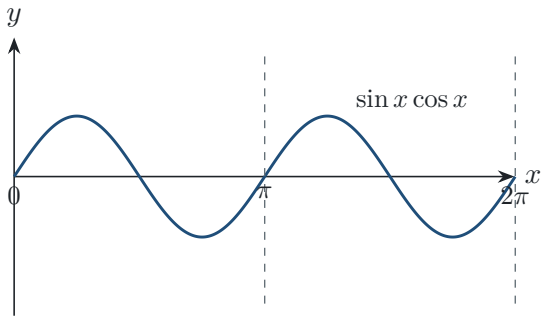


图 23.3　一个周期中 $\sin x \cos x$ 的正面积与负面积相等，所以其积分为 0。

图中函数在 0 到 $\pi/2$ 与 π 到 $3\pi/2$ 取正，在另外两段取负。图像帮助猜想，严格结论来自积化和公式与定积分。

这些正交函数允许我们像抽取向量坐标一样抽取函数坐标。把常数项和各频率的正弦、余弦项排列起来，得到**Fourier 级数**（Fourier series）；各方向的坐标称为**Fourier 系数**（Fourier coefficient）。

推论 23.6（有限 Fourier 投影的系数）　对平方可积函数 f，其到

$$V_N = \operatorname{span}\{1, \cos x, \sin x, \ldots, \cos(Nx), \sin(Nx)\}$$

的**正交投影**（orthogonal projection）写作

$$S_N f = a_0 + \sum_{n=1}^{N} \bigl(a_n \cos(nx) + b_n \sin(nx)\bigr).$$

系数为

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{d}x, \qquad a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x, \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x.$$

证明　正交投影的余差必须与 V_N 中每个基函数正交。常数函数的范数平方是 $\langle 1, 1\rangle = 2\pi$；对 $n \ge 1$，$\cos(nx)$ 与 $\sin(nx)$ 的范数平方都是 π。因此坐标公式分别是 $\langle f, 1\rangle/(2\pi)$、$\langle f, \cos(nx)\rangle/\pi$、$\langle f, \sin(nx)\rangle/\pi$，正是所列三式。 ■

命题 23.7（有限 Fourier 投影的最佳逼近性质）　$S_N f$ 是 V_N 中离 f 最近的函数，即对任意 $g \in V_N$，

$$\|f - g\|^2 = \|f - S_N f\|^2 + \|S_N f - g\|^2 \ge \|f - S_N f\|^2.$$

证明　令 $r = f - S_N f$、$h = S_N f - g$。r 与 V_N 正交，而 $h \in V_N$，所以 $\langle r, h\rangle = 0$。由 $f - g = r + h$ 展开平方范数，交叉项 $2\langle r, h\rangle$ 消失，得到所需等式。等号仅在 $h = 0$、也就是 $g = S_N f$ 时成立。 ■

若某函数恰好等于一个收敛的无限 Fourier 级数，还需要额外的收敛条件才能把无限和逐项积分。本讲只证明有限维投影和系数公式，不断言任意函数都等于自己的 Fourier 级数。

■ **例 23.7**　**把 $f(x) = x$ 投影到前三个方向**　求 $f(x) = x$ 在区间 $[0, 2\pi]$ 上到 $\operatorname{span}\{1, \cos x, \sin x\}$ 的 Fourier 投影。

解　分别计算三个内积。平均值给出 $a_0 = (2\pi)^{-1}\int_0^{2\pi} x\,\mathrm{d}x = \pi$；分部积分得 $\int_0^{2\pi} x\cos x\,\mathrm{d}x = 0$，所以 $a_1 = 0$；而

$$\int_0^{2\pi} x\sin x\,\mathrm{d}x = [-x\cos x + \sin x]_0^{2\pi} = -2\pi,$$

故 $b_1 = -2$。投影是 $p(x) = \pi - 2\sin x$。余差 $r(x) = x - p(x)$ 与 $1, \cos x, \sin x$ 的内积均为 0，这正是正交投影的判据。$f(x) = x$ 的端点值不同，所以这里只把它当作区间上的函数来投影，不讨论它的周期延拓在跳跃点如何收敛。 ♦

■ **例 23.8**　**恢复一组三角多项式系数**　设 $f(x) = 2 + 3\cos x - 4\sin 2x$。求 $a_0, a_1, a_2, a_3, b_1, b_2, b_3$。

解　由正交性，$a_0 = 2$、$a_1 = 3$、$a_2 = a_3 = 0$；$b_1 = 0$、$b_2 = -4$、$b_3 = 0$。代回可核对 $f(0) = 5$、$f(\pi/2) = 2$；后一处的 $\sin(2x) = \sin\pi = 0$。计算系数时用积分逐个筛选频率，不需要从许多函数值猜系数。 ♦

易错提醒　不要把 a_0 的归一化写成 $\pi^{-1}\int f$。常数函数 1 的长度平方是 2π，而 $\cos(nx)$、$\sin(nx)$ 的长度平方是 π；因此 a_0 的分母是 2π，其余系数的分母是 π。

前后联系　这里的 Fourier 级数使用实函数的积分内积。第 25 讲把正弦、余弦合并为复指数，构造 Fourier 矩阵并讨论离散 Fourier 变换与 FFT；那一讲的共轭转置负责复内积，本讲的积分内积是连续情形的对应工具。

本讲小结

- 列随机矩阵满足 $a_{ij} \ge 0$、$\mathbf{1}^{\mathsf{T}} A = \mathbf{1}^{\mathsf{T}}$；它保持非负性、总质量及概率单纯形。
- **转置不改变特征多项式**：$\det(\lambda I - A^{\mathsf{T}}) = \det(\lambda I - A)$。Markov 矩阵有特征值 1，且 $|\lambda| \le 1$。
- 若 A 的所有元素严格为正，则存在唯一概率稳态 $Ap_* = p_*$，且 $A^k p_0 \to p_*$；零元素时可能周期振荡。
- 正交标准基坐标为 $c = Q^{\mathsf{T}} v$；若只正交而未单位化，则 $c_j = \langle v, q_j\rangle / \langle q_j, q_j\rangle$。
- 在 $[0, 2\pi]$ 上，$\|1\|^2 = 2\pi$，而 $\|\cos(nx)\|^2 = \|\sin(nx)\|^2 = \pi$；不同频率彼此正交。
- 有限 Fourier 投影系数为 $a_0 = \frac{1}{2\pi}\int f$、$a_n = \frac{1}{\pi}\int f\cos(nx)$、$b_n = \frac{1}{\pi}\int f\sin(nx)$。

练习

23.1 ★ 判断真假并说明理由：(a) A 是 Markov 矩阵，则 A^k 对每个正整数 k 都是 Markov 矩阵；(b) 每个 Markov 矩阵都满足 $A\mathbf{1}=\mathbf{1}$；(c) 若 A 所有元素严格为正，则每个概率初值的迭代收敛到同一个稳态。

23.2 ★ 对 $A=\begin{bmatrix}0.9 & 0.2\\ 0.1 & 0.8\end{bmatrix}$ 求 A^2，并计算初值 $(0,1000)^{\mathsf T}$ 经过两步后的人口。

23.3 ★ 对该 A 验证 $(2,1)^{\mathsf T}$、$(1,-1)^{\mathsf T}$ 分别对应特征值 1、0.7，并写出从 $(1000,0)^{\mathsf T}$ 出发的稳态极限。

23.4 ★★ 将矩阵 $\begin{bmatrix}0.7 & 0.1 & 0.2\\ 0.1 & 0.6 & 0.3\\ ? & ? & ?\end{bmatrix}$ 补成 Markov 矩阵，求一个概率稳态，并说明补成的矩阵为何对称。

23.5 ★★ 对交换矩阵 $P=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$，求特征值和特征向量；用初值 $(1,0)^{\mathsf T}$ 说明 $P^k p_0$ 不收敛。

23.6 ★★ 求 $f(x)=x$ 在 $[0,2\pi]$ 上到 $\operatorname{span}\{1,\cos x,\sin x\}$ 的正交投影，并验证余差与三个基函数正交。

23.7 ★★ 用积化和公式证明 $\int_0^{2\pi}\cos(2x)\cos(3x)\,\mathrm{d}x=0$ 以及 $\int_0^{2\pi}\sin^2(2x)\,\mathrm{d}x=\pi$。

23.8 ★★★ 对 $g(x)=1-2\cos(2x)+5\sin x$ 求 a_0,a_1,a_2,b_1,b_2，并用 $x=0,\pi/2$ 核对表达式。

练习答案与提示

23.1 (a) 真：A^k 非负，且 $\mathbf{1}^{\mathsf T}A^k=(\mathbf{1}^{\mathsf T}A)A^{k-1}=\mathbf{1}^{\mathsf T}$。(b) 假：对题中的 A，$A\mathbf{1}=(1.1,0.9)^{\mathsf T}\neq\mathbf{1}$；列和条件给出的是 $A^{\mathsf T}\mathbf{1}=\mathbf{1}$。(c) 真，正是定理 23.3 的严格正情形。

23.2 $A^2=\begin{bmatrix}0.83 & 0.34\\ 0.17 & 0.66\end{bmatrix}$，故 $A^2(0,1000)^{\mathsf T}=(340,660)^{\mathsf T}$。两坐标相加仍是 1000。

23.3 $(1000,0)^{\mathsf T}=\frac{1000}{3}(2,1)^{\mathsf T}+\frac{1000}{3}(1,-1)^{\mathsf T}$，所以

$$u_k=\frac{1000}{3}\begin{bmatrix}2+0.7^k\\ 1-0.7^k\end{bmatrix}\longrightarrow\begin{bmatrix}2000/3\\ 1000/3\end{bmatrix}.$$

第二个模态趋于零，极限与起始地点无关。

23.4 补出的第三行为 $(0.2,0.3,0.5)$。完整矩阵对称，行和与列和均为 1，所以 $A\mathbf{1}=\mathbf{1}$。归一化后概率稳态为 $(1/3,1/3,1/3)^{\mathsf T}$。

23.5 $P(1,1)^{\mathsf T}=(1,1)^{\mathsf T}$，$P(1,-1)^{\mathsf T}=-(1,-1)^{\mathsf T}$，故特征值是 $1,-1$。又 $(1,0)^{\mathsf T}=\frac12(1,1)^{\mathsf T}+\frac12(1,-1)^{\mathsf T}$，所以

$$P^k(1,0)^{\mathsf T}=\frac12\big((1,1)^{\mathsf T}+(-1)^k(1,-1)^{\mathsf T}\big),$$

它在 $(1,0)^{\mathsf T}$ 与 $(0,1)^{\mathsf T}$ 之间交替，不收敛。

23.6 $a_0=\pi$、$a_1=0$、$b_1=-2$，所以投影为 $p(x)=\pi-2\sin x$。余差 $r=x-p$ 满足

$$\langle r,1\rangle=2\pi^2-2\pi^2=0,\quad \langle r,\cos x\rangle=0,\quad \langle r,\sin x\rangle=-2\pi+2\pi=0.$$

23.7 $\cos(2x)\cos(3x) = \frac{1}{2}(\cos x + \cos 5x)$，两个非零整数频率在 $[0, 2\pi]$ 上积分均为 0。又 $\sin^2(2x) = \frac{1}{2}(1 - \cos 4x)$，积分为 $\frac{1}{2}(2\pi) = \pi$。

23.8 $a_0 = 1$、$a_1 = 0$、$a_2 = -2$、$b_1 = 5$、$b_2 = 0$；$g(0) = 1 - 2 = -1$，$g(\pi/2) = 1 + 2 + 5 = 8$。

复习课 2

正交、行列式与线性系统

Review: Orthogonality, Determinants, and Linear Systems

对应材料：MIT 18.06SC Unit 2 Exam 与 Solutions（例题 1–4 据原卷改编）；第 13–23 讲各章对应的课时摘要、Problem/Exercises 及解答；MIT 18.06SC Final Review Summary（例题 6、8 参考其中题型）。Unit 2 Exam 未覆盖本复习课全部主题，其他主题据对应课时材料编写；本课不把改编题称为原卷原题。

本讲目标

(1) 能串起正交、投影、最小二乘、QR 与行列式之间的关系。

(2) 能用特征值和特征向量处理矩阵幂、线性微分方程与 Markov 迭代。

(3) 能将有限 Fourier 展开识别为函数空间中的正交投影。

(4) 能检查秩、满列秩、可对角化、列随机和归一化等条件，避开常见误用。

R2.1 知识地图

第 13–16 讲把垂直关系变成最近点、最小二乘和 QR 分解；第 17–19 讲用行列式识别可逆性、解方程并计算体积；第 20–23 讲沿特征方向计算矩阵幂和系统演化，并在 Markov 与 Fourier 问题中再次使用投影和特征值。图 R2.1 标出这几条复习主线。

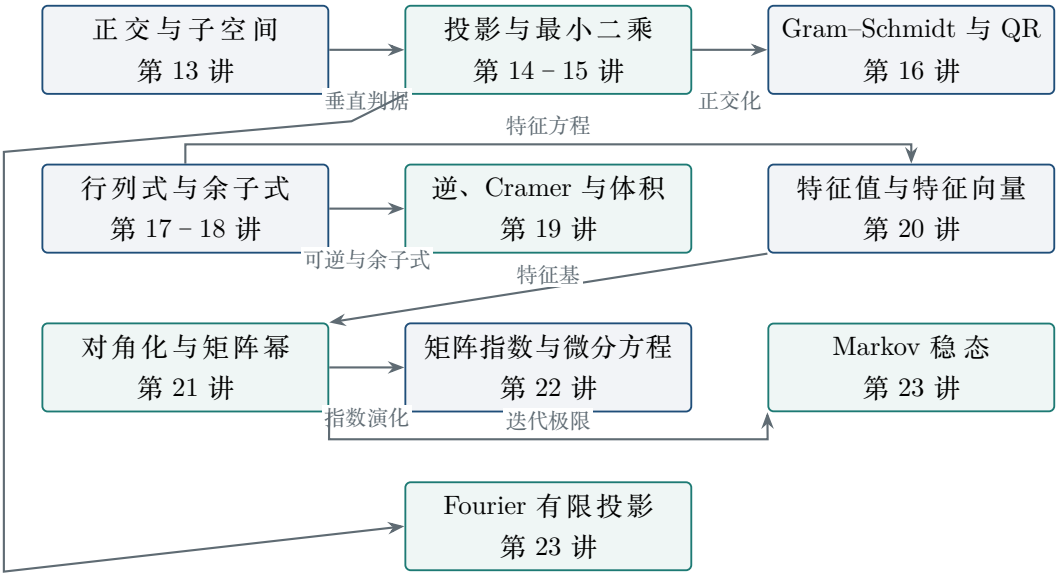


图 R2.1 正交投影连接最小二乘、QR 与 Fourier 系数；行列式连接线性系统和特征方程，特征分解处理矩阵幂与演化。

概念	定义要点	典型用法
正交补	$W^{\perp}=\{y: y^{\mathsf{T}}w=0 \text{ 对所有} w\in W\}$	识别残差与零空间
正交投影	$p\in W,\ b-p\in W^{\perp}$	在子空间内找最近点
最小二乘	$A^{\mathsf{T}}(b-A\hat{x})=0$	方程无精确解时求最佳拟合
QR 分解	$A=QR,\ Q^{\mathsf{T}}Q=I$	用三角方程求最小二乘解
行列式	$\det A\neq 0 \iff A$ 可逆	解方程、面积与特征方程
特征分解	$A=S\Lambda S^{-1}$	计算 A^k 与 e^{At}
Markov / **Fourier**	列和为 1；积分内积下正交	求稳态与有限频率投影

R2.2　核心结论回顾

R2.2.1　正交与逼近（第 13–16 讲）

第 13 讲：正交补与基本子空间

> **定理（正交关系）**
>
> (a) 对 $A\in\mathbb{R}^{m\times n}$，$N(A)=C(A^{\mathsf{T}})^{\perp}$、$N(A^{\mathsf{T}})=C(A)^{\perp}$；若 $\dim W=k$，则 $\dim W^{\perp}=n-k$ 且 $(W^{\perp})^{\perp}=W$。
>
> (b) 两两正交且都非零的向量组线性无关；非零条件不能删去，因为 $(0,v)$ 仍是正交组却线性相关。

关键　每个正交关系都先核对向量所在空间；$Ax=0$ 说明 x 垂直于 A 的行，不是列。

用法　判断 $Ax=b$ 是否有解时，可寻找左零空间中的向量 y：若 $A^{\mathsf{T}}y=0$ 而 $y^{\mathsf{T}}b\neq 0$，则 $b\notin C(A)$。

第 14 讲：最近点与正交投影

> **定理（投影刻画）**
>
> (a) 对子空间 W，p 是 b 在 W 中唯一的最近点，当且仅当 $p\in W$ 且 $b-p\in W^{\perp}$。
>
> (b) 若 $a\neq 0$，则 $\operatorname{proj}_{\operatorname{span}\{a\}}b=\dfrac{a^{\mathsf{T}}b}{a^{\mathsf{T}}a}a$；若 $a=0$，此分母为零，直线方向公式不适用。

关键　把 $b=p+e$ 写成子空间内的分量与正交误差，再用 $\|b-q\|^2=\|e\|^2+\|p-q\|^2$ 比较任意 $q\in W$。

用法　算出候选点后检查 $p\in W$ 和 $e\perp W$；只验证 p 在子空间内还不足以保证它最近。

第 15 讲：投影矩阵与最小二乘

> **定理（投影与法方程）**
>
> (a) 若 A 满列秩，则 $P = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$ 满足 $P^{\mathsf{T}} = P$、$P^2 = P$，并投影到 $C(A)$。
>
> (b) 对任意 A, b，$\hat{x}$ 最小化 $\|Ax - b\|^2$ 当且仅当 $A^{\mathsf{T}}A\hat{x} = A^{\mathsf{T}}b$；拟合向量唯一，系数在 A 列相关时可能不唯一。

关键　令 $e = b - A\hat{x}$；最优条件等价于 $A^{\mathsf{T}}e = 0$，即误差垂直于每一列，而非只要求误差分量和为零。

用法　使用 $(A^{\mathsf{T}}A)^{-1}$ 前先检验列独立；秩亏时仍可求解法方程，但不能写逆矩阵公式。

第 16 讲：正交矩阵、Gram–Schmidt 与 QR

> **定理（正交坐标与 QR）**
>
> (a) $Q^{\mathsf{T}}Q = I$ 时，Q 保持内积和长度；若 Q 是方阵，则 $Q^{-1} = Q^{\mathsf{T}}$。
>
> (b) 列满秩的 A 可写成 $A = QR$，其中 $Q^{\mathsf{T}}Q = I$、R 上三角且对角线为正；Gram–Schmidt 每步减去已有正交方向的投影。

关键　对新列 a_k，先取 $u_k = a_k - \sum_{i<k}(q_i^{\mathsf{T}}a_k)q_i$，再令 $q_k = u_k/\|u_k\|$；输入列独立保证 $u_k \neq 0$。

用法　若列相关，某一步会出现零残差，不能除以其长度；若用 QR 解最小二乘，则解 $R\hat{x} = Q^{\mathsf{T}}b$。

R2.2.2 行列式与线性系统（第 17–19 讲）

第 17 讲：性质与消元

> **定理（行列式的计算规则）**
>
> (a) $n \times n$ 矩阵满足 $\det I = 1$、交换两行变号、固定其余行时对一行线性；三角矩阵的行列式为对角元之积。
>
> (b) 同阶方阵满足 $\det(AB) = \det A \det B$；$\det A \neq 0$ 当且仅当 A 可逆，且 $\det(A^{\mathsf{T}}) = \det A$。

关键　消元时分别记录换行和缩放：行加另一行的倍数不改行列式，换行乘 -1，单行乘 c 则行列式也乘 c。

用法　行列式对每一行分别线性，不代表 $\det(A + B) = \det A + \det B$；计算稠密矩阵时通常用消元。

第 18 讲：排列公式与余子式

> **定理（余子式展开）**
>
> (a) 代数余子式 $C_{ij} = (-1)^{i+j} \det A_{ij}$；沿任一行或列展开都得到 $\det A$。
>
> (b) 固定其他元素，只把 a_{ij} 改为 $a_{ij} + t$，则 $\det(A + tE_{ij}) = \det A + tC_{ij}$，无需假设 A 可逆。

关键　每个排列项从每行、每列各取一个元素；余子式的棋盘符号已包含在 C_{ij} 中，使用它时不要再加一次符号。

用法　稀疏矩阵沿零最多的行或列展开；只有行列式阶数较小时，排列公式才适合直接计算。

第 19 讲：逆、Cramer 法则与体积

> **定理（伴随矩阵与几何因子）**
>
> (a) $\operatorname{adj} A = C^{\mathsf{T}}$ 且 $A \operatorname{adj} A = (\det A) I$；仅当 $\det A \neq 0$ 时，$A^{-1} = \operatorname{adj} A / \det A$。Cramer 法则也要求 $\det A \neq 0$。
>
> (b) 方阵变换把体积乘以 $|\det A|$；有向面积保留符号，三角形面积还要除以 2。

关键　Cramer 法则的第 j 个分子只替换 A 的第 j 列为 b；$\det A = 0$ 时，比值无定义，须回到相容性和零空间。

用法　求单个小系统可用 Cramer；多个右端或大矩阵优先消元。求几何面积时最后取绝对值。

R2.2.3 特征结构与演化（第 20–23 讲）

第 20 讲：特征值与特征向量

> **定理（特征方程）**
>
> (a) 特征值满足 $\det(A - \lambda I) = 0$；对每个根再解 $(A - \lambda I)x = 0$，并排除零向量。
>
> (b) 按代数重数计，特征值之和为 $\operatorname{tr} A$、乘积为 $\det A$；重复根的特征空间仍须单独求出。

关键　特征多项式只给出可能的倍数；特征向量来自对应零空间，不能把重根的代数重数直接当作独立方向数。

用法　先用迹和行列式检查已求的根，再代回特征方程，并用 $Ax = \lambda x$ 检查每个方向。

第 21 讲：对角化与矩阵幂

> **定理（特征基与迭代）**
>
> (a) A 可对角化当且仅当有 n 个线性无关特征向量；若 $A = S\Lambda S^{-1}$，则 $A^k = S\Lambda^k S^{-1}$。
>
> (b) $u_{k+1} = Au_k$ 的解为 $u_k = A^k u_0$；若 $S^{-1}u_0 = c$，则每个特征方向的系数按相应 λ_i^k 缩放。

用法　特征值互异足以保证可对角化；有重根时要数各特征空间的维数。求单条轨迹极限时，只需检查初值中出现的特征方向。

第 22 讲：矩阵指数与微分方程

> **定理（线性系统解）**
>
> (a) $u' = Au$，$u(0) = u_0$ 的解为 $u(t) = e^{At}u_0$，其中 $e^{At} = \sum_{k\geq 0}(At)^k/k!$；若 $A = S\Lambda S^{-1}$，则 $e^{At} = Se^{\Lambda t}S^{-1}$。
>
> (b) 当 A 在复数域上可对角化时，所有初值解都趋于 0 当且仅当各特征值实部都小于 0；实部为正的模态会增长。

用法　先把高阶方程改写为一阶系统，再找特征值和系数；e^{At} 不是对 A 的每个元素分别取指数。

第 23 讲：Markov 矩阵与 Fourier 投影

> **定理（离散与连续正交投影）**
>
> (a) 列 Markov 矩阵 A 满足 $a_{ij} \geq 0$、$\mathbf{1}^{\mathsf{T}}A = \mathbf{1}^{\mathsf{T}}$；它有特征值 1 且 $|\lambda| \leq 1$。若所有元素严格为正，则概率稳态唯一且 $A^k p_0$ 收敛到它。
>
> (b) 在 $[0, 2\pi]$ 上，$1, \cos(nx), \sin(nx)$ 两两正交；到有限频率空间的投影系数为 $a_0 = \frac{1}{2\pi}\int f$、$a_n = \frac{1}{\pi}\int f\cos(nx)$、$b_n = \frac{1}{\pi}\int f\sin(nx)$。

关键　Markov 矩阵的列和按列向量约定检查；Fourier 系数是投影系数，常数函数的平方长度为 2π，其余正弦、余弦的平方长度为 π。

用法　严格正的转移矩阵可用唯一稳态判断长期分布；有限 Fourier 投影只给最佳有限维逼近，不推出任意 Fourier 级数都收敛到原函数。

R2.3　题型与例题

R2.3.1　题型一：正交方向、投影与最小二乘

看到什么　若题目出现「最近点」「最佳拟合」或「Gram–Schmidt」，先把误差写成正交条件，再判断列是否满秩。

例题 R2.1 用正交基列行列式的符号；例题 R2.2 与例题 R2.3 分别练习拟合和投影矩阵。

■ **例 R2.1　Unit 2 Exam 第 1 题改编：正交基的行列式**　设 q_1, q_2, q_3 是 $\mathbb{R}^3$ 的一组正交标准基。求

$$\det[q_1, q_2, q_3], \quad \det[q_1 + q_2, q_2 + q_3, q_3 + q_1], \quad \det[q_1, q_2, q_3]\det[q_2, q_3, q_1].$$

解　令 $\delta = \det[q_1, q_2, q_3]$。因 $Q = [q_1, q_2, q_3]$ 正交，$\delta^2 = \det(Q^{\mathsf{T}}Q) = 1$，故第一项为 ± 1。第二项按列多线性展开，含重复列的六项为零，剩下 $\delta + \det[q_2, q_3, q_1] = 2\delta$，所以取值为 ± 2。第三个行列式由循环置换两次得到，与 δ 相同，乘积为 $\delta^2 = 1$。 ♦

点评　正交标准列先看 $Q^{\mathsf{T}}Q = I$；多线性展开后，重复列项立即为零。列循环三次是偶置换，符号不变。

■ **例 R2.2　Unit 2 Exam 第 2 题改编：21 个时刻的直线拟合**　在 $t = -10, -9, \ldots, 10$ 观测 b_t，只有 $b_0 = 1$，其余为 0。用最小二乘拟合 $C + Dt$，并写出拟合向量和误差。

解　令 A 的两列为全 1 向量和时间向量 $(-10, \ldots, 10)^{\mathsf{T}}$。对称性给 $\sum t = 0$，且 $\sum t^2 = 2(1^2 + \cdots + 10^2) = 770$，所以

$$A^{\mathsf{T}}A = \begin{bmatrix} 21 & 0 \\ 0 & 770 \end{bmatrix}, \qquad A^{\mathsf{T}}b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

法方程给 $\hat{C} = 1/21, \hat{D} = 0$。拟合向量 p 的每个分量是 $1/21$；误差 $e = b - p$ 的中间分量为 $20/21$，其余 20 个分量为 $-1/21$。于是 $\mathbf{1}^{\mathsf{T}}e = 20/21 - 20/21 = 0$，$t^{\mathsf{T}}e = 0$，确有 $A^{\mathsf{T}}e = 0$。 ♦

点评　对称时刻让常数列与斜率列正交，Gram 矩阵变成对角阵；若把中间观测放错位置，右端第二个分量就不再为零。

■ **例 R2.3　Unit 2 Exam 第 3 题改编：同一子空间的两个投影矩阵**　设 $A \in \mathbb{R}^{5\times 3}$ 列独立，Gram–Schmidt 得到 $Q \in \mathbb{R}^{5\times 3}$，且 $C(Q) = C(A)$。写出投影矩阵 P_A, P_Q，说明它们的关系，并给出从新向量 a_4 得到的单位残差。

解　由满列秩投影公式，$P_A = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$；因 $Q^{\mathsf{T}}Q = I$，$P_Q = QQ^{\mathsf{T}}$。二者都投影到同一列空间，故 $P_A = P_Q$。这个 5×5 投影矩阵秩为 3，所以行列式为 0；并有 $P_Q Q = Q$。新残差为 $r = (I - P_A)a_4$，它垂直于前三列；若另有条件 $a_4 \notin C(A)$，则 $r \neq 0$，此时才能定义 $q_4 = r/\|r\|$。若 $a_4 \in C(A)$，残差为零，Gram–Schmidt 应在这里停止。 ♦

点评　换一组同空间的基会改变坐标，却不改变投影。新方向应从 a_4 中减去它在原列空间上的投影。

R2.3.2　题型二：行列式次数、Cramer 法则与几何量

看到什么　行列式含参数时先数每项最多能选几个参数；用 Cramer 前检查非零行列式，用于几何面积时再取绝对值。

例题 R2.4 改编自 Unit 2 Exam 第 4 题；例题 R2.5 把 Cramer 法则与平行四边形面积放在一起。

■ **例 R2.4　Unit 2 Exam 第 4 题改编：行列式的参数次数**　矩阵 $A(x)$ 的首行和首列全为 x，右下角 3×3 子矩阵任意。说明 $\det A(x)$ 的次数至多为 2，并在右下角为 I_3 时计算行列式和零点。

解　行列式的每个排列项从每行、每列各取一个元素。首行与首列不能同时选两个不同的交点，因此一项至多含两个 x，次数至多为 2。右下角为 I_3 时，

$$A(x)=\begin{bmatrix} x & x & x & x\\ x & 1 & 0 & 0\\ x & 0 & 1 & 0\\ x & 0 & 0 & 1 \end{bmatrix}.$$

从第 2–4 列各减去第一列，展开第一行后，右下角变成 $I_3-x\mathbf{1}\mathbf{1}^{\mathsf T}$，其行列式为 $1-3x$。所以 $\det A(x)=x(1-3x)$，次数确为 2，零点为 $x=0,\tfrac13$。 ♦

点评　排列项给出次数上界，单位块特例说明二次项确实能出现。参数取零时矩阵奇异，不能把零行列式当成 Cramer 比值的分母。

■ **例 R2.5　第 19 讲题型改编：Cramer 与面积**　解方程 $\begin{bmatrix}2&1\\1&1\end{bmatrix}x=\begin{bmatrix}1\\0\end{bmatrix}$，并求该矩阵把单位正方形变成的平行四边形面积。

解　$\det A=2-1=1$，故系统有唯一解。替换第一列、第二列分别得 $A_1=\begin{bmatrix}1&1\\0&1\end{bmatrix}$、$A_2=\begin{bmatrix}2&1\\1&0\end{bmatrix}$，所以 $x_1=\det A_1/\det A=1$、$x_2=\det A_2/\det A=-1$。代回得 $A(1,-1)^{\mathsf T}=(1,0)^{\mathsf T}$。面积因子为 $|\det A|=1$。 ♦

点评　Cramer 的每个分子只换对应列；面积用行列式绝对值。若 $\det A=0$，题目中的比值法就不能使用。

R2.3.3　题型三：特征方向、对角化与连续时间演化

看到什么　先从 $\det(A-\lambda I)=0$ 找倍数，再求零空间；只有得到足够多独立方向，才能把矩阵幂或指数化成标量幂或指数。

■ **例 R2.6　Final Review Summary 第 4 题改编：由特征对构造矩阵**　矩阵 A 的特征对为 $(0,(1,2)^{\mathsf T})$ 与 $(3,(2,1)^{\mathsf T})$。构造 A，并计算 A^k（$k\ge 1$）。

解　取 $S=\begin{bmatrix}1&2\\2&1\end{bmatrix}$、$\Lambda=\operatorname{diag}(0,3)$，则

$$S^{-1}=\begin{bmatrix}-1/3&2/3\\2/3&-1/3\end{bmatrix},\qquad A=S\Lambda S^{-1}=\begin{bmatrix}4&-2\\2&-1\end{bmatrix}.$$

因此 $A^k=S\operatorname{diag}(0,3^k)S^{-1}=3^{k-1}A$。直接检查 $A(1,2)^{\mathsf T}=0$、$A(2,1)^{\mathsf T}=3(2,1)^{\mathsf T}$，与给定特征对一致。 ♦

点评　把特征向量按与 Λ 对应的顺序放入 S 的列；交换列时也必须同步交换 Λ 的对角元。

■ 例 R2.7　第 22 讲题型改编：从特征模态解微分方程　解 $y''+3y'+2y=0$，满足 $y(0)=1, y'(0)=0$，并判断解的长期行为。

解　令 $u=(y,y')^{\mathsf{T}}$，则 $u'=Au$，其中 $A=\begin{bmatrix}0 & 1\\ -2 & -3\end{bmatrix}$。特征多项式为 $\lambda^2+3\lambda+2=(\lambda+1)(\lambda+2)$，特征向量可取 $(1,-1)^{\mathsf{T}},(1,-2)^{\mathsf{T}}$。由初值解得 $u(0)=2(1,-1)^{\mathsf{T}}-(1,-2)^{\mathsf{T}}$，故

$$y(t)=2e^{-t}-e^{-2t},\qquad y'(t)=-2e^{-t}+2e^{-2t}.$$

代入得 $y(0)=1, y'(0)=0$；两个特征值都为负，故 $y(t)\to 0$。 ♦

点评　状态变量的排列要与系统矩阵一致；这里采用 (y,y')，所以第二分量的导数才是第一分量。

R2.3.4 题型四：Markov 稳态与 Fourier 投影

看到什么　Markov 题先确认列和为 1，再分离特征值 1 的稳态方向；Fourier 题把系数逐个算成正交投影，不讨论无限级数收敛。

■ 例 R2.8　Final Review Summary 第 3 题改编：三状态 Markov 稳态　令

$$A=\begin{bmatrix}.2 & .4 & .3\\ .4 & .2 & .3\\ .4 & .4 & .4\end{bmatrix},\qquad u_0=(0,10,0)^{\mathsf{T}}.$$

求 $\lim_{k\to\infty}A^k u_0$。

解　各列和为 1，且所有元素为正。特征多项式为 $\lambda(\lambda-1)(\lambda+.2)$。对应特征值 $1,0,-.2$ 的特征向量可取 $(3,3,4)^{\mathsf{T}},(1,1,-2)^{\mathsf{T}},(1,-1,0)^{\mathsf{T}}$。初值分解为

$$u_0=(3,3,4)^{\mathsf{T}}+2(1,1,-2)^{\mathsf{T}}-5(1,-1,0)^{\mathsf{T}}.$$

迭代一次后零特征值分量消失，之后 $u_k=(3,3,4)^{\mathsf{T}}-5(-.2)^k(1,-1,0)^{\mathsf{T}}$，故极限为 $(3,3,4)^{\mathsf{T}}$。总质量 10 保持不变，归一化稳态是 $(.3,.3,.4)^{\mathsf{T}}$。 ♦

点评　初值总质量决定稳态向量的倍数；严格正性保证除 1 外的特征值模都小于 1。只有列随机而有零元素时，迭代可能周期振荡。

■ 例 R2.9　第 23 讲题型改编：投影到前三个 Fourier 方向　在 $[0,2\pi]$ 上，把 $f(x)=x$ 投影到 $V_1=\operatorname{span}\{1,\cos x,\sin x\}$。

解　三函数两两正交，平方长度分别为 $2\pi,\pi,\pi$。因此

$$a_0=\frac{1}{2\pi}\int_0^{2\pi}x\,\mathrm{d}x=\pi,\quad a_1=\frac{1}{\pi}\int_0^{2\pi}x\cos x\,\mathrm{d}x=0,\quad b_1=\frac{1}{\pi}\int_0^{2\pi}x\sin x\,\mathrm{d}x=-2.$$

其中后两式分部积分即可得到。投影为 $p(x)=\pi-2\sin x$；余差 $x-p(x)$ 与 $1,\cos x,\sin x$ 的内积都为 0，所以它确为 V_1 中的最佳平方逼近。 ♦

点评　常数项的分母是 2π，正弦、余弦项的分母是 π；这是基函数长度不同造成的。

R2.4　常见错误

把任意幂等矩阵当成正交投影　错误说法：$P^2=P$ 就表示正交投影。反例：$Q=\begin{bmatrix}1&1\\0&0\end{bmatrix}$ 幂等但不对称。正确条件还需 $P^{\mathsf T}=P$。

只检查残差分量和　错误说法：直线拟合的残差和为零就够了。反例：$e=(1,-1,0)^{\mathsf T}$ 分量和为零，却不垂直于斜率列 $(-1,1,2)^{\mathsf T}$。正确条件是 $A^{\mathsf T}e=0$。

把行列式当成整个矩阵的线性函数　错误说法：$\det(A+B)=\det A+\det B$。取 $A=B=I_2$，左侧为 4，右侧为 2。正确规则是固定其他行时，对单独一行线性。

把重根当成足够多的特征向量　错误说法：二重特征值必然给出两条独立特征方向。矩阵 $\begin{bmatrix}2&1\\0&2\end{bmatrix}$ 只有一条特征直线，不能对角化。要实际计算特征空间维数。

逐元素计算矩阵指数　错误说法：e^{At} 是把 A 的每个元素分别取指数。非对角矩阵的幂会改变多个位置；应使用幂级数，或在可对角化时计算 $Se^{\Lambda t}S^{-1}$。

把所有 Markov 迭代都当成收敛　错误说法：列和为 1 就保证收敛。交换矩阵 $\begin{bmatrix}0&1\\1&0\end{bmatrix}$ 会让 $(1,0)^{\mathsf T}$ 与 $(0,1)^{\mathsf T}$ 周期交替。严格正性是本讲收敛定理的附加条件。

本讲小结

- 用 **正交补**定位零空间：$N(A)=C(A^{\mathsf T})^{\perp}$、$N(A^{\mathsf T})=C(A)^{\perp}$；最近点误差垂直于目标子空间。
- 用 **法方程**求拟合：$A^{\mathsf T}A\hat{x}=A^{\mathsf T}b$，且 $A^{\mathsf T}(b-A\hat{x})=0$；拟合向量唯一，系数未必唯一。
- 用 **QR**分解满列秩矩阵：$A=QR$、$Q^{\mathsf T}Q=I$，最小二乘化为 $R\hat{x}=Q^{\mathsf T}b$。
- 用 **行列式**检验可逆性和体积：$\det(AB)=\det A\det B$，体积因子为 $|\det A|$；余子式可作展开与 Cramer 分子。
- 用 **特征基**计算迭代：$A=S\Lambda S^{-1}\Rightarrow A^k=S\Lambda^k S^{-1}$，特征方向数决定能否对角化。
- 用 **矩阵指数**解 $u'=Au$：$u(t)=e^{At}u_0$；可对角化时，各模态按 $e^{\lambda_i t}$ 演化。
- 用 **稳态与有限投影**处理第 23 讲：严格正的列 Markov 矩阵有唯一概率稳态；有限 Fourier 系数由积分内积确定。

练习

R2.1 ★　设 $W=\operatorname{span}\{(1,1,0)^{\mathsf T},(1,0,1)^{\mathsf T}\}$。求 $W^{\perp}$ 的一组基，并将 $b=(2,1,0)^{\mathsf T}$ 投影到 W；验证误差垂直于 W。

R2.2 ★　对 $A=\begin{bmatrix}1&-1\\1&0\\1&1\end{bmatrix}$、$b=(0,2,1)^{\mathsf T}$，解法方程并求拟合向量和残差。

R2.3 ★★　对 $A=\begin{bmatrix}1&1\\1&0\end{bmatrix}$ 做 Gram–Schmidt，写出薄 QR 分解，并用它求 $\min_x\|Ax-(1,0)^{\mathsf T}\|$。

R2.4 ★　若 B 是三角矩阵，对角线为 $2,-1,3$，求 $\det B$、$\det(B^{\mathsf T}B)$，并判断 B 是否可逆。

R2.5 ★★　计算 $\det\begin{bmatrix}1&2&0\\3&4&0\\5&6&7\end{bmatrix}$，沿含零最多的一列作余子式展开；说明其中一个代数余子式。

R2.6 ★★　用 Cramer 法则解 $\begin{bmatrix}3&2\\1&1\end{bmatrix}x=(7,3)^{\mathsf T}$，并求该矩阵所张成的平行四边形面积。

R2.7 ★★　求 $C=\begin{bmatrix}2&1\\0&2\end{bmatrix}$ 的特征值和每个特征空间维数；判断 C 是否可对角化。

R2.8 ★★　对 $D=\begin{bmatrix}2&1\\0&1\end{bmatrix}$ 求一组 S,Λ 使 $D=S\Lambda S^{-1}$，并计算 D^4。

R2.9 ★★　解系统 $u'=\begin{bmatrix}0&1\\-2&-3\end{bmatrix}u$、$u(0)=(1,0)^{\mathsf T}$，并说明所有解是否趋于零。

R2.10 ★★★ (a) 对 $M=\begin{bmatrix}3/5 & 9/10\\ 2/5 & 1/10\end{bmatrix}$ 求概率稳态并判断迭代收敛；(b) 把 $f(x)=2+3\cos x-4\sin 2x$ 投影到 $\operatorname{span}\{1,\cos x,\sin x,\cos 2x,\sin 2x\}$。

练习答案与提示

R2.1 取 $w_1=(1,1,0)$、$w_2=(1,0,1)$，解 $z^{\mathsf T}w_1=z^{\mathsf T}w_2=0$ 得 $W^{\perp}=\operatorname{span}\{(-1,1,1)\}$. 令 $Q=[w_1\ w_2]$，则 $Q^{\mathsf T}Q=\begin{bmatrix}2&1\\1&2\end{bmatrix}$、$Q^{\mathsf T}b=(3,2)^{\mathsf T}$，系数为 $(4/3,1/3)^{\mathsf T}$。故 $p=(5/3,4/3,1/3)^{\mathsf T}$、$e=b-p=(1/3,-1/3,-1/3)^{\mathsf T}$，且 $w_1^{\mathsf T}e=w_2^{\mathsf T}e=0$。

R2.2 $A^{\mathsf T}A=\operatorname{diag}(3,2)$、$A^{\mathsf T}b=(3,1)^{\mathsf T}$，所以 $\hat x=(1,1/2)^{\mathsf T}$。拟合向量 $A\hat x=(1/2,1,3/2)^{\mathsf T}$，残差 $e=(-1/2,1,-1/2)^{\mathsf T}$；直接算得 $A^{\mathsf T}e=(0,0)^{\mathsf T}$。

R2.3 $q_1=(1,1)^{\mathsf T}/\sqrt2$；第二列减去投影后为 $(1/2,-1/2)^{\mathsf T}$，故 $q_2=(1,-1)^{\mathsf T}/\sqrt2$。于是 $Q=\frac{1}{\sqrt2}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$、$R=\begin{bmatrix}\sqrt2 & 1/\sqrt2\\ 0 & 1/\sqrt2\end{bmatrix}$，且 $QR=A$。这里 $Q^{\mathsf T}(1,0)^{\mathsf T}=(1,1)^{\mathsf T}/\sqrt2$，解 $Rx=Q^{\mathsf T}b$ 得 $x=(0,1)^{\mathsf T}$；拟合为 $(1,0)^{\mathsf T}$，最小误差范数为 0。

R2.4 三角矩阵行列式为对角元之积，故 $\det B=-6$，B 可逆。由 $\det(B^{\mathsf T}B)=\det(B^{\mathsf T})\det B=(\det B)^2=36$。

R2.5 沿第三列展开，只有 $a_{33}=7$ 非零，$C_{33}=\det\begin{bmatrix}1&2\\3&4\end{bmatrix}=4-6=-2$，所以原行列式为 $7C_{33}=-14$。

R2.6 $\det A=3-2=1$。替换第一列、第二列所得行列式分别为 $1,2$，故 $x=(1,2)^{\mathsf T}$；代回得 $(7,3)^{\mathsf T}$。平行四边形面积为 $|\det A|=1$。

R2.7 特征多项式 $(2-\lambda)^2$，唯一特征值为 2。$(C-2I)x=0$ 化为 $x_2=0$，特征空间维数为 1，小于矩阵阶数 2，所以 C 不可对角化。

R2.8 $\lambda=2$ 可取特征向量 $(1,0)^{\mathsf T}$，$\lambda=1$ 可取 $(1,-1)^{\mathsf T}$。令 $S=\begin{bmatrix}1&1\\0&-1\end{bmatrix}$、$\Lambda=\operatorname{diag}(2,1)$，则 $D=S\Lambda S^{-1}$。于是 $D^4=\begin{bmatrix}16&15\\0&1\end{bmatrix}$。

R2.9 A 的特征值为 $-1,-2$，特征向量可取 $(1,-1)^{\mathsf T},(1,-2)^{\mathsf T}$。初值 $(1,0)^{\mathsf T}$ 等于 $2(1,-1)^{\mathsf T}-(1,-2)^{\mathsf T}$，所以

$$u(t)=2e^{-t}(1,-1)^{\mathsf T}-e^{-2t}(1,-2)^{\mathsf T}.$$

代入 $t=0$ 得初值；两个特征值都负，故所有模态衰减，$u(t)\to 0$。

R2.10 (a) 特征值为 1 与 $-3/10$，概率稳态为 $(9/13,4/13)^{\mathsf T}$，因为 $M(9,4)^{\mathsf T}=(9,4)^{\mathsf T}$。另一个特征值模小于 1，所以任意概率初值都收敛到该稳态。(b) 三角函数正交，函数已在所给有限维空间中，投影就是它本身；系数为 $a_0=2,a_1=3,a_2=0,b_1=0,b_2=-4$。

第三部分

正定矩阵、奇异值分解与线性变换

第 24 讲

对称矩阵与正定性

Symmetric Matrices and Positive Definiteness

对应材料：MIT 18.06SC Session 3.1 课时摘要与 Strang 第 25 讲视频；习题 25.1、25.2（Strang 教材 6.4 节第 10 题、6.5 节第 32 题）及解答；参考 Strang《Introduction to Linear Algebra》6.4–6.5 节。

本讲目标

(1) 会用「共轭」技巧证明实对称矩阵的特征值都是实数，并说明不同特征值的特征向量互相垂直。

(2) 理解谱定理 $A = Q\Lambda Q^{\mathsf{T}}$：实对称矩阵恰好是能被正交矩阵对角化的实矩阵；特征值重复时会用 Gram–Schmidt 挑出标准正交的特征向量。

(3) 会把对称矩阵写成投影矩阵的组合 $A = \sum \lambda_k q_k q_k^{\mathsf{T}}$，并用它计算 A^k、A^{-1}。

(4) 知道对称矩阵的正主元个数等于正特征值个数，会对 $A - bI$ 消元来估计特征值的位置。

(5) 掌握正定矩阵的定义和三种等价判别法：特征值全正、主元全正、顺序主子式全正。

24.1 对称矩阵有什么特别

上一讲的 Markov 矩阵有一个特殊的特征值 $\lambda = 1$。这说明：矩阵有特殊结构，特征值和特征向量往往也有特殊性质。本讲研究应用中最常见的一类特殊矩阵：**对称矩阵**（symmetric matrix），即满足 $A^{\mathsf{T}} = A$ 的方阵。最小二乘里的 $A^{\mathsf{T}}A$、力学里的刚度矩阵、统计里的协方差矩阵都是对称矩阵。

先看三个 2×2 矩阵，比较它们的特征值和特征向量。

例 24.1　三个矩阵的对比　求下面三个矩阵的特征值与特征向量，并观察特征向量之间的夹角：

$$S = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \qquad R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

解　对称矩阵 S：$\det(S - \lambda I) = (2-\lambda)^2 - 1 = (\lambda - 3)(\lambda - 1)$，特征值 3 和 1。直接验证

$$S \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad S \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

两个特征向量的点积 $(1,1) \cdot (1,-1) = 0$，**互相垂直**。

上三角矩阵 B：特征值就是对角元 1 和 3。$\lambda = 1$ 时 $B - I = \left[\begin{smallmatrix} 0 & 2 \\ 0 & 2 \end{smallmatrix}\right]$，特征向量 $(1,0)$；$\lambda = 3$ 时 $B - 3I = \left[\begin{smallmatrix} -2 & 2 \\ 0 & 0 \end{smallmatrix}\right]$，特征向量 $(1,1)$。二者点积为 $1 \neq 0$，夹角是 $45°$，不垂直。

旋转矩阵 R 把平面上每个向量转 90°，没有哪个非零实向量转完之后还和自己共线，所以 R 没有实特征向量。代数上，$\det(R-\lambda I)=\lambda^2+1=0$，特征值 $\lambda=\pm i$。取 $x=(i,1)$，

$$Rx=\begin{bmatrix}-1\\ i\end{bmatrix}=i\begin{bmatrix}i\\ 1\end{bmatrix}=ix,$$

所以 $x=(i,1)$ 是 $\lambda=i$ 的特征向量；同理 $(-i,1)$ 对应 $\lambda=-i$。 ♦

表 24.1 汇总了这三个矩阵。只有对称矩阵 S 同时具备两条好性质。

矩阵	特征值	特征向量	特征向量是否垂直
S（对称）	$3,\ 1$（实数）	$(1,1),\ (1,-1)$	垂直
B（上三角）	$1,\ 3$（实数）	$(1,0),\ (1,1)$	夹角 45°
R（旋转）	$i,\ -i$（复数）	$(i,1),\ (-i,1)$	只有复向量

表 24.1　三个 2×2 矩阵的特征值与特征向量。

本讲的核心结论是：**每个实对称矩阵都同时具备这两条性质**。

(1) 特征值全是实数。

(2) 可以选出 n 个互相垂直的特征向量；除以长度后它们构成 $\mathbb{R}^n$ 的一组标准正交基。

复向量的长度

要证明「特征值是实数」，必须先承认它可能是复数，再证明虚部为零。这就要和复向量打交道。对实向量，$x^{\mathsf{T}}x=x_1^2+\cdots+x_n^2$ 是长度的平方，只有 $x=0$ 时才等于 0。这条性质对复向量不成立。例 24.1 中 R 的特征向量 $x=(i,1)$ 就是反例：

$$x^{\mathsf{T}}x=i^2+1^2=-1+1=0,$$

一个非零向量的「长度平方」竟然是 0。补救的办法和复数取模一样：$|z|^2=\bar{z}z$，而不是 z^2。对复向量，用共轭 $\bar{x}$ 去乘：

$$\boxed{\bar{x}^{\mathsf{T}}x=\bar{x}_1x_1+\cdots+\bar{x}_nx_n=|x_1|^2+\cdots+|x_n|^2.}$$

每一项都非负，所以 $x\neq 0$ 时 $\bar{x}^{\mathsf{T}}x>0$。对 $x=(i,1)$，$\bar{x}^{\mathsf{T}}x=(-i)(i)+1\cdot 1=2$，长度是 $\sqrt{2}$。

易错提醒　对复向量，$x^{\mathsf{T}}x$ 可以是 0、负数甚至非实数，不能当作长度的平方。凡是需要「非零向量长度为正」的地方，都要改用 $\bar{x}^{\mathsf{T}}x$。第 25 讲会把 $\bar{x}^{\mathsf{T}}$ 记成 x^{H}（共轭转置），并系统地讲复矩阵。

24.2　实特征值与正交特征向量

> **定理 24.1（对称矩阵的特征值是实数）**　若 A 是实对称矩阵，则 A 的每个特征值都是实数，并且每个特征值都有实的特征向量。

思路　设 $Ax=\lambda x$，其中 λ 和 x 暂时允许是复的。把数 $\bar{x}^{\mathsf{T}}Ax$ 算两遍：一次用 $Ax=\lambda x$，得到 $\lambda\,\bar{x}^{\mathsf{T}}x$；另一次先对方程取共轭、再转置，并用 $A^{\mathsf{T}}=A$，得到 $\bar{\lambda}\,\bar{x}^{\mathsf{T}}x$。因为 $\bar{x}^{\mathsf{T}}x>0$，两式相除就有 $\lambda=\bar{\lambda}$。

证明　设 $Ax=\lambda x$，$x\neq 0$。对方程两边取复共轭。A 的元素是实数，$\bar{A}=A$，所以

$$A\bar{x}=\bar{\lambda}\,\bar{x}.$$

（顺带说明：实矩阵的复特征值总是共轭成对出现。）再转置，并用 $A^{\mathsf{T}}=A$：

$$\bar{x}^{\mathsf{T}}A=\bar{\lambda}\,\bar{x}^{\mathsf{T}}.$$

两边右乘 x 得 $\bar{x}^{\mathsf{T}}Ax=\bar{\lambda}\,\bar{x}^{\mathsf{T}}x$。另一方面，原方程 $Ax=\lambda x$ 两边左乘 $\bar{x}^{\mathsf{T}}$ 得 $\bar{x}^{\mathsf{T}}Ax=\lambda\,\bar{x}^{\mathsf{T}}x$。两式左边相同，所以

$$\lambda\,\bar{x}^{\mathsf{T}}x=\bar{\lambda}\,\bar{x}^{\mathsf{T}}x.$$

由于 $x\neq 0$，$\bar{x}^{\mathsf{T}}x>0$，约去后得 $\lambda=\bar{\lambda}$，即 λ 是实数。

既然 λ 是实数，$A-\lambda I$ 是奇异的实矩阵，它的零空间里有非零实向量，这就是实的特征向量。 ■

证明里用到了两件事：$A^{\mathsf{T}}=A$（把 A 从右边挪到左边），以及 $\bar{x}^{\mathsf{T}}x>0$（才能约去）。缺了任何一条，结论都可能失败。下面的例子说明第二条为什么必须用 $\bar{x}^{\mathsf{T}}x$ 而不是 $x^{\mathsf{T}}x$。

■ **例 24.2　习题 25.1：一个错误的证明**　有人这样「证明」特征值是实数：由 $Ax=\lambda x$ 得 $x^{\mathsf{T}}Ax=\lambda x^{\mathsf{T}}x$，所以

$$\lambda=\frac{x^{\mathsf{T}}Ax}{x^{\mathsf{T}}x},$$

分子分母都是实数，因此 λ 是实数。用旋转矩阵 $R=\left[\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right]$、$\lambda=i$、$x=(i,1)$ 检验这个推理，找出其中的隐含假设。

解　先确认 $Rx=\lambda x$：例 24.1 已算出 $Rx=(-1,i)=i(i,1)$。再算分子分母：

$$x^{\mathsf{T}}Rx=\begin{bmatrix}i & 1\end{bmatrix}\begin{bmatrix}-1\\ i\end{bmatrix}=-i+i=0,\qquad x^{\mathsf{T}}x=i^2+1=0.$$

等式 $x^{\mathsf{T}}Rx=\lambda\,x^{\mathsf{T}}x$ 成立（两边都是 0），但不能除。「证明」的隐含假设是 $x^{\mathsf{T}}x\neq 0$，并且默认 $x^{\mathsf{T}}Ax$ 和 $x^{\mathsf{T}}x$ 是实数，这两点对复向量 x 都不成立。

改用 $\bar{x}^{\mathsf{T}}x$ 能否挽救？$\bar{x}^{\mathsf{T}}x=2\neq 0$，分母没问题了；但 $\bar{x}^{\mathsf{T}}Rx=\begin{bmatrix}-i & 1\end{bmatrix}\begin{bmatrix}-1\\ i\end{bmatrix}=2i$，于是 $\lambda=2i/2=i$，仍不是实数。原因是 R 不对称（$R^{\mathsf{T}}=-R$）：定理 24.1 的证明里，从 $A\bar{x}=\bar{\lambda}\bar{x}$ 转置得到 $\bar{x}^{\mathsf{T}}A=\bar{\lambda}\bar{x}^{\mathsf{T}}$ 这一步要用 $A^{\mathsf{T}}=A$，对 R 这一步给出的是 $\bar{x}^{\mathsf{T}}R=-\bar{\lambda}\bar{x}^{\mathsf{T}}$。 ♦

定理 24.2（不同特征值的特征向量互相垂直）　设 A 是实对称矩阵，$Ax=\lambda_1 x$，$Ay=\lambda_2 y$，且 $\lambda_1\neq\lambda_2$，则 $x^{\mathsf T}y=0$。

证明　由定理 24.1，λ_1,λ_2 是实数，x,y 可取为实向量。把数 $(Ax)^{\mathsf T}y$ 算两遍：

$$(Ax)^{\mathsf T}y=\lambda_1\,x^{\mathsf T}y,\qquad (Ax)^{\mathsf T}y=x^{\mathsf T}A^{\mathsf T}y=x^{\mathsf T}Ay=\lambda_2\,x^{\mathsf T}y.$$

所以 $(\lambda_1-\lambda_2)\,x^{\mathsf T}y=0$。因为 $\lambda_1\neq\lambda_2$，只能 $x^{\mathsf T}y=0$。 ■

关键一步是 $(Ax)^{\mathsf T}y=x^{\mathsf T}(Ay)$：对称矩阵可以在点积的两边「搬来搬去」。例 24.1 的 B 不对称，$(Bx)^{\mathsf T}y=x^{\mathsf T}(B^{\mathsf T}y)$ 里出现的是 $B^{\mathsf T}$ 而不是 B，所以 B 的特征向量不垂直。

特征值重复时，定理 24.2 管不到。这时同一个特征值的特征向量构成一个子空间（特征空间），里面的向量不一定垂直，但我们总可以在里面挑出互相垂直的向量。

■ **例 24.3　重特征值：用 Gram–Schmidt 挑选**　设 J 是 3×3 全 1 矩阵，$A=I+J=\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}$。求 A 的三个两两垂直的单位特征向量。

解　J 的秩为 1，所以 0 是 J 的二重特征值；另一个特征值等于迹 3，特征向量是 $(1,1,1)$。加上 I 后特征值都加 1：A 的特征值是 $4,1,1$。核对：迹 $4+1+1=6=2+2+2$；$\det A=4\cdot1\cdot1=4$。

$\lambda_1=4$：取 $q_1=\frac{1}{\sqrt3}(1,1,1)$。

$\lambda=1$：$A-I=J$，特征空间是 $Jx=0$ 的解，即平面 $x_1+x_2+x_3=0$。这个平面内可取两个向量 $v=(1,-1,0)$ 和 $w=(1,0,-1)$，但 $v^{\mathsf T}w=1\neq0$，它们不垂直。用 Gram–Schmidt：

$$w-\frac{v^{\mathsf T}w}{v^{\mathsf T}v}\,v=(1,0,-1)-\tfrac12(1,-1,0)=\left(\tfrac12,\tfrac12,-1\right),$$

乘以 2 得 $(1,1,-2)$，它仍在平面 $x_1+x_2+x_3=0$ 里，所以仍是特征向量。单位化：

$$q_1=\frac{1}{\sqrt3}\begin{bmatrix}1\\1\\1\end{bmatrix},\qquad q_2=\frac{1}{\sqrt2}\begin{bmatrix}1\\-1\\0\end{bmatrix},\qquad q_3=\frac{1}{\sqrt6}\begin{bmatrix}1\\1\\-2\end{bmatrix}.$$

两两点积：$q_1^{\mathsf T}q_2\propto1-1+0=0$，$q_1^{\mathsf T}q_3\propto1+1-2=0$，$q_2^{\mathsf T}q_3\propto1-1+0=0$。$q_2,q_3$ 与 q_1 垂直是定理 24.2 保证的；q_2 与 q_3 垂直是我们挑出来的。 ♦

想一想　「对称矩阵的任意两个特征向量都垂直」这句话对吗？

答　不对。例 24.3 中 $v=(1,-1,0)$ 和 $w=(1,0,-1)$ 都是 $\lambda=1$ 的特征向量，却不垂直。最极端的例子是 I：每个非零向量都是它的特征向量。正确的说法是：不同特征值的特征向量一定垂直；同一特征值的特征空间里可以选出一组标准正交基。

注 24.1　如果 A 的元素可以是复数，「对称」的正确推广不是 $A^{\mathsf T}=A$，而是 $\bar A^{\mathsf T}=A$，这样的矩阵叫 **Hermite 矩阵**（Hermitian matrix）。把定理 24.1 的证明里的 $\bar A=A$ 换成 $\bar A^{\mathsf T}=A$，同样的两行推导说明 Hermite 矩阵的特征值也是实数。详细内容留到第 25 讲。

■ **例 24.4　一个 Hermite 矩阵**　求 $H=\begin{bmatrix}2 & 3-3i\\ 3+3i & 5\end{bmatrix}$ 的特征值和特征向量。

解　H 不是实对称矩阵，但 $\bar{H}^{\mathsf{T}}=H$：对角元是实数，两个非对角元互为共轭。$\det(H-\lambda I)=(2-\lambda)(5-\lambda)-|3-3i|^2=\lambda^2-7\lambda+10-18=(\lambda-8)(\lambda+1)$，特征值 8 和 -1，都是实数。

$\lambda=8$：$H-8I$ 的第一行是 $(-6,\ 3-3i)$，取 $x=(1,\ 1+i)$，因为 $-6+(3-3i)(1+i)=-6+6=0$。$\lambda=-1$：$H+I$ 的第一行是 $(3,\ 3-3i)$，取 $y=(1-i,\ -1)$，因为 $3(1-i)-(3-3i)=0$。两个特征向量在共轭意义下垂直：

$$\bar{x}^{\mathsf{T}}y=1\cdot(1-i)+(1-i)\cdot(-1)=0.$$

♦

24.3　谱定理：用正交矩阵对角化

第 21 讲中，有 n 个线性无关特征向量的矩阵可以对角化：$A=S\Lambda S^{-1}$，S 的列是特征向量。对称矩阵的特征向量可以选成标准正交的，于是 S 变成正交矩阵 Q，而正交矩阵的逆就是转置（第 16 讲）：$Q^{-1}=Q^{\mathsf{T}}$。

> **定理 24.3（谱定理）**　每个实对称矩阵 A 都可以写成
>
> $$A=Q\Lambda Q^{\mathsf{T}},$$
>
> 其中 Q 是正交矩阵（$Q^{\mathsf{T}}Q=I$），Λ 是实对角矩阵。Λ 的对角元是 A 的特征值，Q 的第 k 列 q_k 是 λ_k 的单位特征向量。反过来，任何形如 $Q\Lambda Q^{\mathsf{T}}$（Q 正交，Λ 实对角）的矩阵都是对称的。

反方向一行就能验证：$(Q\Lambda Q^{\mathsf{T}})^{\mathsf{T}}=(Q^{\mathsf{T}})^{\mathsf{T}}\Lambda^{\mathsf{T}}Q^{\mathsf{T}}=Q\Lambda Q^{\mathsf{T}}$。所以对实矩阵来说，「对称」与「能用正交矩阵对角化」是一回事。数学家把特征值的集合叫做矩阵的**谱**（spectrum），定理因此得名；在力学里它叫**主轴定理**（principal axis theorem）。

思路　特征值互不相同时，定理 24.2 直接给出 n 个互相垂直的特征向量，单位化后排成 Q 即可。难点在于重特征值：要保证二重特征值确实有二维的特征空间，不会像第 21 讲中 $\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]$ 那样「缺」特征向量。办法是对 n 归纳：先用定理 24.1 找到一个实的单位特征向量 q_1，把它扩充成 $\mathbb{R}^n$ 的标准正交基；在这组基下，A 的第一列变成 $(\lambda_1,0,\dots,0)$，对称性让第一行也变成这样，剩下一个 $n-1$ 阶对称矩阵交给归纳假设。

证明　只需证明正方向。对阶数 n 归纳。$n=1$ 时 $A=[a]$，取 $Q=[1]$，$\Lambda=[a]$。

设结论对 $n-1$ 阶实对称矩阵成立，A 是 n 阶实对称矩阵。由定理 24.1，A 有实特征值 λ_1 和实的单位特征向量 q_1。用 Gram–Schmidt 把 q_1 扩充成 $\mathbb{R}^n$ 的标准正交基 $q_1,u_2,\dots,u_n$，排成正交矩阵 $Q_1=\begin{bmatrix}q_1 & u_2 & \cdots & u_n\end{bmatrix}$。考察

$$B=Q_1^{\mathsf{T}}AQ_1.$$

B 的第一列是 $Q_1^{\mathsf{T}}Aq_1=\lambda_1Q_1^{\mathsf{T}}q_1=\lambda_1(1,0,\dots,0)$，因为 q_1 与自己的点积为 1、与 u_j 的点积为 0。

又 $B^{\mathsf{T}} = Q_1^{\mathsf{T}} A^{\mathsf{T}} Q_1 = B$，$B$ 是对称的，所以第一行与第一列相同。于是

$$B = \begin{bmatrix} \lambda_1 & 0 \\ 0 & A' \end{bmatrix},$$

其中 A' 是 $n-1$ 阶实对称矩阵。由归纳假设，$A' = Q'\Lambda' Q'^{\mathsf{T}}$，$Q'$ 正交。令

$$Q = Q_1 \begin{bmatrix} 1 & 0 \\ 0 & Q' \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \Lambda' \end{bmatrix}.$$

两个正交矩阵的乘积仍是正交矩阵。分块计算

$$Q^{\mathsf{T}} A Q = \begin{bmatrix} 1 & 0 \\ 0 & Q'^{\mathsf{T}} \end{bmatrix} B \begin{bmatrix} 1 & 0 \\ 0 & Q' \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & Q'^{\mathsf{T}} A' Q' \end{bmatrix} = \Lambda,$$

即 $A = Q\Lambda Q^{\mathsf{T}}$。由 $AQ = Q\Lambda$ 逐列比较，Q 的每一列都是 A 的单位特征向量，对应的特征值是 Λ 的对角元。 ■

证明的关键在于「第一行也是零」。对一般的矩阵做同样的事，只能让第一列变成 $(\lambda_1, 0, \dots, 0)$，最后得到上三角矩阵（这叫 Schur 分解）；是对称性把上三角变成了对角。

■ **例 24.5**　2×2 **的谱分解**　把 $S = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$ 写成 $Q\Lambda Q^{\mathsf{T}}$ 并验证。

解　由例 24.1，特征值 $3, 1$，特征向量 $(1,1)$ 和 $(1,-1)$。第二个向量换号为 $(-1,1)$，使 Q 是一个旋转（$\det Q = 1$）：

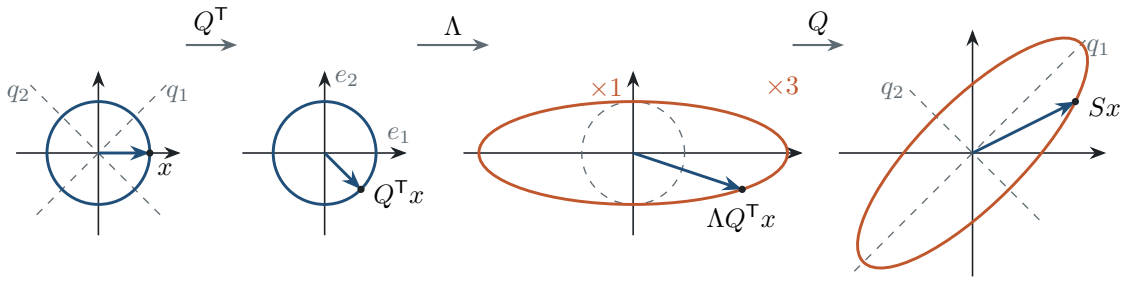
$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

Q 是逆时针旋转 $45°$。验证：

$$Q\Lambda Q^{\mathsf{T}} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3 & -1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} = S.$$

♦

直观理解　从右往左读 $Sx = Q\Lambda Q^{\mathsf{T}} x$。$Q^{\mathsf{T}}$ 把 x 转到「特征向量坐标系」里（q_1, q_2 变成坐标轴 e_1, e_2），Λ 沿两条坐标轴分别拉伸 3 倍和 1 倍，Q 再转回来。所以**对称矩阵的作用就是沿 n 个互相垂直的方向各自伸缩**：单位圆变成一个椭圆，椭圆的主轴是特征向量，半轴长是 $|\lambda_k|$（图 24.1）。一般的 $S\Lambda S^{-1}$ 也是「换坐标、伸缩、换回来」，但坐标轴是斜的。



(a) 原图　(b) 转 $-45°$　(c) 伸缩　(d) 转回 $45°$

图 24.1　$S = Q\Lambda Q^{\mathsf{T}}$ 的几何意义（$S = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$）。单位圆先转 $-45°$，再沿坐标轴拉伸 3 倍和 1 倍，最后转回，变成主轴沿 q_1, q_2 的椭圆。黑点 $x = (1, 0)$ 依次变为 $\frac{1}{\sqrt{2}}(1, -1)$、$\frac{1}{\sqrt{2}}(3, -1)$、$Sx = (2, 1)$。

历史与背景　把二次曲面转到主轴上，是 Euler 和 Lagrange 在 18 世纪研究刚体转动时遇到的问题：刚体的惯性矩阵是对称的，它的特征向量就是转动的主轴。Cauchy 在 1829 年对一般的 n 证明了对称矩阵的特征值都是实数。Sylvester 在 1852 年证明了惯性定律：A 换成 $C^{\mathsf{T}}AC$（C 可逆）时，正、负特征值的个数不变，这是 §24.5 主元符号定理的背景。Hermite 在 1855 年把结论推广到复的 Hermite 矩阵。「谱」这个词来自 Hilbert 在二十世纪初关于积分方程的工作；后来量子力学发现原子光谱的谱线恰好对应某个 Hermite 算子的特征值，这个名字变得名副其实。

想一想　若实矩阵 A 有 n 个两两垂直的实特征向量，A 一定对称吗？

答　一定。实矩阵乘实向量还是实向量，所以 $Ax = \lambda x$ 中的 λ 是实数。把这些向量单位化排成 Q，则 Q 正交，$AQ = Q\Lambda$，于是 $A = Q\Lambda Q^{-1} = Q\Lambda Q^{\mathsf{T}}$，由定理 24.3 的反方向，$A$ 对称。例 24.1 的 B 有实特征值、也能对角化，但特征向量夹角为 $45°$，这正是它不对称的表现。

24.4　谱分解：对称矩阵是投影的组合

把 $Q\Lambda Q^{\mathsf{T}}$ 按第 3 讲「列乘行」的方式展开：$Q\Lambda$ 的列是 $\lambda_k q_k$，Q^{T} 的行是 q_k^{T}，所以

$$A = Q\Lambda Q^{\mathsf{T}} = \begin{bmatrix} \lambda_1 q_1 & \cdots & \lambda_n q_n \end{bmatrix} \begin{bmatrix} q_1^{\mathsf{T}} \\ \vdots \\ q_n^{\mathsf{T}} \end{bmatrix},$$

即

$$\boxed{A = \lambda_1 q_1 q_1^{\mathsf{T}} + \lambda_2 q_2 q_2^{\mathsf{T}} + \cdots + \lambda_n q_n q_n^{\mathsf{T}}.}$$

这叫 A 的**谱分解**（spectral decomposition）。每一项里的 $P_k = q_k q_k^{\mathsf{T}}$ 是秩为 1 的矩阵。因为 q_k 是单位向量，第 14 讲的投影公式 $P = \dfrac{aa^{\mathsf{T}}}{a^{\mathsf{T}}a}$ 告诉我们，P_k 正是**投影到 q_k 方向上的投影矩阵**。由 $q_i^{\mathsf{T}} q_j = 0$（$i \neq j$）、$q_k^{\mathsf{T}} q_k = 1$ 得到：

(1) $P_k^2 = q_k (q_k^{\mathsf{T}} q_k) q_k^{\mathsf{T}} = P_k$，$P_k^{\mathsf{T}} = P_k$：每个 P_k 都是投影矩阵。

(2) $P_i P_j = q_i (q_i^{\mathsf{T}} q_j) q_j^{\mathsf{T}} = 0$（$i \neq j$）：先投影到 q_j，再投影到与之垂直的 q_i，结果为零。

(3) $P_1+\cdots+P_n=QQ^{\mathsf T}=I$：把向量在 n 个垂直方向上的分量加起来，就还原了这个向量。

核心思想　每个对称矩阵都是互相垂直的投影的组合，组合系数是特征值。计算 Ax 时，先把 x 投影到每个特征方向上，再把各个分量乘以对应的特征值，最后相加：

$$Ax=\lambda_1(q_1^{\mathsf T}x)\,q_1+\cdots+\lambda_n(q_n^{\mathsf T}x)\,q_n.$$

由性质 1 和 2，$A^2=\sum_{i,j}\lambda_i\lambda_jP_iP_j=\sum_k\lambda_k^2P_k$。同理，投影矩阵不变，只是系数换成特征值的函数：

$$A^k=\lambda_1^kP_1+\cdots+\lambda_n^kP_n,\qquad A^{-1}=\frac{1}{\lambda_1}P_1+\cdots+\frac{1}{\lambda_n}P_n\quad(\text{各}\lambda_k\neq0).$$

这和 $A^k=Q\Lambda^kQ^{\mathsf T}$ 是同一件事。

■ **例 24.6**　2×2：两个投影　写出 $S=\left[\begin{smallmatrix}2&1\\1&2\end{smallmatrix}\right]$ 的谱分解，求 S^k，并对 $x=(1,0)$ 画出 Sx 的分解。

解　$q_1=\frac{1}{\sqrt2}(1,1)$，$q_2=\frac{1}{\sqrt2}(-1,1)$，

$$P_1=q_1q_1^{\mathsf T}=\frac12\begin{bmatrix}1&1\\1&1\end{bmatrix},\qquad P_2=q_2q_2^{\mathsf T}=\frac12\begin{bmatrix}1&-1\\-1&1\end{bmatrix},\qquad S=3P_1+1\cdot P_2.$$

核对：$3P_1+P_2=\frac12\left[\begin{smallmatrix}4&2\\2&4\end{smallmatrix}\right]=S$，$P_1+P_2=I$。于是

$$S^k=3^kP_1+P_2=\frac12\begin{bmatrix}3^k+1&3^k-1\\3^k-1&3^k+1\end{bmatrix}.$$

$k=2$ 时得 $\left[\begin{smallmatrix}5&4\\4&5\end{smallmatrix}\right]$，与直接计算 S^2 相同；$k=-1$ 时得 $S^{-1}=\frac13\left[\begin{smallmatrix}2&-1\\-1&2\end{smallmatrix}\right]$。

对 $x=(1,0)$：$P_1x=(\frac12,\frac12)$，$P_2x=(\frac12,-\frac12)$，二者之和是 x。乘以特征值后相加，$Sx=3P_1x+P_2x=(\frac32,\frac32)+(\frac12,-\frac12)=(2,1)$，见图 24.2。 ♦

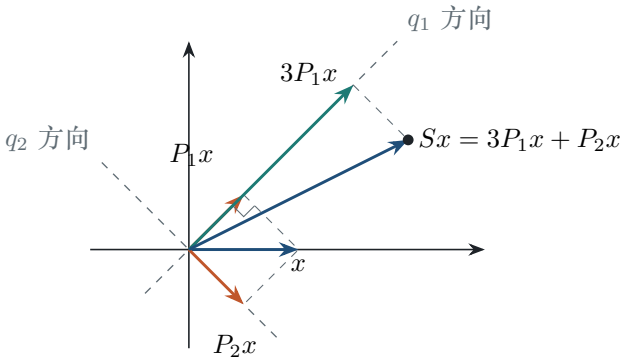


图 24.2　谱分解的几何意义。x 在两个垂直方向上的投影 P_1x,P_2x 分别乘以特征值 3 和 1，再相加得到 Sx。

■ **例 24.7**　3×3：不必求出 q_2,q_3　对例 24.3 的 $A=I+J$，求 A^{-1}、A^5，并解 $Ax=(1,2,3)$。

解　$\lambda_1=4$ 对应 $P_1=q_1q_1^{\mathsf T}=\frac{1}{3}J$。$\lambda=1$ 对应两个投影 P_2+P_3，而由 $P_1+P_2+P_3=I$，不必求出 q_2,q_3：

$$P_2+P_3=I-\tfrac{1}{3}J,\qquad A=4\cdot\tfrac{1}{3}J+1\cdot\left(I-\tfrac{1}{3}J\right).$$

P_2+P_3 是投影到平面 $x_1+x_2+x_3=0$ 上的投影矩阵。于是

$$A^{-1}=\tfrac{1}{4}\cdot\tfrac{1}{3}J+\left(I-\tfrac{1}{3}J\right)=I-\tfrac{1}{4}J=\frac{1}{4}\begin{bmatrix}3&-1&-1\\-1&3&-1\\-1&-1&3\end{bmatrix},\qquad A^5=\tfrac{4^5}{3}J+I-\tfrac{1}{3}J=I+341J.$$

核对：$(I+J)(I-\frac{1}{4}J)=I+\frac{3}{4}J-\frac{1}{4}J^2=I$，因为 $J^2=3J$；A^5 的对角元 342，非对角元 341。最后，$x=A^{-1}b=b-\frac{1}{4}Jb$，$Jb=(6,6,6)$，所以 $x=(1,2,3)-(\frac{3}{2},\frac{3}{2},\frac{3}{2})=(-\frac{1}{2},\frac{1}{2},\frac{3}{2})$。代回：$Ax=(-1+\frac{1}{2}+\frac{3}{2},\ -\frac{1}{2}+1+\frac{3}{2},\ -\frac{1}{2}+\frac{1}{2}+3)=(1,2,3)$。 ♦

谱分解还给出一个估计，§24.6 证明正定判别法时要用到它。

> **引理 24.4（二次型的上下界）**　设 A 是实对称矩阵，最小、最大特征值分别是 $\lambda_{\min}$、$\lambda_{\max}$。则对每个实向量 x，
>
> $$\lambda_{\min}\,x^{\mathsf T}x\ \le\ x^{\mathsf T}Ax\ \le\ \lambda_{\max}\,x^{\mathsf T}x.$$

证明　把 x 在标准正交特征向量下展开：$x=c_1q_1+\cdots+c_nq_n$，$c_k=q_k^{\mathsf T}x$。由正交性，$x^{\mathsf T}x=c_1^2+\cdots+c_n^2$，而 $Ax=\sum\lambda_kc_kq_k$，所以 $x^{\mathsf T}Ax=\lambda_1c_1^2+\cdots+\lambda_nc_n^2$。把每个 λ_k 换成 $\lambda_{\min}$ 或 $\lambda_{\max}$ 即得两个不等式。 ■

特别地，若 A 的特征值全为正，则 $x\neq0$ 时 $x^{\mathsf T}Ax\ge\lambda_{\min}x^{\mathsf T}x>0$。数 $x^{\mathsf T}Ax$ 是第 26 讲的主角，这里只把它当作证明工具。

24.5　主元的符号与特征值的符号

特征值的符号在应用中很要紧。第 22 讲中，$\dfrac{du}{dt}=Au$ 的解由 e^{λ_kt} 组合而成：所有 $\lambda_k<0$ 时解衰减到零（稳定），若有 $\lambda_k>0$，则存在沿该特征向量指数增长的解；初值在该特征方向的分量非零时，解才含有指数增长项。可是 n 上千时，解 $\det(A-\lambda I)=0$ 根本不现实，而消元很便宜。对称矩阵有一个好性质：**主元的符号与特征值的符号一致**。

先回忆对称矩阵的消元（第 4 讲）。若消元不需要换行，$A=LDU$，L 是对角元为 1 的下三角矩阵，D 是主元组成的对角矩阵，U 是对角元为 1 的上三角矩阵。A 对称时 $A=A^{\mathsf T}=U^{\mathsf T}DL^{\mathsf T}$，由分解的唯一性 $U=L^{\mathsf T}$，所以

$$A=LDL^{\mathsf T}.$$

■ **例 24.8　主元与特征值**　比较 $A_1=\left[\begin{smallmatrix}1&3\\3&1\end{smallmatrix}\right]$ 与 $K=\left[\begin{smallmatrix}2&-1&0\\-1&2&-1\\0&-1&2\end{smallmatrix}\right]$ 的主元和特征值。

解　A_1：第二行减去第一行的 3 倍，第二个主元是 $1-9=-8$。主元 $1,-8$，一正一负。特征值：$(1-\lambda)^2=9$，$\lambda=4$（特征向量 $(1,1)$）和 $\lambda=-2$（特征向量 $(1,-1)$），也是一正一负。

K：第二行加第一行的 $\frac{1}{2}$ 倍，得第二个主元 $2-\frac{1}{2}=\frac{3}{2}$；第三行加第二行的 $\frac{2}{3}$ 倍，得第三个主元 $2-\frac{2}{3}=\frac{4}{3}$。主元 $2,\frac{3}{2},\frac{4}{3}$ 全正。特征值是 $2-\sqrt{2}\approx 0.586$、2、$2+\sqrt{2}\approx 3.414$，特征向量分别为 $(1,\sqrt{2},1)$、$(1,0,-1)$、$(1,-\sqrt{2},1)$，也全正。

两个例子里，主元与特征值的数值都不同，但正负个数相同，乘积也相同：$1\cdot(-8)=4\cdot(-2)=-8$，$2\cdot\frac{3}{2}\cdot\frac{4}{3}=4=(2-\sqrt{2})\cdot 2\cdot(2+\sqrt{2})$，都等于 $\det$。 ♦

定理 24.5（主元符号定理）　设 A 是实对称矩阵，消元不需要换行，且 n 个主元都不为零。则 A 的正主元个数等于正特征值个数，负主元个数等于负特征值个数。

思路　$A=LDL^{\mathsf{T}}$ 中，D 的特征值就是主元。让 L 的下三角部分连续地缩到零，矩阵就从 A 连续地变到 D。整个过程中矩阵一直对称，特征值一直是实数；行列式一直等于主元之积，不为零，所以没有特征值能经过 0。一个实数从正变负必须经过 0，因此每个特征值的符号在变形中保持不变。

证明　设 $A=LDL^{\mathsf{T}}$。对 $0\le t\le 1$ 令 $L_t=I+t(L-I)$，它仍是对角元为 1 的下三角矩阵；$L_0=I$，$L_1=L$。令

$$A(t)=L_tDL_t^{\mathsf{T}}.$$

每个 $A(t)$ 都是实对称矩阵，所以它的 n 个特征值都是实数，记为 $\lambda_1(t)\le\cdots\le\lambda_n(t)$。又 $\det A(t)=\det L_t\cdot\det D\cdot\det L_t^{\mathsf{T}}=d_1d_2\cdots d_n\neq 0$，所以对每个 t，0 都不是 $A(t)$ 的特征值。

$A(t)$ 的元素是 t 的多项式，特征多项式的系数也随 t 连续变化，而多项式的根连续地依赖于系数（这一事实我们承认，严格证明属于分析课程）。所以每个 $\lambda_k(t)$ 是 t 的连续函数，且从不为零。由介值定理，$\lambda_k(t)$ 在 $[0,1]$ 上不变号。$t=0$ 时 $A(0)=D$，特征值就是主元；$t=1$ 时 $A(1)=A$。因此 A 的正（负）特征值个数等于 D 的正（负）特征值个数，即正（负）主元个数。 ■

证明里对称性用了两次：保证 $A=LDL^{\mathsf{T}}$ 的形式，以及保证 $A(t)$ 的特征值始终是实数。非对称矩阵的特征值可以离开实轴，从复平面上「绕过」0 而改变实部的符号，定理随之失效。

易错提醒　主元符号定理只对对称矩阵成立。$N=\left[\begin{smallmatrix}1&1\\-4&-3\end{smallmatrix}\right]$ 的主元是 1 和 $-3-(-4)\cdot 1=1$，都是正的；但 $\operatorname{tr}N=-2$，$\det N=1$，特征方程 $(\lambda+1)^2=0$，两个特征值都是 -1。另外，即使对称，主元也不等于特征值，二者只是符号个数相同、乘积相同。

平移：数出大于 b 的特征值

若 $A-bI$ 的消元无需换行且主元均非零，则可使用下面的计数。$A-bI$ 的特征值是 λ_k-b，而且 $A-bI$ 仍然对称。由定理 24.5：

$$A-bI\text{ 的正主元个数}\;=\;A\text{ 的大于 }b\text{ 的特征值个数}.$$

对几个不同的 b 消元，就能把每个特征值夹在一个区间里，完全不用解特征方程。数值计算中的二分法求特征值就基于这个想法。

■ **例 24.9　用平移定位特征值**　不解特征方程，用 $b=\frac{1}{2},\frac{3}{2},\frac{5}{2},\frac{7}{2}$ 定位例 24.8 中 K 的三个特征值。

解　$K-bI$ 的对角元是 $c=2-b$，次对角元是 -1。消元时第 $k+1$ 行加上第 k 行的 $1/d_k$ 倍，所以主元满足 $d_1=c$，$d_{k+1}=c-1/d_k$。例如 $b=\frac{3}{2}$ 时 $c=\frac{1}{2}$，$d_2=\frac{1}{2}-2=-\frac{3}{2}$，$d_3=\frac{1}{2}+\frac{2}{3}=\frac{7}{6}$。结果如下：

b	$K-bI$ 的主元	正主元个数	$\det(K-bI)$	结论
$\frac{1}{2}$	$\frac{3}{2},\ \frac{5}{6},\ \frac{3}{10}$	3	$\frac{3}{8}$	三个特征值都大于 $\frac{1}{2}$
$\frac{3}{2}$	$\frac{1}{2},\ -\frac{3}{2},\ \frac{7}{6}$	2	$-\frac{7}{8}$	两个大于 $\frac{3}{2}$
$\frac{5}{2}$	$-\frac{1}{2},\ \frac{3}{2},\ -\frac{7}{6}$	1	$\frac{7}{8}$	一个大于 $\frac{5}{2}$
$\frac{7}{2}$	$-\frac{3}{2},\ -\frac{5}{6},\ -\frac{3}{10}$	0	$-\frac{3}{8}$	没有大于 $\frac{7}{2}$ 的

所以三个特征值分别落在 $(\frac{1}{2},\frac{3}{2})$、$(\frac{3}{2},\frac{5}{2})$、$(\frac{5}{2},\frac{7}{2})$ 中，与精确值 0.586，2，3.414 相符（图 24.3）。每一行的主元乘积都等于 $\det(K-bI)$，可作核对。 ♦

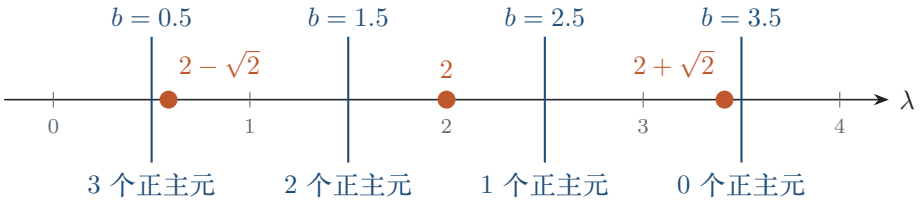


图 24.3　平移定位 K 的特征值。蓝线是试探的 b，下方是 $K-bI$ 的正主元个数，即大于 b 的特征值个数；橙点是精确的特征值。相邻两条蓝线的计数之差就是其间特征值的个数。

想一想　例 24.9 为什么不取 $b=1$？对 $K-I$ 消元试试。

答　$K-I=\begin{bmatrix}1&-1&0\\-1&1&-1\\0&-1&1\end{bmatrix}$，第一个主元 1，第二个主元 $1-1=0$，消元卡住了。可是 $\det(K-I)=-1\neq 0$，1 不是 K 的特征值。零主元只说明左上角的 2×2 子矩阵 $\begin{bmatrix}1&-1\\-1&1\end{bmatrix}$ 奇异，不说明整个矩阵奇异。实际计算时把 b 稍微挪一下即可：$b=\frac{9}{10}$ 时主元为 $\frac{11}{10}$，$\frac{21}{110}$，$-\frac{869}{210}$，两正一负，说明恰有两个特征值大于 0.9。

24.6　正定矩阵

所有特征值都为正的对称矩阵最受欢迎：它可逆，$\dfrac{du}{dt}=-Au$ 的解全部衰减，对应的二次函数有唯一的最小值（第 26 讲）。

定义 24.6（正定矩阵）　实对称矩阵 A 称为**正定的**（positive definite），如果它的所有特征值都大于 0；称为**半正定的**（positive semidefinite），如果所有特征值都 ≥ 0。类似地有负定（全部 <0）、半负定（全部 ≤ 0），其余的称为**不定的**（indefinite）。

例如 I、正对角元的对角矩阵、例 24.8 的 K、例 24.3 的 $I+J$ 都是正定的；$\begin{bmatrix}1&3\\3&1\end{bmatrix}$（特征值 $4,-2$）是不定的；$\begin{bmatrix}1&2\\2&4\end{bmatrix}$（特征值 $5,0$）是半正定的但不是正定的。

易错提醒　「正定」只对对称矩阵定义。$B=\left[\begin{smallmatrix}1&4\\0&1\end{smallmatrix}\right]$ 的特征值都是 1，但它不叫正定：取 $x=(1,-1)$，$x^{\mathsf T}Bx=1-4+1=-2<0$。第 26 讲会说明，正定真正的含义是 $x^{\mathsf T}Ax>0$ 对所有 $x\neq 0$ 成立，而这个性质与特征值全正等价的前提是 A 对称。

按定义判断正定要求出全部特征值，代价太大。下面给出两个便宜的判别法。记 A_k 为 A 左上角的 $k\times k$ 子矩阵，$D_k=\det A_k$ 称为第 k 个**顺序主子式**（leading principal minor）。

引理 24.7（主元等于主子式之比）　若 A 的消元不需要换行，主元为 $d_1,\dots,d_n$，则

$$D_k=d_1d_2\cdots d_k,\qquad d_1=D_1,\quad d_k=\frac{D_k}{D_{k-1}}\ (k\geq 2).$$

反过来，若 $D_1,\dots,D_n$ 都不为零，则消元不需要换行。

证明　消元的前 k 步只用前 k 行去消它们下方的元素，左上角 $k\times k$ 块的变化与单独对 A_k 消元完全相同，所以 A_k 的主元就是 $d_1,\dots,d_k$，而行列式等于主元之积，得 $D_k=d_1\cdots d_k$。反方向对 k 归纳：若前 $k-1$ 个主元非零，则 $d_k=D_k/D_{k-1}\neq 0$，消元可以继续。 ■

定理 24.8（正定的三种判别法）　对实对称矩阵 A，下面三条等价：

(a) 所有特征值 $\lambda_k>0$；

(b) 消元不需要换行，且所有主元 $d_k>0$；

(c) 所有顺序主子式 $D_k>0$，$k=1,\dots,n$。

证明　按 (c)⇒(b)⇒(a)⇒(c) 的顺序证明。

(c)⇒(b)：D_k 全不为零，由引理 24.7，消元不需要换行，且 $d_1=D_1>0$，$d_k=D_k/D_{k-1}>0$。

(b)⇒(a)：n 个主元都为正，由定理 24.5，A 有 n 个正特征值，即全部特征值为正。

(a)⇒(c)：固定 k。A_k 是对称矩阵，设 μ 是它的特征值，$y\in\mathbb{R}^k$ 是实特征向量，$y\neq 0$。把 y 补零得 $x=(y,0,\dots,0)\in\mathbb{R}^n$，则 $x^{\mathsf T}Ax=y^{\mathsf T}A_ky=\mu\,y^{\mathsf T}y$。由引理 24.4，$x^{\mathsf T}Ax\geq\lambda_{\min}(A)\,x^{\mathsf T}x>0$，所以 $\mu>0$。A_k 的特征值全为正，$D_k=\det A_k$ 是它们的乘积，也为正。 ■

■ **例 24.10　课时摘要的例子**　用三种方法判断 $A=\left[\begin{smallmatrix}5&2\\2&3\end{smallmatrix}\right]$ 是否正定。

解　主元：第二行减去第一行的 $\frac{2}{5}$ 倍，$d_2=3-\frac{2}{5}\cdot 2=\frac{11}{5}$。主元 $5,\frac{11}{5}$ 全正。

主子式：$D_1=5>0$，$D_2=15-4=11>0$。与引理一致：$d_2=D_2/D_1=\frac{11}{5}$。

特征值：$\lambda^2-8\lambda+11=0$，$\lambda=4\pm\sqrt{5}\approx 6.236,\ 1.764$，全正。三种方法都判定 A 正定。主元 $5,\frac{11}{5}$ 与特征值 $4\pm\sqrt{5}$ 数值不同，乘积都是 11。 ♦

易错提醒　只检查 $\det A>0$ 不够。$\left[\begin{smallmatrix}-1&0\\0&-3\end{smallmatrix}\right]$ 的行列式是 $3>0$，却是负定的，因为 $D_1=-1<0$。每一个顺序主子式都要检查。另外，把「$>$」改成「$\geq$」后主子式判别法不再成立：$\left[\begin{smallmatrix}0&0\\0&-1\end{smallmatrix}\right]$ 的 $D_1=0$，$D_2=0$，都 ≥ 0，但它有特征值 -1，不是半正定的。

■ **例 24.11**　**含参数的 2×2 矩阵**　b 取何值时 $A=\left[\begin{smallmatrix}2&b\\b&4\end{smallmatrix}\right]$ 正定？何时半正定？

解　$D_1=2>0$ 总成立，$D_2=8-b^2$。所以 A 正定 $\iff b^2<8\iff |b|<2\sqrt{2}$。用特征值核对：$\operatorname{tr}A=6$，$\det A=8-b^2$，$\lambda=3\pm\sqrt{1+b^2}$。较小的特征值 $3-\sqrt{1+b^2}>0\iff 1+b^2<9\iff b^2<8$，结论相同。

$b=\pm 2\sqrt{2}$ 时 $D_2=0$，特征值为 6 和 0，第二个主元 $d_2=4-\frac{b^2}{2}=0$：A 半正定但不正定。$|b|>2\sqrt{2}$ 时较小的特征值为负，A 不定。所以 A 半正定 $\iff |b|\leq 2\sqrt{2}$。

b	$D_2=8-b^2$	$d_2=4-b^2/2$	特征值 $3\pm\sqrt{1+b^2}$	类型
1	7	$\frac{7}{2}$	4.414，1.586	正定
2	4	2	5.236，0.764	正定
$2\sqrt{2}$	0	0	6，0	半正定
3	-1	$-\frac{1}{2}$	6.162，-0.162	不定

♦

■ **例 24.12**　**3×3：Strang 6.5 的一道题**　s 取何值时 $A=\left[\begin{smallmatrix}s&-4&-4\\-4&s&-4\\-4&-4&s\end{smallmatrix}\right]$ 正定？

解　主子式：$D_1=s$，$D_2=s^2-16$，$D_3=\det A=(s+4)^2(s-8)$。三个都为正要求 $s>0$、$s>4$、$s>8$，合起来是 $s>8$。

用特征值更快：$A=(s+4)I-4J$，J 为全 1 矩阵，特征值 $3,0,0$（例 24.3），所以 A 的特征值是 $s+4-12=s-8$ 和 $s+4$（二重）。全正 $\iff s>8$。取 $s=9$：主元 9，$\frac{65}{9}$，$\frac{13}{5}$，主子式 9，65，169，全正。$s=8$ 时 $D_3=0$，A 半正定，$(1,1,1)$ 在零空间中。$s=7$ 时主子式为 7，33，-121，最后一个为负。♦

想一想　对 2×2 对称矩阵，「$\det A>0$ 且 $\operatorname{tr}A>0$」能推出正定吗？3×3 呢？

答　2×2 时能：$\det A=\lambda_1\lambda_2>0$ 说明两个特征值同号，$\operatorname{tr}A=\lambda_1+\lambda_2>0$ 说明它们都为正。3×3 时不能：$\operatorname{diag}(5,-1,-1)$ 的行列式 $5>0$，迹 $3>0$，但有两个负特征值。原因是三个数的乘积为正只能说明负数有偶数个。

前后联系　正定矩阵把本课程的几条主线连在一起：消元的主元、行列式、特征值。第 26 讲给出第四个、也是最根本的判别法 $x^{\mathsf T}Ax>0$（$x\neq 0$），并解释它的几何意义：$f(x)=x^{\mathsf T}Ax$ 的图像是一个向上开口的「碗」，在原点取最小值；还会说明列线性无关时 $A^{\mathsf T}A$ 正定。第 28 讲的奇异值分解就是把谱定理用在半正定矩阵 $A^{\mathsf T}A$ 上。

本讲小结

- 实对称矩阵的特征值都是实数（用 $\bar{x}^{\mathsf{T}}x>0$ 的共轭技巧证明），不同特征值的特征向量互相垂直。
- 谱定理：$A=Q\Lambda Q^{\mathsf{T}}$，$Q$ 正交。实矩阵对称 $\iff$ 能被正交矩阵对角化。重特征值时用 Gram–Schmidt 在特征空间里挑出标准正交基。
- 谱分解 $A=\sum \lambda_k q_k q_k^{\mathsf{T}}$：对称矩阵是互相垂直的投影的组合；$A^m=\sum_j \lambda_j^m q_j q_j^{\mathsf{T}}$；$A$ 可逆时，$A^{-1}=\sum_j \lambda_j^{-1} q_j q_j^{\mathsf{T}}$。
- 对称矩阵消元无需换行且主元均非零时，正（负）主元个数等于正（负）特征值个数；$A-bI$ 满足同样条件时，可数出大于 b 的特征值个数。非对称矩阵没有这个性质。
- 正定：对称且特征值全正。等价于主元全正，也等价于顺序主子式全正。只看 $\det A>0$ 不够。

练习

24.1 ★ 判断真假，假的给出反例：(a) 实矩阵的特征值都是实数，则它是对称的；(b) 对称矩阵的主元就是它的特征值；(c) 可逆对称矩阵 A 的逆 A^{-1} 也对称，且与 A 有相同的特征向量；(d) 对角元全为正的对称矩阵是正定的。

24.2 ★ 求 $A=\left[\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right]$ 的 Q 和 Λ，写出谱分解，并求 A^{10}。

24.3 ★ 判断下列矩阵是正定、半正定、负定还是不定：$\left[\begin{smallmatrix}1&2\\2&5\end{smallmatrix}\right]$，$\left[\begin{smallmatrix}1&2\\2&3\end{smallmatrix}\right]$，$\left[\begin{smallmatrix}-2&1\\1&-2\end{smallmatrix}\right]$，$\left[\begin{smallmatrix}4&2&0\\2&4&2\\0&2&4\end{smallmatrix}\right]$，$\left[\begin{smallmatrix}1&1&1\\1&1&1\\1&1&1\end{smallmatrix}\right]$。

24.4 ★★ 求使下列矩阵正定的参数范围：(a) $\left[\begin{smallmatrix}1&b\\b&9\end{smallmatrix}\right]$；(b) $\left[\begin{smallmatrix}t&3&0\\3&t&4\\0&4&t\end{smallmatrix}\right]$。对 (b) 取 $t=6$ 求出主元。

24.5 ★★ 构造一个 2×2 对称矩阵，使它的特征值为 4 和 -2，且 $(1,2)$ 是 $\lambda=4$ 的特征向量。

24.6 ★★ 仿照定理 24.1 的证明：(a) 若实矩阵 A 满足 $A^{\mathsf{T}}=-A$（反对称），则它的特征值都是纯虚数（实部为零）；(b) 若 Q 是实正交矩阵，则它的特征值都满足 $|\lambda|=1$。

24.7 ★★ 设 A 是实对称矩阵。证明：(a) 若 $A^2=0$，则 $A=0$；(b) 若 $A^3=I$，则 $A=I$。对非对称矩阵，两个结论各举一个反例。

24.8 ★★ 不求特征值，用主元判断 $A=\left[\begin{smallmatrix}1&2&0\\2&1&2\\0&2&1\end{smallmatrix}\right]$ 有几个正特征值；再对 $A-2I$、$A+2I$ 消元，确定每个特征值所在的区间。最后与精确值 $1,\ 1\pm 2\sqrt{2}$ 比较。

24.9 ★★ （习题 25.2）下列集合在矩阵乘法下是否构成群（乘积和逆仍在集合中）：(a) 正定对称矩阵；(b) 正交矩阵；(c) 固定 A 的全体 e^{tA}，$t\in\mathbb{R}$；(d) 行列式为 1 的矩阵。对 (a)，用 $A=\left[\begin{smallmatrix}2&1\\1&1\end{smallmatrix}\right]$，$B=\left[\begin{smallmatrix}1&1/2\\1/2&1\end{smallmatrix}\right]$ 检验。

24.10 ★★★ 设 A,B 都正定。(a) 证明存在正定矩阵 R 使 $R^2=A$，并对 $A=\left[\begin{smallmatrix}5&4\\4&5\end{smallmatrix}\right]$ 求出 R。(b) 证明 AB 的特征值都是正实数（提示：$AB=R(RBR)R^{-1}$）。(c) 对第 9 题的 A,B 验证 (b)。

练习答案与提示

24.1 (a) 假：例 24.1 的 $B=\left[\begin{smallmatrix}1&2\\0&3\end{smallmatrix}\right]$ 特征值 $1,3$，不对称。(b) 假：$\left[\begin{smallmatrix}5&2\\2&3\end{smallmatrix}\right]$ 的主元 $5,\frac{11}{5}$，特征值 $4\pm\sqrt{5}$；只有符号个数和乘积相同。(c) 真：$A^{-1}=Q\Lambda^{-1}Q^{\mathsf T}$ 对称，特征向量仍是 Q 的列，特征值变为 $1/\lambda_k$。(d) 假：$\left[\begin{smallmatrix}1&3\\3&1\end{smallmatrix}\right]$ 的特征值 $4,-2$。

24.2 特征值 3(特征向量 $(1,1)$)和 -1(特征向量 $(-1,1)$)。$Q=\frac{1}{\sqrt2}\left[\begin{smallmatrix}1&-1\\1&1\end{smallmatrix}\right]$，$\Lambda=\operatorname{diag}(3,-1)$，$A=3P_1-P_2$，$P_1,P_2$ 同例 24.6。$A^k=3^kP_1+(-1)^kP_2$，$A^{10}=\frac12\left[\begin{smallmatrix}3^{10}+1&3^{10}-1\\3^{10}-1&3^{10}+1\end{smallmatrix}\right]=\left[\begin{smallmatrix}29525&29524\\29524&29525\end{smallmatrix}\right]$。

24.3 主元依次为：$1,1$，正定（特征值 $3\pm2\sqrt2$）；$1,-1$，不定（特征值 $2\pm\sqrt5$）；$-2,-\frac32$，负定（特征值 $-1,-3$）；$4,3,\frac83$，正定（特征值 $4,\ 4\pm2\sqrt2$）。全 1 矩阵秩为 1，特征值 $3,0,0$，半正定但不正定；消元时第二个主元为 0，$D_2=D_3=0$。

24.4 (a) $D_2=9-b^2>0$，即 $|b|<3$。(b) $D_1=t$，$D_2=t^2-9$，$D_3=t(t^2-16)-9t=t(t^2-25)$。三个都为正 $\iff t>5$。$t=6$ 时主子式 $6,27,66$，主元 $6,\ \frac92,\ \frac{22}{9}$。

24.5 第二个特征向量必须垂直于 $(1,2)$，取 $(-2,1)$。$q_1=\frac{1}{\sqrt5}(1,2)$，$q_2=\frac{1}{\sqrt5}(-2,1)$，$A=4q_1q_1^{\mathsf T}-2q_2q_2^{\mathsf T}=\frac45\left[\begin{smallmatrix}1&2\\2&4\end{smallmatrix}\right]-\frac25\left[\begin{smallmatrix}4&-2\\-2&1\end{smallmatrix}\right]=\frac15\left[\begin{smallmatrix}-4&12\\12&14\end{smallmatrix}\right]$。核对：$A(1,2)=(4,8)$，$A(-2,1)=(4,-2)$。

24.6 (a) 由 $Ax=\lambda x$ 取共轭得 $A\bar x=\bar\lambda\bar x$，转置并用 $A^{\mathsf T}=-A$ 得 $-\bar x^{\mathsf T}A=\bar\lambda\bar x^{\mathsf T}$，右乘 x：$-\bar x^{\mathsf T}Ax=\bar\lambda\,\bar x^{\mathsf T}x$。又 $\bar x^{\mathsf T}Ax=\lambda\,\bar x^{\mathsf T}x$，所以 $-\lambda=\bar\lambda$，实部为零。例：$R=\left[\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right]$ 的特征值 $\pm i$。(b) $Qx=\lambda x$，$Q\bar x=\bar\lambda\bar x$。$(Q\bar x)^{\mathsf T}(Qx)=\bar x^{\mathsf T}Q^{\mathsf T}Qx=\bar x^{\mathsf T}x$，另一方面它等于 $\bar\lambda\lambda\,\bar x^{\mathsf T}x$，所以 $|\lambda|^2=1$。

24.7 (a) $A=Q\Lambda Q^{\mathsf T}$，$A^2=Q\Lambda^2Q^{\mathsf T}=0$ 推出 $\lambda_k^2=0$，$\Lambda=0$，$A=0$。反例 $\left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right]$。(b) $A^3=Q\Lambda^3Q^{\mathsf T}=I$ 推出 $\lambda_k^3=1$，而 λ_k 是实数，只能 $\lambda_k=1$，$A=QIQ^{\mathsf T}=I$。反例：转 120° 的旋转矩阵 $\left[\begin{smallmatrix}-1/2&-\sqrt3/2\\\sqrt3/2&-1/2\end{smallmatrix}\right]$，转三次回到原处，但它不是 I（它的特征值是复数）。

24.8 用递推 $d_{k+1}=c-4/d_k$（c 为对角元）。A 的主元 $1,-3,\frac73$：两正一负，两个正特征值。$A-2I$ 的主元 $-1,3,-\frac73$：一个特征值大于 2。$A+2I$ 的主元 $3,\frac53,\frac35$：全部特征值大于 -2。所以三个特征值分别在 $(-2,0)$、$(0,2)$、$(2,\infty)$ 中，与 $1-2\sqrt2\approx-1.83$、1、$1+2\sqrt2\approx3.83$ 一致。

24.9 (a) 不是。A 的主元 $2,\frac12$，B 的主元 $1,\frac34$，都正定，但 $AB=\left[\begin{smallmatrix}5/2&2\\3/2&3/2\end{smallmatrix}\right]$ 不对称，不在集合中(逆仍正定，特征值为 $1/\lambda_k$)。(b) 是：$(Q_1Q_2)^{\mathsf T}Q_1Q_2=Q_2^{\mathsf T}Q_1^{\mathsf T}Q_1Q_2=I$，$Q^{-1}=Q^{\mathsf T}$ 也正交。(c) 是：$e^{sA}e^{tA}=e^{(s+t)A}$（sA 与 tA 可交换），逆为 e^{-tA}。(d) 是：$\det AB=\det A\det B=1$，$\det A^{-1}=1/\det A=1$。

24.10 (a) 令 $R=Q\Lambda^{1/2}Q^{\mathsf T}$，$\Lambda^{1/2}=\operatorname{diag}(\sqrt{\lambda_k})$。$R$ 对称，特征值 $\sqrt{\lambda_k}>0$，故正定。再利用 $Q^{\mathsf T}Q=I$，得到

$$R^2=Q\Lambda^{1/2}(Q^{\mathsf T}Q)\Lambda^{1/2}Q^{\mathsf T}=A.$$

对 $A=9P_1+P_2$(P_1,P_2 同例 24.6)，$R=3P_1+P_2=\left[\begin{smallmatrix}2&1\\1&2\end{smallmatrix}\right]$，$R^2=\left[\begin{smallmatrix}5&4\\4&5\end{smallmatrix}\right]$。(b) $AB=R^2B=R(RBR)R^{-1}$ 与 $C=RBR$ 相似，特征值相同。$C^{\mathsf T}=R^{\mathsf T}B^{\mathsf T}R^{\mathsf T}=C$，所以特征值是实数。

对 C 的实特征向量 y，$\mu\,y^{\mathsf T}y=y^{\mathsf T}Cy=(Ry)^{\mathsf T}B(Ry)>0$（$R$ 可逆，$Ry\neq0$，再用引理 24.4），故 $\mu>0$。(c) $\operatorname{tr}AB=4$，$\det AB=\det A\det B=1\cdot\frac34=\frac34$，特征值 $2\pm\frac{\sqrt{13}}{2}\approx3.803,\ 0.197$，都为正，尽管 AB 不对称。

第 25 讲

复矩阵与快速 Fourier 变换

Complex Matrices; Fast Fourier Transform

对应材料：MIT 18.06SC Session 3.2「Complex matrices; fast Fourier transform」课时摘要（Strang 第 26 讲视频）；Session 3.2 习题 26.1、26.2 及解答；Strang《Introduction to Linear Algebra》第 4 版 §10.1–10.3（第 5 版 §9.1–9.3）；ZoomNotes Part 9。

本讲目标

(1) 知道为什么实矩阵也离不开复数；会用模、共轭和极坐标处理复数，理解 n 次单位根和「单位根之和为零」。

(2) 掌握复向量的长度 $\|z\|^2 = z^{\mathsf{H}} z$ 与内积 $y^{\mathsf{H}} x$，明白转置时为什么必须取共轭。

(3) 理解 Hermite 矩阵（$A^{\mathsf{H}} = A$）和酉矩阵（$Q^{\mathsf{H}} Q = I$）分别是对称矩阵和正交矩阵在复数域里的对应物，会证明并使用它们的特征值性质。

(4) 熟悉 Fourier 矩阵 F_n：会写出 F_2, F_4，证明 $F_n^{\mathsf{H}} F_n = nI$，会用 $F_n^{-1} = \overline{F_n}/n$ 在「系数」和「采样值」之间来回变换。

(5) 理解 FFT 的分解 $F_{2m} = \begin{bmatrix} I & D \\ I & -D \end{bmatrix} \begin{bmatrix} F_m & \\ & F_m \end{bmatrix} P$，知道它为什么把乘法次数从 n^2 降到约 $\frac{1}{2} n \log_2 n$。

25.1 复数回顾与单位根

第 20 讲已经看到，实矩阵可以有复特征值。把平面逆时针旋转 90° 的矩阵

$$Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

的特征值是 $\pm i$，其中 $\lambda = i$ 的一个特征向量是 $x = (1, -i)$：$Qx = (i, 1) = i\,(1, -i)$。如果照搬实向量的长度公式，会得到

$$x^{\mathsf{T}} x = 1^2 + (-i)^2 = 1 - 1 = 0.$$

一个非零向量的「长度」竟然是 0。这说明一旦特征值和特征向量变成复的，长度、内积、对称、正交这些概念都要重新检查。本讲做两件事：先把 $\mathbb{R}^n$ 上的这套几何搬到 $\mathbb{C}^n$ 上（第 2–4 节）；再认识最重要的复矩阵，即 Fourier 矩阵 F_n，以及把 F_n 乘向量的计算量从 n^2 降到约 $\frac{1}{2} n \log_2 n$ 的快速 Fourier 变换（第 5–6 节）。

25.1.1 复数的两种写法

复数 $z = a + ib$（$a, b \in \mathbb{R}$）对应平面上的点 (a, b)，$a = \operatorname{Re} z$ 是实部，$b = \operatorname{Im} z$ 是虚部。它的**共轭**（complex conjugate）是 $\overline{z} = a - ib$，即关于实轴的镜像。下面这个恒等式是本讲一切的出发点：

$$z\,\overline{z} = (a+ib)(a-ib) = a^2 - i^2 b^2 = a^2 + b^2 = |z|^2.$$

模的平方是 z 乘它的共轭，不是 z^2。由此还得到 $\dfrac{1}{z} = \dfrac{\overline{z}}{|z|^2}$。

第二种写法是**极坐标形式**（polar form）。设 $r = |z|$，θ 是从正实轴转到 z 的角，由 Euler 公式

$$z = r\cos\theta + ir\sin\theta = re^{i\theta}.$$

极坐标让乘法变得简单：$(re^{i\theta})(Re^{i\varphi}) = rR\,e^{i(\theta+\varphi)}$，**模相乘，角相加**。共轭则是把角取反：$\overline{re^{i\theta}} = re^{-i\theta}$。特别地，单位圆上的数 $e^{i\theta}$ 满足 $\overline{e^{i\theta}} = e^{-i\theta} = 1/e^{i\theta}$，即**单位圆上取共轭就是取倒数**。

例如 $z = 1+i = \sqrt{2}\,e^{i\pi/4}$。按极坐标，$z^2 = 2e^{i\pi/2} = 2i$；直接展开，$(1+i)^2 = 1+2i+i^2 = 2i$，两者一致。又 $\overline{z} = 1 - i$，$z\overline{z} = 1 - i^2 = 2 = |z|^2$。

25.1.2 n 次单位根

设 n 是正整数。方程 $z^n = 1$ 在复数域里恰有 n 个解，称为 **n 次单位根**（n-th roots of unity）。它们都在单位圆上（因为 $|z|^n = 1$ 推出 $|z| = 1$），把圆周 n 等分。记

$$\boxed{w = w_n = e^{2\pi i/n} = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}}$$

则全部 n 次单位根就是 $1, w, w^2, \ldots, w^{n-1}$：$w^k$ 的角是 $2\pi k/n$，$(w^k)^n = e^{2\pi i k} = 1$。因为 $w^n = 1$，**w 的指数只需模 n 看**，例如 $w^{n+3} = w^3$，$w^{-1} = w^{n-1}$。又因为 w 在单位圆上，$\overline{w} = w^{-1}$。图 25.1 画出了 $n = 8$ 的情形。

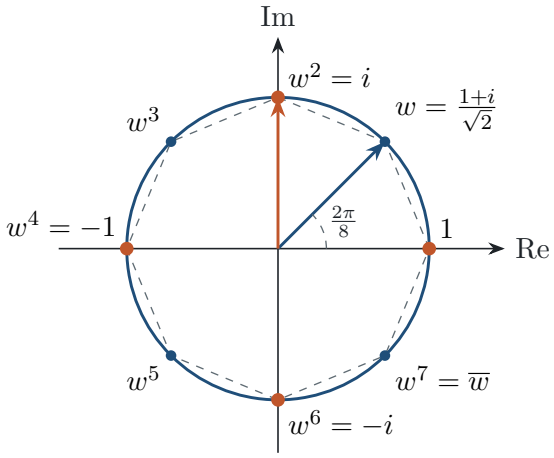


图 25.1　8 次单位根 $w^k = e^{2\pi i k/8}$ 把单位圆 8 等分。橙色点是偶次幂 $1, w^2, w^4, w^6$，它们恰好是 4 次单位根 $1, i, -1, -i$；共轭 $\overline{w} = w^7$ 是 w 关于实轴的镜像。

下面这个求和公式贯穿第 5、6 节，Fourier 矩阵的列互相正交就靠它。

引理 25.1（单位根求和）　设 $w=e^{2\pi i/n}$，k 为整数，则

$$\sum_{j=0}^{n-1} w^{jk} = 1 + w^k + w^{2k} + \cdots + w^{(n-1)k} = \begin{cases} n, & n \mid k, \\ 0, & n \nmid k. \end{cases}$$

证明　若 $n \mid k$，则 $w^k = 1$，每一项都是 1，和为 n。若 $n \nmid k$，则 $q = w^k \neq 1$，用等比数列求和：

$$\sum_{j=0}^{n-1} q^j = \frac{1-q^n}{1-q} = \frac{1-(w^n)^k}{1-q} = \frac{1-1}{1-q} = 0. \qquad \blacksquare$$

几何上也很好理解。$n \nmid k$ 时，把和记为 s，乘以 $q=w^k$ 相当于把每一项向前挪一位：$qs = q + q^2 + \cdots + q^n$，而 $q^n = 1$，所以 qs 和 s 是同样 n 个数之和，$qs = s$。由 $q \neq 1$ 得 $s = 0$。换句话说，绕圆心均匀分布的单位向量，合力为零。

例 25.1　8 次单位根　设 $n=8$，$w = e^{2\pi i/8}$。(1) 把 w, w^2, w^4 写成 $a+ib$ 的形式；(2) 验证 $1 + w + \cdots + w^7 = 0$；(3) 用引理 25.1 计算 $\sum_{j=0}^{7} w^{2j}$、$\sum_{j=0}^{7} w^{4j}$、$\sum_{j=0}^{7} w^{8j}$。

解　(1) $w = \cos\frac{\pi}{4} + i\sin\frac{\pi}{4} = \frac{1+i}{\sqrt{2}}$。平方：$w^2 = \frac{(1+i)^2}{2} = \frac{2i}{2} = i$，再平方 $w^4 = i^2 = -1$，$w^8 = 1$。所以 $w_8^2 = i = w_4$：**8 次单位根的平方是 4 次单位根**，这正是图 25.1 中橙色点的来历，也是第 6 节 FFT 的出发点。

(2) 由 $w^4 = -1$ 得 $w^{j+4} = -w^j$，八项两两配对抵消：

$$(1 + w^4) + (w + w^5) + (w^2 + w^6) + (w^3 + w^7) = 0 + 0 + 0 + 0 = 0.$$

(3) $\sum w^{2j}$ 对应 $k = 2$，$8 \nmid 2$，和为 0；具体地，它是 $1 + i - 1 - i$ 重复两遍。$\sum w^{4j}$ 对应 $k = 4$，$8 \nmid 4$，和为 0，即 $1-1+1-1+1-1+1-1$。$\sum w^{8j}$ 对应 $k = 8$，每项都是 1，和为 8。 ♦

25.2　复向量的长度与内积

25.2.1　为什么要取共轭

在 $\mathbb{R}^n$ 中，$\|x\|^2 = x^{\mathsf{T}}x = x_1^2 + \cdots + x_n^2$。对复向量这个公式失效：开头的例子里 $x^{\mathsf{T}}x = 0$；再如 $z = (1, i)$，$z^{\mathsf{T}}z = 1 + i^2 = 0$。$z^{\mathsf{T}}z$ 一般是复数，根本不能当长度。

我们真正想要的是每个分量模平方之和：$|z_1|^2 + \cdots + |z_n|^2$。由 $|z_j|^2 = \overline{z_j}\, z_j$，只要在转置的同时取共轭就行：

$$\|z\|^2 = \overline{z_1}z_1 + \cdots + \overline{z_n}z_n = \begin{bmatrix} \overline{z_1} & \cdots & \overline{z_n} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.$$

定义 25.2（共轭转置）　矩阵 A 的**共轭转置**（conjugate transpose，Hermitian transpose）是 $A^{\mathsf H}=\overline{A}^{\mathsf T}$，即先转置、再把每个元素取共轭（顺序无所谓），$(A^{\mathsf H})_{ij}=\overline{a_{ji}}$。对复向量 z，$z^{\mathsf H}=\overline{z}^{\mathsf T}$ 是一个行向量，z 的长度定义为

$$\boxed{\|z\|^2=z^{\mathsf H}z=|z_1|^2+\cdots+|z_n|^2}$$

上标 H 取自 Hermite 的名字。对 $z=(1,i)$：$z^{\mathsf H}=\begin{bmatrix}1 & -i\end{bmatrix}$，$z^{\mathsf H}z=1\cdot 1+(-i)\cdot i=1+1=2$，长度 $\sqrt{2}$，这才合理。实向量没有虚部，$z^{\mathsf H}=z^{\mathsf T}$，新定义包含了旧定义。

这个长度还有一个实数解释。写 $z_j=a_j+ib_j$，则 $|z_j|^2=a_j^2+b_j^2$，所以 $\|z\|^2=\sum(a_j^2+b_j^2)$，正好是把 z 看成 $\mathbb{R}^{2n}$ 中的实向量 $(a_1,b_1,\dots,a_n,b_n)$ 时的普通长度平方。

25.2.2　复内积

长度确定后，内积也跟着确定：$x^{\mathsf H}x$ 是 x 和自己的内积，那么 x 与 y 的内积就该是 $y^{\mathsf H}x$。

定义 25.3（复内积与正交）　$x,y\in\mathbb{C}^n$ 的**内积**（inner product）定义为

$$\boxed{y^{\mathsf H}x=\overline{y_1}x_1+\overline{y_2}x_2+\cdots+\overline{y_n}x_n}$$

若 $y^{\mathsf H}x=0$，称 x 与 y **正交**（orthogonal）。特别地 $x^{\mathsf H}x=\|x\|^2\geq 0$，等号仅当 $x=0$。

与实内积相比，有几条规则要重新记：

(1) **交换次序要取共轭**：$x^{\mathsf H}y=\overline{y^{\mathsf H}x}$。所以 $y^{\mathsf H}x=0$ 与 $x^{\mathsf H}y=0$ 等价，正交仍然是对称的关系。

(2) 数乘：$y^{\mathsf H}(cx)=c\,(y^{\mathsf H}x)$，但 $(cy)^{\mathsf H}x=\bar c\,(y^{\mathsf H}x)$，因为 $(cy)^{\mathsf H}=\bar c\,y^{\mathsf H}$。

(3) 乘积：$(AB)^{\mathsf H}=B^{\mathsf H}A^{\mathsf H}$，与 $(AB)^{\mathsf T}=B^{\mathsf T}A^{\mathsf T}$ 同理，多一步逐项取共轭，而 $\overline{AB}=\overline{A}\,\overline{B}$。

(4) 矩阵可以「搬过去」：$(Ax)^{\mathsf H}y=x^{\mathsf H}(A^{\mathsf H}y)$。这一条由第 3 条直接得到，下一节的证明全靠它。

■ **例 25.2　长度、内积与正交的计算**　(1) 求 $u=(1+i,\,2-i)$ 的长度；(2) 对 u 与 $x=(1,i)$ 计算 $u^{\mathsf H}x$ 和 $x^{\mathsf H}u$；(3) 判断 $x=(1,i)$ 与 $y=(1,-i)$ 是否正交，并与 $x^{\mathsf T}y$ 比较。

解　(1) $|1+i|^2=1+1=2$，$|2-i|^2=4+1=5$，所以 $\|u\|^2=u^{\mathsf H}u=7$，$\|u\|=\sqrt{7}$。若误用 $u^{\mathsf T}u=(1+i)^2+(2-i)^2=2i+(3-4i)=3-2i$，得到的不是实数。

(2) $u^{\mathsf H}=\begin{bmatrix}1-i & 2+i\end{bmatrix}$，

$$u^{\mathsf H}x=(1-i)\cdot 1+(2+i)\cdot i=1-i+2i+i^2=i,\qquad x^{\mathsf H}u=1\cdot(1+i)+(-i)(2-i)=1+i-2i+i^2=-i.$$

两者互为共轭，符合规则 1。

(3) $x^{\mathsf H}y=\bar 1\cdot 1+\bar i\cdot(-i)=1+(-i)(-i)=1+i^2=0$，所以 $x\perp y$。而不取共轭时 $x^{\mathsf T}y=1+i(-i)=1+1=2\neq 0$。这两个向量正是开头旋转矩阵 Q 的两个特征向量：$Q(1,i)=(-i,1)=-i\,(1,i)$，$Q(1,-i)=i\,(1,-i)$。Q 是实正交矩阵，它的特征向量理应正交，**只有用带共轭的内积才看得出这一点**。 ♦

易错提醒　计算 $x^{\mathsf H}y$ 时，不能漏掉第一个向量的共轭。例如 $x=(1,i)$、$y=(1,-i)$ 满足 $x^{\mathsf T}y=2$，但 $x^{\mathsf H}y=0$，所以不取共轭的点积不能判断正交。

易错提醒　复内积的共轭放在哪一边，各书约定不同。这里使用 $y^{\mathsf H}x=\sum_j \overline{y_j}x_j$；若另一本书把共轭放在第二个向量上，交换顺序时的共轭规则也要随之调整，但长度与正交判断不变。

25.3　Hermite 矩阵

第 24 讲的实对称矩阵 $A^{\mathsf T}=A$ 有两条好性质：特征值是实数，特征向量可以取成互相正交的。到了复数域，应该要求 $A^{\mathsf T}=A$ 还是 $A^{\mathsf H}=A$？既然长度和内积里的 T 都换成了 H，自然的猜测是后者。下面会看到猜测正确，而且 $A^{\mathsf T}=A$ 的复矩阵保不住这些性质。

定义 25.4（Hermite 矩阵）　满足 $A^{\mathsf H}=A$ 的复方阵称为 **Hermite 矩阵**（Hermitian matrix），即对所有 i,j 有 $a_{ji}=\overline{a_{ij}}$。

取 $i=j$ 得 $a_{ii}=\overline{a_{ii}}$，所以对角元都是实数；对角线两侧的元素互为共轭。例如

$$\begin{bmatrix} 2 & 3+i \\ 3-i & 5 \end{bmatrix}\ \text{是 Hermite 矩阵，}\qquad \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}\ \text{不是（因为 } a_{21}=i\neq\overline{a_{12}}=-i\text{）.}$$

实对称矩阵没有虚部，$A^{\mathsf H}=A^{\mathsf T}=A$，所以实对称矩阵都是 Hermite 矩阵。

定理 25.5（Hermite 矩阵的三条性质）　设 $A^{\mathsf H}=A$。则

(1) 对任意 $x\in\mathbb{C}^n$，$x^{\mathsf H}Ax$ 是实数；

(2) A 的特征值都是实数；

(3) 属于不同特征值的特征向量互相正交。

证明思路　一个数是实数，当且仅当它等于自己的共轭；而 1×1 矩阵的共轭转置就是取共轭。所以证明「某个数是实数」，只要对它取 H，再用 $A^{\mathsf H}=A$ 把它变回自己。第 2、3 条都从 $Ax=\lambda x$ 两边左乘一个 $x^{\mathsf H}$ 或 $y^{\mathsf H}$ 开始。

证明　(1) $x^{\mathsf H}Ax$ 是 1×1 矩阵，$\overline{x^{\mathsf H}Ax}=(x^{\mathsf H}Ax)^{\mathsf H}=x^{\mathsf H}A^{\mathsf H}(x^{\mathsf H})^{\mathsf H}=x^{\mathsf H}Ax$，所以它是实数。

(2) 设 $Ax=\lambda x$，$x\neq0$。左乘 $x^{\mathsf H}$：$x^{\mathsf H}Ax=\lambda\,x^{\mathsf H}x=\lambda\|x\|^2$。左边由 (1) 是实数，$\|x\|^2>0$ 也是实数，所以 $\lambda=x^{\mathsf H}Ax/\|x\|^2$ 是实数。

(3) 设 $Ax=\lambda x$，$Ay=\mu y$，$\lambda\neq\mu$。把 $y^{\mathsf H}Ax$ 用两种方式计算：

$$y^{\mathsf H}(Ax)=\lambda\,y^{\mathsf H}x,\qquad y^{\mathsf H}Ax=(A^{\mathsf H}y)^{\mathsf H}x=(Ay)^{\mathsf H}x=(\mu y)^{\mathsf H}x=\overline{\mu}\,y^{\mathsf H}x=\mu\,y^{\mathsf H}x.$$

最后一步用到 (2)：μ 是实数。两式相减得 $(\lambda-\mu)\,y^{\mathsf H}x=0$，而 $\lambda\neq\mu$，故 $y^{\mathsf H}x=0$。 ■

证明里关键的一步是 $\overline{\mu}=\mu$：第 3 条要靠第 2 条。另外 $A^{\mathsf H}=A$ 用了两次，一次在 (1) 中把

$x^{\mathsf H}A^{\mathsf H}x$ 换回 $x^{\mathsf H}Ax$，一次在 (3) 中把 A 从 x 那边搬到 y 这边。若只知道 $A^{\mathsf T}=A$，这两步都做不了。

Hermite 矩阵也能用一组标准正交特征向量对角化。

定理 25.6（Hermite 谱定理）　若 $A\in\mathbb{C}^{n\times n}$ 且 $A^{\mathsf H}=A$，则存在酉矩阵 U 和实对角矩阵 Λ，使

$$A=U\Lambda U^{\mathsf H}.$$

思路　按矩阵阶数归纳。先取一个单位特征向量 q，把它扩充为 $\mathbb{C}^n$ 的标准正交基；在这组基下，A 的第一列只有左上角可能非零。Hermite 对称性迫使第一行也如此，剩下的 $n-1$ 阶块仍是 Hermite 矩阵，可用归纳假设对角化。

证明　$n=1$ 时，Hermite 条件说明 A 的唯一元素为实数，结论成立。设结论对 $n-1$ 阶成立。由定理 25.5，A 有实特征值 λ 和非零特征向量；单位化后记为 q。把 q 扩充成标准正交基，令 Q_1 的列为这组基，并置 $B=Q_1^{\mathsf H}AQ_1$。则 $B^{\mathsf H}=B$，且 B 的第一列为

$$Q_1^{\mathsf H}Aq=\lambda Q_1^{\mathsf H}q=\lambda e_1.$$

因为 B 是 Hermite 矩阵且 λ 为实数，第一行也只有首项可能非零。故 $B=\operatorname{diag}(\lambda,C)$，其中 C 是 $n-1$ 阶 Hermite 矩阵。由归纳假设，$C=V\Lambda' V^{\mathsf H}$，其中 V 酉、Λ' 实对角。令 $U=Q_1\operatorname{diag}(1,V)$、$\Lambda=\operatorname{diag}(\lambda,\Lambda')$，便有 $A=U\Lambda U^{\mathsf H}$，且 U 酉、Λ 实对角。 ■

■ **例 25.3　一个 2×2 Hermite 矩阵**　设 $A=\begin{bmatrix}2 & 3+i\\ 3-i & 5\end{bmatrix}$。(1) 取 $x=(1,i)$，验证 $x^{\mathsf H}Ax$ 是实数；(2) 求特征值和特征向量，并检验正交性；(3) 写出 $A=U\Lambda U^{\mathsf H}$。

解　(1) $Ax=\begin{bmatrix}2+(3+i)i\\ (3-i)+5i\end{bmatrix}=\begin{bmatrix}1+3i\\ 3+4i\end{bmatrix}$，于是

$$x^{\mathsf H}Ax=1\cdot(1+3i)+(-i)(3+4i)=1+3i-3i-4i^2=5.$$

虚部恰好抵消。A 和 x 都是复的，结果却是实数，这就是定理 25.5(1)。

(2) 特征多项式用到 $(3+i)(3-i)=|3+i|^2=10$：

$$\det(A-\lambda I)=(2-\lambda)(5-\lambda)-10=\lambda^2-7\lambda=\lambda(\lambda-7),$$

所以 $\lambda_1=0,\lambda_2=7$，都是实数（迹 7，行列式 0）。$\lambda_1=0$：解 $2x_1+(3+i)x_2=0$，取 $v_1=(3+i,\,-2)$。$\lambda_2=7$：解 $-5x_1+(3+i)x_2=0$，取 $v_2=(3+i,\,5)$；验算 $Av_2=(21+7i,\,35)=7v_2$。正交性：

$$v_1^{\mathsf H}v_2=\overline{(3+i)}(3+i)+(-2)\cdot 5=10-10=0,$$

而不取共轭时 $v_1^{\mathsf T}v_2=(3+i)^2-10=-2+6i\neq 0$。

(3) $\|v_1\|^2=10+4=14$，$\|v_2\|^2=10+25=35$。令

$$U=\begin{bmatrix}\frac{3+i}{\sqrt{14}} & \frac{3+i}{\sqrt{35}}\\ \frac{-2}{\sqrt{14}} & \frac{5}{\sqrt{35}}\end{bmatrix},\qquad \Lambda=\begin{bmatrix}0 & 0\\ 0 & 7\end{bmatrix},$$

则 $A = U\Lambda U^{\mathsf H} = 0\cdot u_1u_1^{\mathsf H} + 7\,u_2u_2^{\mathsf H}$。验算第二项：

$$\frac{7}{35}\,v_2v_2^{\mathsf H} = \frac{1}{5}\begin{bmatrix}3+i\\5\end{bmatrix}\begin{bmatrix}3-i & 5\end{bmatrix} = \frac{1}{5}\begin{bmatrix}10 & 15+5i\\15-5i & 25\end{bmatrix} = \begin{bmatrix}2 & 3+i\\3-i & 5\end{bmatrix} = A.$$

A 有零特征值，所以秩为 1，谱分解只剩一项。 ♦

■ **例 25.4　复对称矩阵不是正确的推广**　$B = \begin{bmatrix}1 & i\\ i & 1\end{bmatrix}$ 满足 $B^{\mathsf T} = B$，但 $B^{\mathsf H} = \begin{bmatrix}1 & -i\\ -i & 1\end{bmatrix} \neq B$。检查定理 25.5 的前两条结论是否成立。

解　取 $y = (1,1)$：$By = (1+i,\,1+i)$，$y^{\mathsf H}By = 2+2i$，不是实数，结论 (1) 不成立。特征多项式 $(1-\lambda)^2 - i^2 = (1-\lambda)^2 + 1 = 0$，得 $\lambda = 1\pm i$；具体地 $B(1,1) = (1+i)(1,1)$，$B(1,-1) = (1-i)(1,-1)$。特征值不是实数，结论 (2) 也不成立。 ♦

易错提醒　复数域里「对称」有两种含义：$A^{\mathsf T} = A$（复对称）与 $A^{\mathsf H} = A$（Hermite）。只有后者继承实对称矩阵的性质，前者可能连对角化都做不到：$C = \begin{bmatrix}1 & i\\ i & -1\end{bmatrix}$ 满足 $C^{\mathsf T} = C$，但 $C^2 = 0$，$C \neq 0$，所以它的特征值只有 0 而且不可对角化。两种对称只在实矩阵上重合。

想一想　(1) Hermite 矩阵的迹和行列式一定是实数吗？(2) iI 是 Hermite 矩阵吗？

答　(1) 是。迹是对角元之和，对角元是实数；行列式是特征值之积，特征值是实数。(2) 不是，$(iI)^{\mathsf H} = -iI$。它的对角元不是实数，特征值 i 也不是实数。

25.4　酉矩阵

实正交矩阵 Q 的列是单位正交向量，写成矩阵就是 $Q^{\mathsf T}Q = I$。复数域里列向量 $q_1,\dots,q_n$ 单位正交的意思是 $q_i^{\mathsf H}q_j = 0$（$i \neq j$）且 $q_i^{\mathsf H}q_i = 1$；把它们排成 Q，$Q^{\mathsf H}Q$ 的 (i,j) 元正是 $q_i^{\mathsf H}q_j$。

定义 25.7（酉矩阵）　列向量单位正交的复方阵称为**酉矩阵**（unitary matrix），即 $Q^{\mathsf H}Q = I$。等价地 $Q^{-1} = Q^{\mathsf H}$，因而也有 $QQ^{\mathsf H} = I$，行向量同样单位正交。实正交矩阵都是酉矩阵。

定理 25.8（酉矩阵的性质）　设 $Q^{\mathsf H}Q = I$。则

(1) Q 保持内积和长度：$(Qy)^{\mathsf H}(Qx) = y^{\mathsf H}x$，$\|Qx\| = \|x\|$；

(2) Q 的特征值都满足 $|\lambda| = 1$；

(3) 属于不同特征值的特征向量互相正交。

证明　(1) $(Qy)^{\mathsf H}(Qx) = y^{\mathsf H}Q^{\mathsf H}Qx = y^{\mathsf H}x$；取 $y = x$ 得 $\|Qx\|^2 = \|x\|^2$。

(2) 设 $Qx = \lambda x$，$x \neq 0$。由 (1)，$\|x\| = \|Qx\| = \|\lambda x\| = |\lambda|\|x\|$，所以 $|\lambda| = 1$。

(3) 设 $Qx=\lambda x$，$Qy=\mu y$，$\lambda\neq\mu$。由 (1)，$y^{\mathsf H}x=(Qy)^{\mathsf H}(Qx)=\overline{\mu}\lambda\, y^{\mathsf H}x$。若 $y^{\mathsf H}x\neq 0$，则 $\overline{\mu}\lambda=1$；又 $|\mu|=1$ 给出 $\overline{\mu}=1/\mu$，于是 $\lambda=\mu$，矛盾。故 $y^{\mathsf H}x=0$。 ■

条件 $Q^{\mathsf H}Q=I$ 不能去掉：$Q=2I$ 把单位向量 e_1 的长度变成 2，特征值也等于 2，不在单位圆上。

■ **例 25.5　一个酉矩阵**　设 $Q=\dfrac{1}{\sqrt{2}}\begin{bmatrix}1 & i\\ i & 1\end{bmatrix}$。(1) 验证 Q 是酉矩阵；(2) 求特征值并在单位圆上标出；(3) 对 $x=(1,0)$ 验证 $\|Qx\|=\|x\|$。

解　(1) $Q^{\mathsf H}=\dfrac{1}{\sqrt{2}}\begin{bmatrix}1 & -i\\ -i & 1\end{bmatrix}$，

$$Q^{\mathsf H}Q=\frac{1}{2}\begin{bmatrix}1\cdot 1+(-i)i & 1\cdot i+(-i)\cdot 1\\ -i\cdot 1+1\cdot i & (-i)i+1\cdot 1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}=I.$$

(2) Q 就是例 25.4 的 B 除以 $\sqrt{2}$，特征向量不变，特征值也除以 $\sqrt{2}$：

$$\lambda_1=\frac{1+i}{\sqrt{2}}=e^{i\pi/4},\qquad \lambda_2=\frac{1-i}{\sqrt{2}}=e^{-i\pi/4},$$

对应特征向量 $(1,1)$ 与 $(1,-1)$，二者正交。两个特征值正好是图 25.1 中的 w 与 $w^7=\overline{w}$，模都是 1。

(3) $Qx=\frac{1}{\sqrt{2}}(1,\,i)$，$\|Qx\|^2=\frac{1}{2}(1+1)=1=\|x\|^2$。 ♦

例 25.4 和本例放在一起看：B 不是 Hermite 矩阵，特征值不在实轴上；$B/\sqrt{2}$ 是酉矩阵，特征值落在单位圆上。

直观理解　把 1×1 矩阵 $[z]$ 看成一个复数：$z^{\mathsf H}=\overline{z}$。Hermite 条件 $\overline{z}=z$ 说 z 是实数；酉条件 $\overline{z}z=1$ 说 z 在单位圆上；$\overline{z}=-z$（**反 Hermite**（skew-Hermitian），$A^{\mathsf H}=-A$）说 z 是纯虚数。$n\times n$ 的情形把这个图景搬到特征值上：Hermite 矩阵的特征值在实轴上，酉矩阵的在单位圆上，反 Hermite 矩阵的在虚轴上。

表 25.1 把第 2–4 节的对应关系列在一起。规律只有一条：**凡是实数情形出现转置 T 的地方，复数情形都换成共轭转置 H。**

表 25.1　$\mathbb{R}^n$ 与 $\mathbb{C}^n$ 的对照

	实数情形 $\mathbb{R}^n$	复数情形 $\mathbb{C}^n$
长度平方	$x^{\mathsf{T}}x=\sum x_j^2$	$z^{\mathsf{H}}z=\sum\lvert z_j\rvert^2$
内积	$y^{\mathsf{T}}x$	$y^{\mathsf{H}}x$，且 $x^{\mathsf{H}}y=\overline{y^{\mathsf{H}}x}$
正交	$y^{\mathsf{T}}x=0$	$y^{\mathsf{H}}x=0$
转置规则	$(AB)^{\mathsf{T}}=B^{\mathsf{T}}A^{\mathsf{T}}$	$(AB)^{\mathsf{H}}=B^{\mathsf{H}}A^{\mathsf{H}}$
对称 / Hermite	$A^{\mathsf{T}}=A$	$A^{\mathsf{H}}=A$
谱分解	$A=Q\Lambda Q^{\mathsf{T}}$，$\Lambda$ 实	$A=U\Lambda U^{\mathsf{H}}$，$\Lambda$ 实
正交 / 酉	$Q^{\mathsf{T}}Q=I$	$Q^{\mathsf{H}}Q=I$
特征值	对称：实数；正交：$\lvert\lambda\rvert=1$	Hermite：实数；酉：$\lvert\lambda\rvert=1$

想一想　若 Q 既是 Hermite 矩阵又是酉矩阵，它的特征值可能是哪些数？举一个例子。

答　Hermite 给出 λ 是实数，酉给出 $\lvert\lambda\rvert=1$，所以 $\lambda=\pm1$。这时 $Q^2=Q^{\mathsf{H}}Q=I$。例如交换两个分量的置换矩阵 $\begin{bmatrix}0&1\\1&0\end{bmatrix}$，特征值 1 和 -1，特征向量 $(1,1)$ 和 $(1,-1)$。

25.5　Fourier 矩阵与离散 Fourier 变换

给定四个周期样本 $(1,2,0,-1)$，怎样分别看出其中的低频和高频成分？第 23 讲把周期函数展开成 Fourier 级数；有限个等距样本则由**离散 Fourier 变换**（discrete Fourier transform, DFT）分解。

本讲固定采用正指数的前向约定。给定正整数 n 和样本 $f_0,\dots,f_{n-1}$，令 $w=e^{2\pi i/n}$，定义未归一化的频率系数

$$c_k=\sum_{j=0}^{n-1}w^{jk}f_j,\qquad 0\le k<n.$$

对应的反变换是

$$f_j=\frac{1}{n}\sum_{k=0}^{n-1}w^{-jk}c_k.$$

正负指数的约定有时会对调；对调时只需把 F_n 与 $\overline{F_n}$ 互换，归一化因子仍是 $1/n$。

25.5.1 定义与例子

定义 25.9（Fourier 矩阵） 设 n 是正整数，$w = e^{2\pi i/n}$。$n \times n$ **Fourier 矩阵** (Fourier matrix) F_n 的 (j,k) 元是 w^{jk}，**行列下标都从 0 数到 $n-1$**：

$$F_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w & w^2 & \cdots & w^{n-1} \\ 1 & w^2 & w^4 & \cdots & w^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & w^{n-1} & w^{2(n-1)} & \cdots & w^{(n-1)^2} \end{bmatrix}.$$

第 k 列的元素 $w^{jk} = e^{ikx_j}$ 是复指数在采样点上的取值；由于 $F_n^{\mathsf{T}} = F_n$，同一组元素也出现在第 k 行，给出 $c_k = \sum_j w^{jk} f_j$。所以这里 F_n 把样本送到频率系数，F_n^{-1} 把系数还原为样本。$F_n^{\mathsf{T}} = F_n$ 只表示转置对称，通常并不表示 $F_n^{\mathsf{H}} = F_n$；元素模为 1，指数按模 n 化简。

■ **例 25.6　写出 F_2 与 F_4（习题 26.1）**　写出 F_2 和 F_4，并检查 F_4 的第 1 列与第 3 列是否正交。

解　$n=2$ 时 $w = e^{i\pi} = -1$，$F_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$。$n=4$ 时 $w = e^{i\pi/2} = i$，指数按模 4 化简（$i^4 = 1$，$i^6 = i^2 = -1$，$i^9 = i$）：

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}.$$

第 1 列 $c_1 = (1, i, -1, -i)$，第 3 列 $c_3 = (1, -i, -1, i)$。带共轭的内积为

$$c_1^{\mathsf{H}} c_3 = 1 \cdot 1 + (-i)(-i) + (-1)(-1) + (i)(i) = 1 - 1 + 1 - 1 = 0,$$

两列正交。不取共轭则 $c_1^{\mathsf{T}} c_3 = 1 + 1 + 1 + 1 = 4$；更离谱的是 $c_1^{\mathsf{T}} c_1 = 1 - 1 + 1 - 1 = 0$，而真正的长度平方 $c_1^{\mathsf{H}} c_1 = 4$。 ♦

■ **例 25.7　四个样本的 DFT**　对样本 $f = (1, 2, 0, -1)^{\mathsf{T}}$，用 F_4 求未归一化的 Fourier 系数 c，再用逆变换还原 f。

解　按 F_4 的四行逐项求和。这里 $w = i$：

$$\begin{aligned} c_0 &= 1 + 2 + 0 - 1 = 2, \\ c_1 &= 1 + 2i + 0 + (-i)(-1) = 1 + 3i, \\ c_2 &= 1 - 2 + 0 + 1 = 0, \\ c_3 &= 1 - 2i + 0 + i(-1) = 1 - 3i. \end{aligned}$$

所以 $c=(2,1+3i,0,1-3i)^{\mathsf{T}}$。逆变换是 $f=\frac{1}{4}\overline{F_4}c$；例如第一项为 $(2+(1+3i)+(1-3i))/4=1$，其余三项依次为 $2,0,-1$，恰好还原原样本。这也说明前向 DFT 的系数未除以 4，归一化因子全部放在逆变换中。 ♦

25.5.2　F_n 的列正交，逆矩阵几乎不用算

定理 25.10（Fourier 矩阵的正交性）　$F_n^{\mathsf{H}}F_n=nI$。因此

$$\boxed{F_n^{-1}=\frac{1}{n}F_n^{\mathsf{H}}=\frac{1}{n}\overline{F_n},}$$

并且 $F_n/\sqrt{n}$ 是酉矩阵。

证明　$F_n^{\mathsf{H}}F_n$ 的 (j,k) 元是第 j 列与第 k 列的内积。由 $\overline{w}=w^{-1}$，

$$\sum_{l=0}^{n-1}\overline{w^{lj}}\,w^{lk}=\sum_{l=0}^{n-1}w^{l(k-j)}.$$

由引理 25.1（取 $k-j$ 代替 k），当 $k=j$ 时和为 n；当 $k\neq j$ 时 $0<|k-j|<n$，$n\nmid(k-j)$，和为 0。所以 $F_n^{\mathsf{H}}F_n=nI$。又 $F_n^{\mathsf{T}}=F_n$，故 $F_n^{\mathsf{H}}=\overline{F_n}^{\mathsf{T}}=\overline{F_n}$。 ■

$\overline{F_n}$ 也是 Fourier 矩阵，只不过把 w 换成 $\overline{w}=e^{-2\pi i/n}$。所以**求逆和正变换是同一种运算**，下一节的快速算法对两个方向都适用。

■ **例 25.8　单位脉冲与常数向量**　计算 F_4e_0 与 $F_4(1,1,1,1)$，其中 $e_0=(1,0,0,0)$；再用 F_4^{-1} 把 $(1,1,1,1)$ 变回去。

解　F_4e_0 是 F_4 的第 0 列：$F_4e_0=(1,1,1,1)$。$F_4(1,1,1,1)$ 是各行之和：第 0 行和为 4，其余各行是单位根之和，由引理 25.1 为 0，所以 $F_4(1,1,1,1)=(4,0,0,0)=4e_0$。反过来，$F_4^{-1}(1,1,1,1)=\frac{1}{4}\overline{F_4}(1,1,1,1)=\frac{1}{4}(4,0,0,0)=e_0$。 ♦

Strang 的说法是：**一个单位脉冲包含等量的所有频率**，它的未归一化 Fourier 系数全是 1；反过来，常数样本只含频率 0。这里 F_4 用两次把 e_0 变成 $4e_0$，但这不说明 $F_4^2=4I$：一般 $F_n^2\neq nI$，练习第 10 题会算出具体反例。把变换还原要用 $\overline{F_n}/n$。

25.6　快速 Fourier 变换：偶数项与奇数项

直接计算 F_4x 要做 16 次复数乘法。把输入按偶、奇下标分开后，只需两次 F_2 变换和少量合并运算。这个拆分来自下面的矩阵恒等式。

定理 25.11（Fourier 矩阵的偶奇分解）　设 $m \geq 1$，$w = e^{2\pi i/(2m)}$，并令

$$D = \operatorname{diag}(1, w, \dots, w^{m-1}), \qquad P_{2m}x = (x_0, x_2, \dots, x_{2m-2}, x_1, x_3, \dots, x_{2m-1})^{\mathsf{T}}.$$

右侧的 F_m 使用原始 $2m$ 次单位根的平方 $w^2 = e^{2\pi i/m}$。则

$$F_{2m} = \begin{bmatrix} I_m & D \\ I_m & -D \end{bmatrix} \begin{bmatrix} F_m & 0 \\ 0 & F_m \end{bmatrix} P_{2m}.$$

思路　先用置换 P_{2m} 把偶数下标和奇数下标的输入分别排在一起，再让两个 F_m 独立变换。最后用 D 调整奇数部分的相位，并把结果相加、相减。证明只需分别计算输出的第 j 项和第 $j+m$ 项。

证明　令 $x_{\mathrm{e}} = (x_0, x_2, \dots, x_{2m-2})^{\mathsf{T}}$，$x_{\mathrm{o}} = (x_1, x_3, \dots, x_{2m-1})^{\mathsf{T}}$，并记 $a = F_m x_{\mathrm{e}}$、$b = F_m x_{\mathrm{o}}$。对 $0 \leq j < m$，$F_{2m}x$ 的第 j 个分量为

$$\begin{aligned} y_j &= \sum_{\ell=0}^{m-1} w^{j(2\ell)} x_{2\ell} + \sum_{\ell=0}^{m-1} w^{j(2\ell+1)} x_{2\ell+1} \\ &= a_j + w^j b_j. \end{aligned}$$

第 $j+m$ 个分量中，偶数项的因子多出 $w^{2m\ell} = 1$，奇数项多出 $w^m = -1$，所以 $y_{j+m} = a_j - w^j b_j$。因此

$$\begin{bmatrix} y_0 \\ \vdots \\ y_{m-1} \\ y_m \\ \vdots \\ y_{2m-1} \end{bmatrix} = \begin{bmatrix} I_m & D \\ I_m & -D \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix},$$

而 $\begin{bmatrix} a \\ b \end{bmatrix} = \operatorname{diag}(F_m, F_m) P_{2m} x$。合并两式即得结论。 ∎

分解里的 D 要使用 $2m$ 次单位根 w，不能误用 m 次单位根；奇数部分相对偶数部分多出的相位正是 w^j。

例 25.9　用 F_4 做一次蝶形计算　取 $x = (1, 2, 3, 4)^{\mathsf{T}}$，按定理 25.11 的顺序计算 $F_4 x$，并与直接相乘比较。

解　先重排输入：$P_4 x = (1, 3, 2, 4)^{\mathsf{T}}$。偶数项和奇数项分别做 F_2 变换：

$$a = F_2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}, \qquad b = F_2 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ -2 \end{bmatrix}.$$

这里 $D=\operatorname{diag}(1,i)$，所以

$$F_4x=\begin{bmatrix}a_0+b_0\\ a_1+ib_1\\ a_0-b_0\\ a_1-ib_1\end{bmatrix}=\begin{bmatrix}10\\ -2-2i\\ -2\\ -2+2i\end{bmatrix}.$$

直接用 F_4 的四行也得到 10、$1+2i-3-4i=-2-2i$、$1-2+3-4=-2$、$1-2i-3+4i=-2+2i$。两种算法结果相同，分解没有改变变换，只减少了重复计算。 ♦

图 25.2 把这个计算写成数据流：先分离偶、奇样本，再各算一次 F_2，最后用 $D=\operatorname{diag}(1,i)$ 合并。

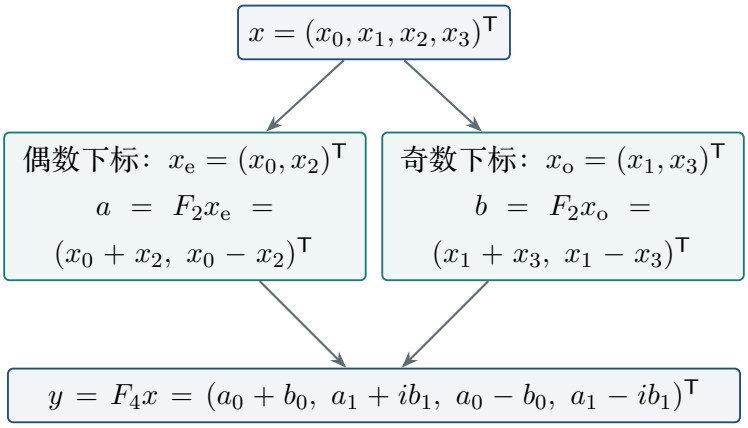


图 25.2　偶奇重排后分别计算两次 F_2，再用 $D=\operatorname{diag}(1,i)$ 相加减得到 F_4x。

左、右两枝都只作用在两个输入上；只有奇数枝的第二个结果乘以 i。这正对应分解矩阵里的 D，也解释了为什么 D 要保留第四次单位根 i。

易错提醒　把 D 错写成 $\operatorname{diag}(1,-1)$ 会把 F_4e_1 算成 $(1,-1,-1,1)^{\mathsf T}$；直接用 F_4 得到 $(1,i,-1,-i)^{\mathsf T}$。正确的 $D=\operatorname{diag}(1,i)$ 使用的是四次单位根 $1,i$，不是平方根 $1,-1$。

把这个分解递归用于两个 F_m，就得到**快速 Fourier 变换**（fast Fourier transform, FFT）。对 $n=2^p$，令 $M(n)$ 表示未优化计数：把每个 D 的对角元都按一次复数乘法计算，不把重排和加减计入这个数，则

$$M(1)=0,\qquad M(n)=2M(n/2)+n/2.$$

每层的 $n/2$ 来自对角矩阵 D，两个子问题各有 $n/2$ 个输入。归纳可得

$$\boxed{M(n)=\frac{n}{2}\log_2 n\qquad (n=2^p).}$$

直接用稠密矩阵乘法，按每个矩阵元做一次复数乘法，需要 n^2 次；FFT 还要做 $n\log_2 n$ 次加减，置换 P 只重新排列数据。以 $n=1024$ 为例，朴素计数是 1,048,576 次乘法，FFT 的对角乘法计数是 $512\cdot 10=5{,}120$ 次，少约 205 倍。许多旋转因子是 $1,-1,i,-i$，实际实现可把它们化成复制、变号或交换实虚部，因此乘法还会更少。

FFT 的拆分原则　遇到偶数长度的 Fourier 变换，先把样本按下标奇偶分组；每组各做一半长度的变换，再乘上旋转因子并相加、相减。把这一步重复到长度为 1，就把一次 n^2 规模的稠密乘法改成约 $\frac{1}{2}n\log_2 n$ 次旋转因子乘法。

想一想　这个偶奇分解只对 $n=2^p$ 成立吗?

答　矩阵恒等式 F_{2m} 对任意正整数 m 都成立；$n=2^p$ 是反复二等分并一直降到 F_1 的条件。若长度不是 2 的幂，仍可在偶数长度上做一次拆分，但递归会遇到奇数长度或不同因子；相应 FFT 需要按这些因子继续分解，不能直接套用上面的二进制乘法计数。

本讲小结

- 复数 $z=a+ib$ 满足 $|z|^2=z\bar{z}$；n 次单位根 $w=e^{2\pi i/n}$ 满足 $w^n=1$。
- 单位根求和：$\sum_{j=0}^{n-1}w^{jk}=n$ 若 $n\mid k$，否则为 0。
- $y^{\mathsf{H}}x=\sum_j \overline{y_j}x_j$，$\|x\|^2=x^{\mathsf{H}}x=\sum_j|x_j|^2$；换序会取共轭。
- 若 $A^{\mathsf{H}}=A$，则 $x^{\mathsf{H}}Ax$ 为实数；特征值为实数，不同特征值的特征向量正交。
- 若 $Q^{\mathsf{H}}Q=I$，则 Q 保持内积和长度，且每个特征值满足 $|\lambda|=1$。
- $F_n=(w^{jk})_{j,k=0}^{n-1}$ 满足 $F_n^{\mathsf{H}}F_n=nI$，故 $F_n^{-1}=\overline{F_n}/n$。
- $n=2^p$ 时，FFT 的偶奇分解把乘法次数降为 $\frac{n}{2}\log_2 n$，加减次数为 $n\log_2 n$。

练习

25.1 ★　判断真假，假的给出反例：(a) 对任意 $z\in\mathbb{C}^n$，$\|z\|^2=z^{\mathsf{T}}z$；(b) Hermite 矩阵的对角元都是实数；(c) 酉矩阵的特征值模都为 1；(d) 复矩阵满足 $A^{\mathsf{T}}=A$ 就一定是 Hermite 矩阵。

25.2 ★　对 $u=(1+i,2-i)$、$x=(1,i)$，求 $\|u\|^2$、$\|u\|$、$u^{\mathsf{H}}x$ 与 $x^{\mathsf{H}}u$。

25.3 ★★　对 $H=\begin{bmatrix}2 & 3+i\\ 3-i & 5\end{bmatrix}$，求特征值和一组正交特征向量，并核验它们的内积为 0。

25.4 ★　求 F_2，并验证 $F_2^{\mathsf{H}}F_2=2I$（习题 26.1）。

25.5 ★★　对样本 $f=(1,2,0,-1)^{\mathsf{T}}$，求 $c=F_4f$，再用 F_4^{-1} 还原 f。

25.6 ★　设 $w=e^{2\pi i/8}$。用等比数列求和证明 $8\nmid k$ 时 $\sum_{j=0}^{7}w^{jk}=0$；再求 $k=2,4,8$ 时的和。

25.7 ★★　求习题 26.2 中 F_4 的 D 和偶奇置换矩阵 P_4，并验证分解式。

25.8 ★★　用偶奇分解计算 F_4x，其中 $x=(1,2,3,4)^{\mathsf{T}}$，并与直接矩阵乘法比较。

25.9 ★★　按本讲的乘法计数，比较 $n=64$、$n=1024$ 时直接乘 F_n 与 FFT 的复数乘法次数。

25.10 ★★★　计算 F_4^2 及 F_4^2x，其中 $x=(1,2,3,4)^{\mathsf{T}}$，说明为什么逆变换不能简单地「再乘一次 F_4」。

练习答案与提示

25.1 (a) 假：$z=(1,i)$ 时 $z^{\mathsf T}z=0$，但 $\|z\|^2=2$。(b) 真，因为 $a_{jj}=\overline{a_{jj}}$。(c) 真，由定理 25.8(2)。(d) 假：$B=\begin{bmatrix}1 & i\\ i & 1\end{bmatrix}$ 满足 $B^{\mathsf T}=B$，但 $B^{\mathsf H}\neq B$，其特征值为 $1+i$、$1-i$。

25.2 $\|u\|^2=|1+i|^2+|2-i|^2=2+5=7$，故 $\|u\|=\sqrt{7}$。$u^{\mathsf H}x=(1-i)+(2+i)i=i$，而 $x^{\mathsf H}u=\bar{i}=-i$。

25.3 $\det(H-\lambda I)=\lambda(\lambda-7)$，特征值为 $0,7$。可取 $v_0=(3+i,-2)$ 与 $v_7=(3+i,5)$；直接乘得 $Hv_0=0$、$Hv_7=7v_7$。$v_0^{\mathsf H}v_7=|3+i|^2-10=0$。

25.4 $F_2=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$，所以 $F_2^{\mathsf H}F_2=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}=2I$。

25.5 逐行相乘得 $c=(2,1+3i,0,1-3i)^{\mathsf T}$。逆变换 $\overline{F_4}c/4$ 的四项为 $(1,2,0,-1)^{\mathsf T}$，所以恢复了原样本。

25.6 若 $8\nmid k$，则 $q=w^k\neq 1$ 且 $q^8=1$，故 $\sum_{j=0}^{7}w^{jk}=\sum_{j=0}^{7}q^j=(1-q^8)/(1-q)=0$。$k=2,4$ 的和均为 0；$k=8$ 时各项为 1，和为 8。

25.7 $D=\begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix}$，$P_4=\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}$。代入得

$$\begin{bmatrix}I & D\\ I & -D\end{bmatrix}\begin{bmatrix}F_2 & 0\\ 0 & F_2\end{bmatrix}P_4=\begin{bmatrix}1&1&1&1\\1&i&-1&-i\\1&-1&1&-1\\1&-i&-1&i\end{bmatrix}=F_4.$$

25.8 令 $a=F_2(1,3)^{\mathsf T}=(4,-2)^{\mathsf T}$、$b=F_2(2,4)^{\mathsf T}=(6,-2)^{\mathsf T}$。合并为 $F_4x=(10,-2-2i,-2,-2+2i)^{\mathsf T}$；直接用 F_4 的行点积也得相同结果。

25.9 直接乘法的计数为 n^2，偶奇递归的对角乘法计数为 $\frac{n}{2}\log_2 n$。$n=64$ 时是 4096 对 192；$n=1024$ 时是 $1{,}048{,}576$ 对 $5{,}120$。

25.10 由 F_4^2 的 (j,k) 元 $\sum_{\ell=0}^{3}i^{\ell(j+k)}$ 可知，当 $j+k\equiv 0\pmod 4$ 时该元为 4，否则为 0。因此

$$F_4^2=\begin{bmatrix}4&0&0&0\\0&0&0&4\\0&0&4&0\\0&4&0&0\end{bmatrix},\qquad F_4^2(1,2,3,4)^{\mathsf T}=(4,16,12,8)^{\mathsf T}.$$

这会把分量按 $0,3,2,1$ 的顺序重排并乘以 4，并非恒等矩阵；逆变换应乘 $\overline{F_4}/4$。

第 26 讲

正定矩阵与极小值

Positive Definite Matrices and Minima

对应材料：MIT 18.06SC Session 3.3（Lecture 27）课时摘要；习题 27.1、27.2 及解答；Strang《ZoomNotes for Linear Algebra》6.3 节。

本讲目标

(1) 会把实对称矩阵写成二次型 $x^{\mathsf{T}}Ax$，并从它的正负判断图像是碗、平面方向还是鞍面。

(2) 会用特征值、能量、主元和顺序主子式判别正定，能在 2×2 情形完成平方。

(3) 会对高维对称矩阵检查顺序主子式，并把消元主元写成二次型中的平方系数。

(4) 能用 Hessian 判断非退化驻点的局部极小值，区分局部结论与二次函数的全局极小值。

(5) 会处理非对称矩阵的二次型，并证明两个正定矩阵的乘积虽可不对称、特征值仍为正。

26.1 二次型的正负和几何

矩阵 $A=\begin{bmatrix}2 & 6\\ 6 & 18\end{bmatrix}$ 作用在向量 $x=(x_1,x_2)^{\mathsf{T}}$ 上，先算 $x^{\mathsf{T}}Ax$：

$$x^{\mathsf{T}}Ax=2x_1^2+12x_1x_2+18x_2^2=2(x_1+3x_2)^2.$$

这个数不会为负，却会在非零向量 $(3,-1)^{\mathsf{T}}$ 上等于 0。只看矩阵元素的正负，很难一眼看出这条零方向；二次型把矩阵与几何联系了起来。

定义 26.1（正定性与二次型） 设 A 是实对称矩阵。函数 $q_A(x)=x^{\mathsf{T}}Ax$ 称为由 A 确定的**二次型**（quadratic form）。若 $q_A(x)>0$ 对每个 $x\neq 0$ 成立，称 A**正定**（positive definite）；若 $q_A(x)\geq 0$ 对所有 x 成立，称 A**半正定**（positive semidefinite）。相应地，若 $q_A(x)<0$ 对每个 $x\neq 0$ 成立，称 A**负定**（negative definite）；若 $q_A(x)\leq 0$ 对所有 x 成立，称 A**半负定**（negative semidefinite）。若 q_A 既取正值又取负值，称 A**不定**（indefinite）。

例子可以直接从平方看出：I 正定，因为 $x^{\mathsf{T}}Ix=\|x\|^2>0$；上面的矩阵半正定但不正定，因为 $q_A(3,-1)=0$；$-I$ 负定；而 $\operatorname{diag}(1,-1)$ 不定，因为 $(1,0)$ 给出 1，$(0,1)$ 给出 -1。

■ **例 26.1　一条正方向和一条负方向**　对 $S=\begin{bmatrix}2 & 6\\ 6 & 7\end{bmatrix}$，判断 $q_S(x,y)=2x^2+12xy+7y^2$ 的符号。

解　先沿坐标方向试值：$q_S(1,0)=2>0$。再沿 $x=-3y$ 附近检查，取 $(x,y)=(-3,1)$，得 $q_S(-3,1)=18-36+7=-11<0$。因此 q_S 既有正值又有负值，S 不定。混合项 $12xy$ 可以压过两个正平方项，所以对角元都正仍不足以保证正定。 ♦

易错提醒　「对称矩阵的对角元全正，所以矩阵正定」是错误的。上例的两个对角元是 $2,7$，但 $q_S(-3,1)=-11$。还必须检查交叉项；正定要求每个非零方向上的总和都严格为正。

图 26.1 把三种符号情形放在同一张等值线图上。椭圆只在原点达到 0，半正定例子有整条零线，不定例子的零等值线则把正值区和负值区分开。

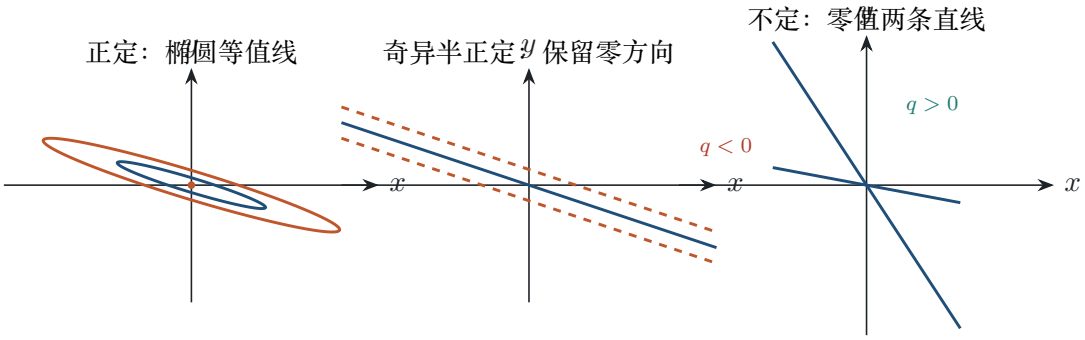


图 26.1　正定二次型的正等值线是椭圆；奇异半正定二次型有非零零方向；不定二次型跨过零值线改变符号。

椭圆图画的是 $q_B=\frac{1}{2}$ 和 $q_B=2$，其中 $B=\begin{bmatrix}2&6\\6&20\end{bmatrix}$；半正定图画的是 $q_P=0$ 与 $q_P=2$；右图画的是 $q_S=0$ 的两条方向线。我们会在下一节验证这些曲线分别对应正定、半正定和不定矩阵。

想一想　半正定矩阵与正定矩阵只差一个「$>$」和「$\geq$」，为什么这一个符号会改变几何图像？

答　正定要求除原点外 q_A 处处大于 0，所以零等值集只有原点；半正定允许某些非零向量满足 $q_A(x)=0$。对称半正定矩阵可沿这些向量方向保持平坦，例如 $q_P(x_1,x_2)=2(x_1+3x_2)^2$ 在直线 $x_1=-3x_2$ 上为零。

26.2　2×2：四种判据是一件事

对称矩阵 $A=\begin{bmatrix}a&b\\b&c\end{bmatrix}$ 的二次型是

$$q_A(x,y)=ax^2+2bxy+cy^2.$$

交叉项 $2bxy$ 的正负随方向改变，所以 $a>0$、$c>0$ 还不够。真正起作用的是交叉项能否被两个平方项控制。

定理 26.2（2×2 正定判据）　设 $A=\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ 为实对称矩阵。以下条件等价：

(a) A 的两个特征值都为正；

(b) $x^{\mathsf T}Ax>0$ 对每个非零 $x\in\mathbb{R}^2$ 成立；

(c) $a>0$ 且 $ac-b^2>0$；

(d) 消元的两个主元 a 与 $c-b^2/a$ 都为正。

直观理解　正交特征向量把椭圆的主轴转回坐标轴，$x^{\mathsf T}Ax$ 就是沿两条主轴的平方和；消元则不转轴，而是把同一表达式改写成两个平方。特征值和主元的数值不同，但它们记录的是同一个正负性。

思路　先用谱定理证明 (a) 与 (b) 等价。再把 q_A 配成一个关于 x 的平方加一个关于 y 的平方，直接得到 (c)；消元的第二个主元正是第二个平方前的系数，得到 (d)。

证明　**(a) 与 (b)**。由第 24 讲的谱定理，存在正交矩阵 Q 使 $A=Q\,\mathrm{diag}(\lambda_1,\lambda_2)Q^{\mathsf T}$。令 $z=Q^{\mathsf T}x$，则

$$x^{\mathsf T}Ax=\lambda_1 z_1^2+\lambda_2 z_2^2.$$

若两个特征值都正，右边对每个 $z\neq 0$ 为正。反过来取单位特征向量 $x=q_i$，有 $x^{\mathsf T}Ax=\lambda_i$，所以 (b) 推出两个特征值都正。

(b) 与 (c)。若 (b) 成立，取 $(x,y)=(1,0)$ 得 $a>0$。再取 $x=-b/a$、$y=1$，交叉项恰好抵消，得

$$q_A\left(-\frac{b}{a},1\right)=c-\frac{b^2}{a}=\frac{ac-b^2}{a}>0.$$

于是 $ac-b^2>0$。反过来，若 $a>0$ 且 $ac-b^2>0$，则

$$q_A(x,y)=a\left(x+\frac{b}{a}y\right)^2+\frac{ac-b^2}{a}y^2.$$

两个系数都正。若 $y\neq 0$，第二项为正；若 $y=0$ 且 $(x,y)\neq(0,0)$，第一项为正。因此 (b) 成立。

(c) 与 (d)。消去第一列下方的 b，第二个主元为 $c-b^2/a$。因为 $a>0$，它为正当且仅当 $ac-b^2>0$。四个条件等价。 ■

主元条件与行列式条件对应得很直接：$D_1=a$、$D_2=ac-b^2$，所以 $D_2=d_1d_2$。不过要先有 $D_1>0$；只看总行列式的符号会漏掉整条负定的可能。

对称矩阵的特征值为实数，而 $\det A$ 是两个特征值的乘积，$\mathrm{tr}\,A=a+c$ 是它们的和。因此行列式还可以配合迹给出完整分类。若 $\det A<0$，两个特征值异号，矩阵不定；若 $\det A>0$，两个特征值同号，再由 $\mathrm{tr}\,A$ 的正负分别判为正定、负定。比如 B 的行列式为 4、迹为 22，所以正定；$-I$ 的行列式为 1、迹为 -2，所以负定。

边界 $\det A=0$ 表示至少一个特征值为零；此时另一个特征值就是 $\mathrm{tr}\,A$。迹为正时矩阵半正定，迹为负时半负定，迹也为零时两个特征值都为零，实对称矩阵便是零矩阵。例子 $P=\begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}$

的行列式为零、迹为 20，故半正定且有一个零方向；$-P$ 则半负定。这些边界情形说明严格正定必须排除零特征值。

■ **例 26.2　越过正定的边界**　对 $A_t=\begin{bmatrix}2&6\\6&t\end{bmatrix}$，求 t 使 A_t 正定、半正定或不定。

解　直接看两个判别量：$D_1=2$，$D_2=2t-36$。所以 A_t 正定当且仅当 $t>18$。当 $t=18$ 时，

$$x^{\mathsf T}A_{18}x=2(x+3y)^2,$$

矩阵半正定，特征值为 $20,0$，零空间由 $(3,-1)^{\mathsf T}$ 张成。若 $t<18$，取 $(1,0)$ 得正值 2，取 $(-3,1)$ 得 $t-18<0$，故矩阵不定。这个参数正好穿过零特征值所在的边界。 ♦

想一想　为什么条件里既要 $a>0$，又要 $ac-b^2>0$？
答　行列式 $ac-b^2>0$ 只说明两个实特征值同号，不能判断它们是正还是负。矩阵 $-I$ 的行列式是 $1>0$，但二次型恒为负；$a>0$ 排除了负定情形。反过来，$a>0$ 单独也不够，例如 A_7 的第一个主元为正而第二个为 -11。

■ **例 26.3　平方完成显示三种形状**　分别把 $S=\begin{bmatrix}2&6\\6&7\end{bmatrix}$ 与 $B=\begin{bmatrix}2&6\\6&20\end{bmatrix}$ 的二次型配方，并判断类型。

解　按定理 26.2 的第二主元计算。对 S，主元是 $7-6^2/2=-11$，所以

$$x^{\mathsf T}Sx=2(x+3y)^2-11y^2.$$

一项正平方、一项负平方说明它是不定的。对 B，第二主元是 $20-6^2/2=2$，所以

$$x^{\mathsf T}Bx=2(x+3y)^2+2y^2.$$

两个平方系数都正，只有 $x=y=0$ 时等于零；B 正定，其等值线是椭圆。核对特征值判据：$\operatorname{tr}B=22$、$\det B=4$，特征多项式为 $\lambda^2-22\lambda+4$，根为 $11\pm3\sqrt{13}$。因为 $121>117$，较小的根也为正。 ♦

26.3　更高维：消元把能量写成平方和

2×2 的行列式只有一个，容易检查；维数增加后，单看 $\det A$ 仍只得到一个符号。对称矩阵的每个左上角方块都要检查。

定义 26.3（顺序主子式）　设 A_k 是 A 左上角的 $k\times k$ 子矩阵。数 $D_k=\det A_k$ 称为 A 的第 k 个**顺序主子式**（leading principal minor）。约定 $D_0=1$。

定理 26.4（正定的顺序主子式判据）　实对称矩阵 A 正定，当且仅当所有顺序主子式都为正：

$$D_1 > 0,\ D_2 > 0,\ \dots,\ D_n > 0.$$

在不换行的对称消元中，第 k 个主元为 $d_k = D_k/D_{k-1}$；因此正定也等价于所有主元为正。

思路　若所有 $D_k > 0$，消元的主元 $d_k = D_k/D_{k-1}$ 都存在且为正，分解 $A = LDL^{\mathsf T}$ 把 $x^{\mathsf T}Ax$ 写成正系数平方和。反过来，正定矩阵的每个左上角主块仍正定，因为可以把向量补零；它的特征值都正，行列式也正。

证明　**顺序主子式全正推出正定。**由 $D_0 = 1$ 与 $D_k > 0$，每个 $d_k = D_k/D_{k-1}$ 都为正，消元无需换行。对称消元给出 $A = LDL^{\mathsf T}$，其中 L 是可逆的单位下三角矩阵，$D = \operatorname{diag}(d_1, \dots, d_n)$。于是

$$x^{\mathsf T}Ax = (L^{\mathsf T}x)^{\mathsf T}D(L^{\mathsf T}x) = \sum_{k=1}^{n} d_k\big((L^{\mathsf T}x)_k\big)^2.$$

若 $x \neq 0$，可逆性保证 $L^{\mathsf T}x \neq 0$；至少一个平方非零，且对应系数为正，故 $x^{\mathsf T}Ax > 0$。

正定推出顺序主子式全正。固定 k，取任意非零 $y \in \mathbb{R}^k$ 并补零成 $x = (y, 0, \dots, 0) \in \mathbb{R}^n$。则 $y^{\mathsf T}A_k y = x^{\mathsf T}Ax > 0$，所以 A_k 正定。由谱定理，A_k 的所有特征值都正；$D_k = \det A_k$ 是这些特征值的乘积，因此 $D_k > 0$。这对每个 k 都成立。 ■

■ **例 26.4　三阶差分矩阵**　对

$$K = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix},$$

计算顺序主子式、主元和特征值，再把 $x^{\mathsf T}Kx$ 写成平方和。

解　按左上角的 1×1、2×2、3×3 子矩阵计算，得到 $D_1 = 2$、$D_2 = 3$、$D_3 = 4$。定理 26.4 判定 K 正定；相邻主子式之比给主元 $2, \frac{3}{2}, \frac{4}{3}$。

直接展开二次型并配方：

$$x^{\mathsf T}Kx = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3 = 2\left(x_1 - \frac{1}{2}x_2\right)^2 + \frac{3}{2}\left(x_2 - \frac{2}{3}x_3\right)^2 + \frac{4}{3}x_3^2.$$

三个平方的系数恰好是三个主元。特征多项式为 $(\lambda-2)(\lambda^2-4\lambda+2)$，故特征值是 $2-\sqrt{2}, 2, 2+\sqrt{2}$，都为正。以 $x = (1,2,3)^{\mathsf T}$ 核对，前两个平方为 0，最后一项为 12，而矩阵乘法也给 $x^{\mathsf T}Kx = 12$。 ♦

三维图像 $z = x^{\mathsf T}Kx$ 是一个向上开的碗；水平截面 $x^{\mathsf T}Kx = 1$ 是椭球面。若 $K = Q\Lambda Q^{\mathsf T}$，换到特征向量坐标后方程变为 $\lambda_1u_1^2 + \lambda_2u_2^2 + \lambda_3u_3^2 = 1$。第 i 条主轴的半轴长为 $1/\sqrt{\lambda_i}$：特征值越大，沿该方向的截面越窄。

定理 26.5（$A^{\mathsf T}A$ 的正定性）　任意实矩阵 C 都使 $C^{\mathsf T}C$ 半正定；$C^{\mathsf T}C$ 正定，当且仅当 C 的列线性无关。

证明　对任意 x，

$$x^{\mathsf{T}}C^{\mathsf{T}}Cx=(Cx)^{\mathsf{T}}(Cx)=\|Cx\|^2\geq 0.$$

若 C 的列线性无关，$Cx=0$ 只能有 $x=0$，故非零 x 给出严格正值。若列相关，则存在 $x\neq 0$ 使 $Cx=0$，于是 $x^{\mathsf{T}}C^{\mathsf{T}}Cx=0$，不可能正定。 ■

■ **例 26.5　列是否独立，决定严格正性**　比较 $C=\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}$ 与 $D=\begin{bmatrix}1&1\\2&2\end{bmatrix}$ 所得到的 $C^{\mathsf{T}}C$ 和 $D^{\mathsf{T}}D$。

解　先算 $C^{\mathsf{T}}C=\begin{bmatrix}2&1\\1&2\end{bmatrix}$，其顺序主子式为 $2,3$，所以正定；这与 C 的两列线性无关相符。再算 $D^{\mathsf{T}}D=\begin{bmatrix}5&5\\5&5\end{bmatrix}$，取 $x=(1,-1)^{\mathsf{T}}$ 得 $Dx=0$，故 $x^{\mathsf{T}}D^{\mathsf{T}}Dx=0$。所以 $D^{\mathsf{T}}D$ 只有半正定性。 ♦

26.4　Hessian 与局部极小值

注 26.1　[阅读准备] 本节需要偏导数、梯度和多元二阶 Taylor 展开。尚未学过多元微积分时，可暂缓本节及对应练习，先掌握前面的二次型与正定判据；后续相似矩阵和 SVD 的证明只需要这些线性代数结论，不依赖 Hessian 判别法。

两个函数 $f(x,y)=2x^2+12xy+20y^2$ 与 $g(x,y)=2x^2+12xy+7y^2$ 的一阶偏导在原点都为零；原点都是驻点。可是 f 在原点取严格最小值，g 沿方向 $(-3,1)$ 取负值，原点是鞍点。一阶导数只能找到水平切平面，还需要二阶导数判断曲面的弯曲方向。

定义 26.6（Hessian 矩阵）　设 $f:\mathbb{R}^n\to\mathbb{R}$ 在点 x 的二阶偏导连续。由所有二阶偏导组成的矩阵

$$H_f(x)=\left(\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right)_{i,j=1}^{n}$$

称为 f 在 x 处的**Hessian 矩阵**（Hessian matrix）。二阶偏导连续时，$H_f(x)$ 对称。

对刚才两个函数，$H_f(0)=\begin{bmatrix}4&12\\12&40\end{bmatrix}$，其顺序主子式为 $4,16$，正定；而 $H_g(0)=\begin{bmatrix}4&12\\12&14\end{bmatrix}$，行列式为 -88，不定。Hessian 把「一个向上、一个鞍形」的差异变成矩阵的正负性。

定理 26.7（Hessian 的二阶判别）　设 f 在 x_* 的一个邻域内二阶偏导连续，且 $\nabla f(x_*)=0$。则：

(a) 若 $H_f(x_*)$ 正定，x_* 是严格局部极小点；

(b) 若 $H_f(x_*)$ 负定，x_* 是严格局部极大点；

(c) 若 x_* 是局部极小点，$H_f(x_*)$ 半正定；

(d) 若 $H_f(x_*)$ 不定，x_* 是鞍点，既非局部极小点也非局部极大点。

在两个变量时，这个判别可直接写成偏导数形式。令

$$H_f(x_*)=\begin{bmatrix} f_{xx}(x_*) & f_{xy}(x_*) \\ f_{xy}(x_*) & f_{yy}(x_*) \end{bmatrix}, \qquad \Delta = f_{xx}(x_*)f_{yy}(x_*)-f_{xy}(x_*)^2.$$

在驻点且二阶偏导连续的前提下，$\Delta>0$ 且 $f_{xx}(x_*)>0$ 时 Hessian 正定，故为严格局部极小；$\Delta>0$ 且 $f_{xx}(x_*)<0$ 时 Hessian 负定，故为严格局部极大；$\Delta<0$ 时 Hessian 不定，故为鞍点。$\Delta=0$ 时二阶判别没有结论，须检查高阶项或直接比较函数值。这个规则只是 2×2 主子式判据应用于 Hessian：迹或某个二阶偏导单独为正，并不足以判定极小值。

例如前面的 f 满足 $\Delta=16>0$、$f_{xx}(0)=4>0$，所以原点是严格局部极小；g 的 $\Delta=-88<0$，所以原点是鞍点。若把 f 换成 $-f$，Hessian 负定，原点便成为严格局部极大。边界 $\Delta=0$ 时，x^2+y^4 与 x^2-y^4 在原点有相同的半正定 Hessian，却分别给出严格极小和鞍点，说明退化情形必须继续检查高阶项。

思路　沿任意方向 v 看一元函数 $g(t)=f(x_*+tv)$。它在 $t=0$ 的二阶导数是 $v^{\mathsf T}H_f(x_*)v$，所以局部极小要求每个方向的二阶变化非负。若 Hessian 正定，其最小特征值给所有方向一个统一的正下界；二阶 Taylor 展开保证足够靠近 x_* 时，余项不会改变这个正号。

证明　记 $H=H_f(x_*)$。由二阶 Taylor 展开，对 $h\to 0$ 有

$$f(x_*+h)=f(x_*)+\frac{1}{2}h^{\mathsf T}Hh+r(h), \qquad \frac{r(h)}{\|h\|^2}\longrightarrow 0.$$

(a) 正定给严格局部极小。设 H 的最小特征值为 $\mu>0$。由谱定理，$h^{\mathsf T}Hh\ge\mu\|h\|^2$。取充分小的非零 h，使 $|r(h)|\le\frac{\mu}{4}\|h\|^2$，则

$$f(x_*+h)-f(x_*)\ge\frac{\mu}{2}\|h\|^2-|r(h)|\ge\frac{\mu}{4}\|h\|^2>0.$$

所以 x_* 是严格局部极小点。**(b) 负定给严格局部极大。**函数 $-f$ 在 x_* 的梯度仍为零，其 Hessian 是 $-H$，为正定。由 (a)，$-f$ 在 x_* 严格局部极小，等价于 f 在 x_* 严格局部极大。**(c) 局部极小要求 Hessian 半正定。**对任意方向 v，令 $g(t)=f(x_*+tv)$。g 在 0 取局部极小值，所以 $g''(0)\ge 0$。链式法则给 $g''(0)=v^{\mathsf T}Hv$，对每个 v 都成立，故 H 半正定。

(d) 不定 Hessian 给鞍点。取单位特征向量 u_+、u_-，对应特征值分别为正数 λ_+ 和负数 λ_-。沿 $h=tu_+$，二阶主项为 $\frac{1}{2}\lambda_+t^2>0$；沿 $h=tu_-$，主项为 $\frac{1}{2}\lambda_-t^2<0$。余项除以 t^2 趋于 0，故充分小的非零 t 下两种符号都保留；任意邻域内都有函数值高于和低于 $f(x_*)$ 的点。 ■

图 26.2 把驻点处 Hessian 的三种情形并排比较。正定给出严格局部极小；不定给出鞍点；半正定但奇异时，二阶信息还不能下结论。

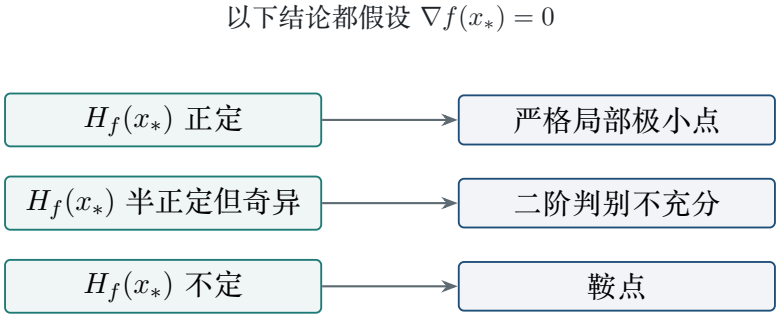


图 26.2　驻点处 Hessian 正定或不定时可分类；半正定且奇异时需要更高阶信息。

■ 例 26.6　**一阶导数相同，二阶形状不同**　在原点比较 $f(x,y) = 2x^2 + 12xy + 20y^2$ 与 $g(x,y) = 2x^2 + 12xy + 7y^2$，说明它们为何分别有极小值和鞍点。

解　先求梯度：$\nabla f = (4x + 12y,\, 12x + 40y)$，$\nabla g = (4x + 12y,\, 12x + 14y)$，两者在原点均为零。$H_f(0)$ 的顺序主子式 $4, 16$ 都正，所以它正定；由定理 26.7，原点是严格局部极小点。由于 f 是齐次二次式且 $f = x^{\mathsf{T}}Bx$ 正定，$f(x) > f(0) = 0$ 对每个 $x \neq 0$ 成立，原点还是全局唯一极小点。

$H_g(0)$ 的行列式为 -88，两个特征值异号，故它不定；定理 26.7 判定原点为鞍点。直接沿 $(1,0)$ 与 $(-3,1)$ 检查，$g(1,0) = 2$、$g(-3,1) = -11$，也能看见一正一负。 ♦

驻点处的 Hessian 是局部判据。对二次函数，Hessian 不随位置改变，而且平方完成能给出全局结论。

定理 26.8（正定二次函数的唯一极小值）　设 A 是实对称正定矩阵，$r \in \mathbb{R}^n$，$c \in \mathbb{R}$，并令 $F(x) = x^{\mathsf{T}}Ax + r^{\mathsf{T}}x + c$。则 F 有唯一全局极小点

$$x_* = -\frac{1}{2}A^{-1}r,$$

极小值为 $F(x_*) = c - \frac{1}{4}r^{\mathsf{T}}A^{-1}r$。

证明　因为 A 正定，所以 A 可逆。展开平方：

$$F(x) = \left(x + \frac{1}{2}A^{-1}r\right)^{\mathsf{T}} A \left(x + \frac{1}{2}A^{-1}r\right) + c - \frac{1}{4}r^{\mathsf{T}}A^{-1}r.$$

第一项非负，且只在 $x = -\frac{1}{2}A^{-1}r$ 时为零，因此该点给出唯一全局极小值，公式也随之成立。 ■

■ 例 26.7　**把碗的底部移到别处**　求 $F(x,y) = 2x^2 + 12xy + 20y^2 - 4x$ 的全局最小值。

解　这里 $A=\begin{bmatrix}2 & 6\\ 6 & 20\end{bmatrix}$ 正定，$r=(-4,0)^{\mathsf T}$。由定理 26.8，

$$A^{-1}=\frac{1}{4}\begin{bmatrix}20 & -6\\ -6 & 2\end{bmatrix},\qquad x_*=-\frac{1}{2}A^{-1}r=(10,-3)^{\mathsf T}.$$

代入极小值公式得 $F(x_*)=-20$。直接平方完成也可核对：

$$F(x,y)=-20+2\big((x-10)+3(y+3)\big)^2+2(y+3)^2.$$

两个平方同时为零只能有 $(x,y)=(10,-3)$，所以这是唯一全局极小点，而非仅仅某个局部极小点。 ♦

■ **例 26.8**　**半正定但奇异时，二阶判别会停下来**　比较 $u(x,y)=x^2+y^4$ 与 $v(x,y)=x^2-y^4$ 在原点的 Hessian 和极值类型。

解　两函数的梯度在原点都为零，Hessian 都是 $\begin{bmatrix}2 & 0\\ 0 & 0\end{bmatrix}$，半正定但奇异。$u(x,y)>0$ 对每个非零 (x,y) 成立，所以原点是严格全局极小点。$v(1,0)=1$ 而 $v(0,1)=-1$，原点附近同时有更大和更小的值，所以它是鞍点。相同的二阶项没有记录 y^4 的正负；这说明 Hessian 半正定且奇异时要回到函数本身继续判断。 ♦

易错提醒　「Hessian 正定就有极小值」少了驻点条件。$h(t)=(t-1)^2$ 在 $t=0$ 的二阶导数为 $2>0$，但 $h'(0)=-2$，且 $h(1)=0<h(0)=1$；二阶判别必须在 $\nabla h=0$ 的点使用。

想一想　Hessian 正定是否也是局部极小值的必要条件？

答　不是。函数 $p(t)=t^4$ 在 0 有严格极小值，但 $p''(0)=0$，Hessian 只有半正定性。定理 26.7(a) 是充分条件；(b) 给出的必要条件只是半正定。若 Hessian 奇异，二阶判别本身不能定性。

方法：先找平衡点，再看曲率　求光滑函数的极小值时，先解 $\nabla f=0$ 找驻点，再检查 Hessian。正定给严格局部极小；对二次函数，正定二次项还允许平方完成，从而直接得到唯一全局极小点。

26.5　非对称矩阵的两种情形

正定矩阵按定义是对称的，但非对称矩阵仍能产生二次型，也能出现在矩阵乘积中。这里的两道原习题分别检验这两点：二次型只看对称部分；正定矩阵的乘积可不对称，却仍有正特征值。

■ **例 26.9**　**正定矩阵的乘积**　设 A,B 都是实对称正定矩阵。证明 AB 的每个特征值都是正实数，并说明 AB 不必对称。（习题 27.1）

解　取任意特征值 λ 及对应的非零复特征向量 x，使 $ABx=\lambda x$。使用复内积并把 B 移到左边：

$$(Bx)^*A(Bx)=(ABx)^*(Bx)=\overline{\lambda}\,x^*Bx.$$

对复向量 $z=u+iv\neq 0$，对称性使展开式中的交叉项相消，故 $z^{*}Az=u^{\mathsf{T}}Au+v^{\mathsf{T}}Av>0$；对 B 也同样成立。于是 $x^{*}Bx>0$，从而 $Bx\neq 0$；A 正定又给出 $(Bx)^{*}A(Bx)>0$。等式两边除以 $x^{*}Bx$，可知 $\overline{\lambda}$ 必为正实数，因而 λ 也是正实数。

这个证明也修正了「直接对实特征向量取点积」的省略：实矩阵的特征值未必一开始就已知为实数，使用共轭转置才能覆盖所有特征向量。下一例用具体矩阵看见乘积确实可能不对称。 ♦

■ **例 26.10　乘积不对称，特征值仍为正**　取 $A=\begin{bmatrix}2&1\\1&2\end{bmatrix}$、$B=\begin{bmatrix}3&1\\1&2\end{bmatrix}$，求 AB 并验证上一例的结论。

解　A 的主元为 $2,\frac{3}{2}$，B 的主元为 $3,\frac{5}{3}$，所以两者都正定。相乘得

$$AB=\begin{bmatrix}7&4\\5&5\end{bmatrix},\qquad \det(\lambda I-AB)=\lambda^{2}-12\lambda+15.$$

特征值是 $6\pm\sqrt{21}$，都为正，因为 $6>\sqrt{21}$。而 AB 的两个非对角元 4 与 5 不相等，所以 AB 不对称，也不能把它称作正定矩阵。 ♦

■ **例 26.11　非对称矩阵的二次型**　对 $C=\begin{bmatrix}1&5\\7&9\end{bmatrix}$，求 $x^{\mathsf{T}}Cx$，并判断它的取值是否变号。（习题 27.2）

解　按矩阵乘法计算：

$$x^{\mathsf{T}}Cx=x(x+5y)+y(7x+9y)=x^{2}+12xy+9y^{2}.$$

取 $(x,y)=(1,0)$ 得 $1>0$；取 $(2,-2)$ 得 $4-48+36=-8<0$，所以这个二次型会变号。为说明原因，写 $C=S+K$，其中 $S=(C+C^{\mathsf{T}})/2=\begin{bmatrix}1&6\\6&9\end{bmatrix}$ 对称，$K=(C-C^{\mathsf{T}})/2$ 反对称。标量 $x^{\mathsf{T}}Kx$ 等于它自己的转置 $x^{\mathsf{T}}K^{\mathsf{T}}x=-x^{\mathsf{T}}Kx$，故为 0；因此 $x^{\mathsf{T}}Cx=x^{\mathsf{T}}Sx$，二次型的符号由 S 决定。此处 $\det S=-27<0$，与取到正负值一致。 ♦

易错提醒　非对称矩阵不能直接套用对称矩阵的正定判别式。对 C，二次型由对称部分 $S=(C+C^{\mathsf{T}})/2$ 决定，所以检查 $\det C$ 不是判别 $x^{\mathsf{T}}Cx$ 正负的通用方法；本例应看 $\det S=-27$，而 $\det C=-26$ 是另一个数。

本讲小结

- 实对称矩阵 A 正定 $\iff x^{\mathsf T}Ax>0$ 对每个 $x\neq 0$；半正定允许非零零方向。
- 对 $A=\begin{bmatrix}a & b\\ b & c\end{bmatrix}$，正定 $\iff a>0$ 且 $ac-b^2>0$；此时 $x^{\mathsf T}Ax=a(x+by/a)^2+(ac-b^2)y^2/a$。
- 一般维数下，A 正定 $\iff D_k>0$ 对所有左上角主子式成立；主元 $d_k=D_k/D_{k-1}$ 是平方和的系数。
- $x^{\mathsf T}C^{\mathsf T}Cx=\|Cx\|^2$，所以 $C^{\mathsf T}C$ 半正定；它正定 $\iff C$ 的列线性无关。
- 在驻点 $\nabla f(x_*)=0$，Hessian 正定给严格局部极小；局部极小要求 Hessian 半正定；Hessian 不定给鞍点。
- 若 $F(x)=x^{\mathsf T}Ax+r^{\mathsf T}x+c$ 且 A 正定，则 $x_*=-\frac{1}{2}A^{-1}r$ 是唯一全局极小点，$F(x_*)=c-\frac{1}{4}r^{\mathsf T}A^{-1}r$。
- 非对称 C 的二次型只取决于 $S=(C+C^{\mathsf T})/2$；两个实对称正定矩阵的乘积可不对称，但特征值仍为正实数。

练习

26.1 ★ 判断真假并给出理由或反例：(a) 实对称矩阵的对角元全正就一定正定；(b) 2×2 对称矩阵的行列式为正就一定正定；(c) 驻点处 Hessian 半正定就一定是局部极小点。

26.2 ★ 对 $P=\begin{bmatrix}2 & 6\\ 6 & 18\end{bmatrix}$、$S=\begin{bmatrix}2 & 6\\ 6 & 7\end{bmatrix}$、$B=\begin{bmatrix}2 & 6\\ 6 & 20\end{bmatrix}$ 分别判别正定性，并对半正定或不定的矩阵给出相应方向。

26.3 ★ 对 $A_t=\begin{bmatrix}2 & 6\\ 6 & t\end{bmatrix}$，求正定、半正定和不定的参数范围。

26.4 ★★ 对三阶差分矩阵 $K=\begin{bmatrix}2 & -1 & 0\\ -1 & 2 & -1\\ 0 & -1 & 2\end{bmatrix}$，计算顺序主子式、主元和特征值，并将 $x^{\mathsf T}Kx$ 配成平方和。

26.5 ★★ 证明 $C^{\mathsf T}C$ 半正定，且正定当且仅当 C 的列线性无关；用本讲的 3×2 矩阵 C 和 2×2 矩阵 D 检查两种情况。

26.6 ★★ 对 $f(x,y)=2x^2+12xy+20y^2$ 与 $g(x,y)=2x^2+12xy+7y^2$，用梯度和 Hessian 判别原点的类型。

26.7 ★★ 求 $F(x,y)=2x^2+12xy+20y^2-4x$ 的驻点与全局最小值，并用平方完成核对。

26.8 ★★ （习题 27.2）对 $C=\begin{bmatrix}1 & 5\\ 7 & 9\end{bmatrix}$ 求 $x^{\mathsf T}Cx$，给出正值和负值，并求其对称部分的行列式。

26.9 ★★★ （习题 27.1）若 A,B 都是实对称正定矩阵，证明 AB 的所有特征值为正实数。证明中要说明如何处理复特征向量。

26.10 ★★★ 设 $u = x^2 + y^4$、$v = x^2 - y^4$。验证二者在原点的 Hessian 相同，却有不同的极值类型；再说明 $p(t) = t^4$ 对「Hessian 正定是极小值必要条件」构成什么反例。

练习答案与提示

26.1 (a) 假：$S = \left[\begin{smallmatrix} 2 & 6 \\ 6 & 7 \end{smallmatrix}\right]$ 对角元为正，但 $q_S(-3,1) = -11$。(b) 假：$-I$ 的行列式为 1，二次型却恒为负；2×2 正定还需左上角元为正。(c) 假：$u = x^2 + y^4$ 与 $v = x^2 - y^4$ 在原点有同一个半正定奇异 Hessian，但 $v(0,1) = -1 < v(0,0) = 0$。

26.2 P 半正定，因为 $q_P = 2(x+3y)^2$，并在 $(3,-1)$ 方向为零。S 不定：$q_S(1,0) = 2$、$q_S(-3,1) = -11$。B 正定，因为 $q_B = 2(x+3y)^2 + 2y^2$，两个平方同时为零只可能在原点。

26.3 $D_1 = 2$、$D_2 = 2t - 36$。因此 $t > 18$ 时正定；$t = 18$ 时半正定且有零方向 $(3,-1)$；$t < 18$ 时用 $(1,0)$ 得正值 2，用 $(-3,1)$ 得负值 $t - 18$，所以不定。

26.4 $D_1 = 2$、$D_2 = 3$、$D_3 = 4$，主元为 $2, \frac{3}{2}, \frac{4}{3}$。特征多项式为 $(\lambda-2)(\lambda^2-4\lambda+2)$，特征值为 $2-\sqrt{2}, 2, 2+\sqrt{2}$。平方分解为

$$x^{\mathsf T}Kx = 2\left(x_1 - \frac{1}{2}x_2\right)^2 + \frac{3}{2}\left(x_2 - \frac{2}{3}x_3\right)^2 + \frac{4}{3}x_3^2.$$

26.5 对任意 x，$x^{\mathsf T}C^{\mathsf T}Cx = \|Cx\|^2 \ge 0$。若列线性无关，$Cx = 0$ 只可能有 $x = 0$，所以 $C^{\mathsf T}C$ 正定；若列相关，则存在非零 x 使 $Cx = 0$，所以只能半正定。例中 $C^{\mathsf T}C = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$ 正定，而 $D^{\mathsf T}D = \left[\begin{smallmatrix} 5 & 5 \\ 5 & 5 \end{smallmatrix}\right]$ 在 $(1,-1)$ 上为零。

26.6 两个梯度在原点为零。$H_f(0) = \left[\begin{smallmatrix} 4 & 12 \\ 12 & 40 \end{smallmatrix}\right]$ 的顺序主子式是 $4, 16$，正定，所以原点是严格局部极小点；$H_g(0) = \left[\begin{smallmatrix} 4 & 12 \\ 12 & 14 \end{smallmatrix}\right]$ 的行列式为 -88，不定，所以原点是鞍点。

26.7 $A = \left[\begin{smallmatrix} 2 & 6 \\ 6 & 20 \end{smallmatrix}\right]$、$r = (-4,0)^{\mathsf T}$，得 $(x_*, y_*) = (10,-3)$。平方完成：

$$F(x,y) = -20 + 2\big((x-10) + 3(y+3)\big)^2 + 2(y+3)^2.$$

故唯一全局最小值是 -20，在 $(10,-3)$ 取得。

26.8 $x^{\mathsf T}Cx = x(x+5y) + y(7x+9y) = x^2 + 12xy + 9y^2$。$(1,0)$ 给 1，$(2,-2)$ 给 -8，所以它不定。对称部分 $S = \left[\begin{smallmatrix} 1 & 6 \\ 6 & 9 \end{smallmatrix}\right]$，$\det S = 9 - 36 = -27$。

26.9 取任意复特征对 $ABx = \lambda x$，$x \neq 0$。若 $z = u + iv \neq 0$，实对称性使 $z^*Az = u^{\mathsf T}Au + v^{\mathsf T}Av > 0$，对 B 也成立。因此 $x^*Bx > 0$，特别地 $Bx \neq 0$。于是

$$(Bx)^*A(Bx) = (ABx)^*(Bx) = \overline{\lambda}\, x^*Bx.$$

等式两边除以正实数 x^*Bx，可知 $\overline{\lambda}$、进而 λ 是正实数。不能只假设 x 是实向量，因为实矩阵的特征向量可能是复数。

26.10 两者在原点的梯度为零，Hessian 都是 $\operatorname{diag}(2,0)$。u 在原点严格取最小值；$v(1,0) = 1$、$v(0,1) = -1$，所以原点是鞍点。$p(t) = t^4$ 在原点有极小值而 $p''(0) = 0$，说明 Hessian 正定不是极小值的必要条件。

第 27 讲

相似矩阵与 Jordan 标准形

Similar Matrices and Jordan Form

对应材料：MIT 18.06SC Session 3.4 课时摘要与 Strang 第 28 讲视频 Similar Matrices and Jordan Form；习题 28.1、28.2 及解答（Strang《Introduction to Linear Algebra》6.6 节第 12、20 题）；补充材料依 Strang 教材 6.6 节。

本讲目标

(1) 会证明 $A^{\mathsf{T}}A$ 半正定，并证明它正定当且仅当 A 的列线性无关；知道正定矩阵的逆、和仍正定，且可不交换行地消元。

(2) 掌握相似的定义 $B = M^{-1}AM$，从同一线性变换的换基公式推出相似，并逐列核对具体的换基计算。

(3) 证明相似矩阵有相同的特征多项式、特征值和特征空间维数，辨认迹、行列式、秩等相似不变量。

(4) 用 2×2 的重根例子说明重复特征值不决定相似类，也不决定是否可对角化。

(5) 读懂 Jordan 块与 Jordan 标准形的结论，解释块大小怎样记录特征向量，并验证具体的 Jordan 链和相似变换。

27.1 先补一块：$A^{\mathsf{T}}A$ 是正定的

Strang 在这一讲开头先回到正定矩阵。上一讲给出的判别法（特征值、主元、顺序主子式、能量 $x^{\mathsf{T}}Ax > 0$）都是对一个给定的对称矩阵来用的。实际问题里正定矩阵往往以 $A^{\mathsf{T}}A$ 的形式出现：最小二乘的法方程 $A^{\mathsf{T}}A\hat{x} = A^{\mathsf{T}}b$ 就是这样，其中 A 是 $m \times n$ 的长方形矩阵。这一节回答两个问题：正定性在哪些运算下保持，以及 $A^{\mathsf{T}}A$ 什么时候正定。

命题 27.1（逆与和） 设 A, B 都是 $n \times n$ 对称正定矩阵，则

(1) A^{-1} 对称正定；

(2) $A + B$ 对称正定。

证明 (1) $(A^{-1})^{\mathsf{T}} = (A^{\mathsf{T}})^{-1} = A^{-1}$，所以 A^{-1} 对称。A 的特征值 $\lambda_1, \dots, \lambda_n$ 全为正，A^{-1} 的特征值是 $1/\lambda_1, \dots, 1/\lambda_n$，也全为正。

(2) $(A + B)^{\mathsf{T}} = A^{\mathsf{T}} + B^{\mathsf{T}} = A + B$，所以 $A + B$ 对称。$A + B$ 的特征值和 A、B 的特征值

没有简单关系，这里改用能量判别：对任意 $x \neq 0$，

$$x^{\mathsf{T}}(A+B)x = x^{\mathsf{T}}Ax + x^{\mathsf{T}}Bx > 0 + 0 = 0.$$

■

命题 27.2（正定矩阵的主元）　设 A 是实对称正定矩阵。对 A 作普通高斯消元，无须交换行，且每个主元都为正数。

证明　对阶数归纳。1×1 时 $A=[a]$ 且 $a>0$。当 $n\geq 2$ 时写成

$$A = \begin{bmatrix} a & b^{\mathsf{T}} \\ b & C \end{bmatrix}, \qquad a>0, \quad b\in\mathbb{R}^{n-1}.$$

消去第一列后余下的对称矩阵是 $S = C - bb^{\mathsf{T}}/a$。对任意 $y\neq 0$，取 $z = (-b^{\mathsf{T}}y/a,\, y)^{\mathsf{T}} \neq 0$，则直接展开得 $z^{\mathsf{T}}Az = y^{\mathsf{T}}Sy > 0$，所以 S 仍正定。归纳假设可对 S 不交换行消元且得到正主元；首个主元 a 也为正，结论成立。 ■

这个结论说主元不会为零，并不表示主元在浮点计算中一定「足够大」。若矩阵只半正定，主元可以为零，例如 $\operatorname{diag}(1,0)$；因此正定条件不能删掉。

第 (2) 条说明，判别法要挑着用：证明 A^{-1} 正定时看特征值最方便，证明 $A+B$ 正定时看能量最方便。

定理 27.3（$A^{\mathsf{T}}A$ 的正定性）　设 A 是 $m\times n$ 实矩阵，则 $A^{\mathsf{T}}A$ 是 $n\times n$ 对称矩阵，并且

(1) 对任意 x，$x^{\mathsf{T}}(A^{\mathsf{T}}A)x = \|Ax\|^2 \geq 0$，即 $A^{\mathsf{T}}A$ 半正定；

(2) $A^{\mathsf{T}}A$ 正定 $\iff$ A 的 n 列线性无关 $\iff$ $\operatorname{rank} A = n$。

证明　$(A^{\mathsf{T}}A)^{\mathsf{T}} = A^{\mathsf{T}}(A^{\mathsf{T}})^{\mathsf{T}} = A^{\mathsf{T}}A$，对称。移动括号得到 $x^{\mathsf{T}}A^{\mathsf{T}}Ax = (Ax)^{\mathsf{T}}(Ax) = \|Ax\|^2 \geq 0$。等号成立当且仅当 $Ax=0$。若 A 的列线性无关，则 $Ax=0$ 只有零解，于是 $x\neq 0$ 时 $\|Ax\|^2>0$，$A^{\mathsf{T}}A$ 正定。反之若列线性相关，取 $x\neq 0$ 使 $Ax=0$，则 $x^{\mathsf{T}}A^{\mathsf{T}}Ax = 0$，$A^{\mathsf{T}}A$ 不是正定的。 ■

■ **例 27.1　列独立与 $A^{\mathsf{T}}A$ 的正定性**　设 $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$，并令 $C = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$。分别求 $A^{\mathsf{T}}A$、$C^{\mathsf{T}}C$，判断它们的正定性。

解　按定理把二次型化为向量长度。计算得

$$A^{\mathsf{T}}A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \qquad C^{\mathsf{T}}C = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}.$$

A 的两列 $(1,1,0)$、$(0,1,1)$ 不成比例，所以线性无关，定理 27.3 保证 $A^{\mathsf{T}}A$ 正定。

数值上，$\det(A^{\mathsf{T}}A) = 3 > 0$，且左上角元素为 $2>0$，也符合第 24 讲的二阶判别。

记 $G = A^{\mathsf{T}}A$。对它求逆并与单位阵相加：

$$G^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \qquad G + I = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

两者都对称正定：G^{-1} 由前面的命题保证；$G+I$ 对任意 $x \neq 0$ 满足 $x^{\mathsf{T}}(G+I)x = x^{\mathsf{T}}Gx + \|x\|^2 > 0$。这也展示了证明和正定性时，二次型往往比重新求特征值更直接。

C 的第二列等于第一列，取 $x = (-1, 1)^{\mathsf{T}} \neq 0$ 就有 $Cx = 0$，因此 $x^{\mathsf{T}}C^{\mathsf{T}}Cx = \|Cx\|^2 = 0$。所以 $C^{\mathsf{T}}C$ 只有半正定，不是正定。若 $m < n$，$m \times n$ 矩阵不可能有 n 个线性无关的列，故 $A^{\mathsf{T}}A$ 必不正定。 ♦

想一想　若 $A^{\mathsf{T}}A$ 正定，能否推出 A 是方阵？

答　不能。上例的 A 是 3×2 矩阵，却有线性无关的两列，因此 $A^{\mathsf{T}}A$ 是 2×2 正定矩阵。正定条件要求列独立，不要求 A 为方阵。

27.2　相似矩阵就是换一组基

对角矩阵 $D = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$ 作用很容易看：第一坐标乘 3，第二坐标不变。矩阵 $B = \begin{bmatrix} 3 & 4 \\ 0 & 1 \end{bmatrix}$ 看起来更复杂，但它描述的可能是同一个线性变换，只是基不同。需要一个精确关系来表达「只是换了坐标」。

定义 27.4（相似矩阵）　设 A, B 是同一数域上的 $n \times n$ 矩阵。若存在可逆矩阵 M 使

$$B = M^{-1}AM,$$

就称 A 与 B **相似**（similar），记作 $A \sim B$。

这是一个等价关系：$A = I^{-1}AI$，所以自反；若 $B = M^{-1}AM$，则 $A = MBM^{-1}$，所以对称；若 $B = M^{-1}AM$ 且 $C = N^{-1}BN$，则 $C = (MN)^{-1}A(MN)$，所以传递。可逆性是定义中的关键条件：若 M 不可逆，M^{-1} 不存在，坐标也不能一一换回。

为什么是这个乘法顺序？设变换 T 在旧基 $\mathcal{E}$、新基 $\mathcal{F} = (v_1, \dots, v_n)$ 下的矩阵分别为 A, B。令 $M = [v_1 \ \cdots \ v_n]_{\mathcal{E}}$，列向量是新基在旧基下的坐标。

新坐标 y 对应的旧坐标是 $x = My$。在旧基下应用 A 后，用 M^{-1} 换回新基坐标：

$$By = M^{-1}A(My), \qquad B = M^{-1}AM.$$

这个等式也能按列读。B 的第 j 列是 $T(v_j)$ 在新基 $\mathcal{F}$ 下的坐标，因为 Be_j 正是第 j 列。图 27.1 用一个具体向量展示「旧坐标—新坐标—作用—换回」的顺序；矩阵乘积中相邻的 $M^{-1}M$ 会抵消。

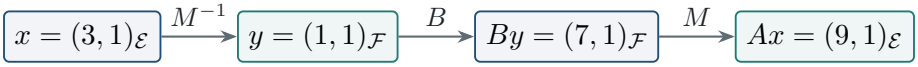


图 27.1　同一变换的输入与输出可用两组基表示；M^{-1} 与 M 负责换坐标，中间的 B 才是新基下的变换矩阵。

■ **例 27.2　逐列求新基下的矩阵**　令 $A=D=\begin{bmatrix}3&0\\0&1\end{bmatrix}$，新基为 $v_1=(1,0)^{\mathsf{T}}$、$v_2=(2,1)^{\mathsf{T}}$。求新基下的矩阵 B，并用 $M^{-1}AM$ 核对。

解　先算基向量的像，再把像写回新基坐标。$T(v_1)=(3,0)=3v_1$，所以 B 的第一列是 $(3,0)^{\mathsf{T}}$。而 $T(v_2)=(6,1)=4v_1+v_2$，所以第二列是 $(4,1)^{\mathsf{T}}$。因此

$$B=\begin{bmatrix}3&4\\0&1\end{bmatrix}.$$

把新基向量排成列，$M=\begin{bmatrix}1&2\\0&1\end{bmatrix}$，$M^{-1}=\begin{bmatrix}1&-2\\0&1\end{bmatrix}$。直接计算

$$M^{-1}AM=\begin{bmatrix}1&-2\\0&1\end{bmatrix}\begin{bmatrix}3&0\\0&1\end{bmatrix}\begin{bmatrix}1&2\\0&1\end{bmatrix}=\begin{bmatrix}3&4\\0&1\end{bmatrix}=B.$$

逐列方法把「矩阵的列是什么」说清楚；乘积方法则能快速检查结果。两种计算必须一致。　♦

定理 27.5（相似矩阵保留谱信息）　设 $A,B\in\mathbb{F}^{n\times n}$，其中 $\mathbb{F}=\mathbb{R}$ 或 $\mathbb{C}$。若 $B=M^{-1}AM$，M 可逆，则 A,B 的特征多项式相同；它们按代数重数计的特征值相同，并且每个特征值对应的特征空间维数相同。此外，$\operatorname{tr}A=\operatorname{tr}B$、$\det A=\det B$、$\operatorname{rank}A=\operatorname{rank}B$。

证明　先比较特征多项式。对任意标量 t，

$$tI-B=M^{-1}(tI-A)M,$$

故

$$\det(tI-B)=\det(M^{-1})\det(tI-A)\det(M)=\det(tI-A).$$

两个多项式完全相同，因此根及其重数都相同。

再比较特征空间。若 $Ax=\lambda x$ 且 $x\neq0$，则 $y=M^{-1}x\neq0$，并且

$$By=M^{-1}AMM^{-1}x=M^{-1}Ax=\lambda y.$$

反过来用 M 可从 B 的特征向量恢复 A 的特征向量，所以 M^{-1} 给出两个特征空间之间的线性双射，维数相同。

最后，行列式的乘法性给出 $\det B=\det A$。利用 $\operatorname{tr}(XY)=\operatorname{tr}(YX)$，有 $\operatorname{tr}B=\operatorname{tr}(M^{-1}AM)=\operatorname{tr}(AMM^{-1})=\operatorname{tr}A$。左乘可逆矩阵不改变列关系，右乘可逆矩阵只是对列作可逆线性组合，故秩也不变。　■

例 27.2 中，D 和 B 的迹都是 4、行列式都是 3、秩都是 2，但 D 对称而 B 不对称，矩阵条目也不同。因此，对称性和具体条目不由一般相似变换保持。若 A 对称且 M 正交，则 $B = M^{\mathsf{T}}AM$ 仍对称；本讲的相似定义不要求 M 正交。

反过来并不成立：几个不变量取值相同，只能说明还没有被它们区分，不能据此断定相似。下一节会看到，特征值、迹、行列式都相同的两个矩阵，仍可能因特征空间维数不同而不相似。

想一想　相似矩阵的特征向量必须是同一个坐标向量吗?

答　不必。$D = \operatorname{diag}(3,1)$ 的特征向量 e_2 属于特征值 1；其相似矩阵 $B = \begin{bmatrix} 3 & 4 \\ 0 & 1 \end{bmatrix}$ 的该特征空间是 $\operatorname{span}\{(-2,1)\}$，而 $Be_2 = (4,1)$，不是 e_2 的倍数。相似变换保持特征空间维数，并由 M^{-1} 把向量坐标对应起来；它不要求两套坐标里出现相同的向量。

同一个抵消也适用于矩阵的幂，因此相似关系能传递到由矩阵多项式定义的运算。

命题 27.6（相似与多项式）　若 $A \sim B$，则对任意一元多项式 p，有 $p(A) \sim p(B)$。

证明　设 $B = M^{-1}AM$。由归纳法，$B^k = M^{-1}A^kM$ 对每个整数 $k \geq 0$ 成立；$k = 0$ 时是 $I = M^{-1}IM$，归纳步中间的 MM^{-1} 抵消。若 $p(t) = \sum_{k=0}^{d} c_k t^k$，则

$$p(B) = \sum_{k=0}^{d} c_k B^k = M^{-1}\left(\sum_{k=0}^{d} c_k A^k\right) M = M^{-1}p(A)M.$$

因此 $p(A)$ 与 $p(B)$ 相似。$p(t) = t^2$ 就是教材习题 28.2(a)。 ■

若 A 有一组特征向量基，换基矩阵会把它直接变成对角矩阵。这是对角化的矩阵写法，也是相似关系最重要的一个应用。

命题 27.7（特征向量基给出对角矩阵）　设 $A \in \mathbb{F}^{n\times n}$，并有一组基 $v_1, \ldots, v_n$ 满足 $Av_j = \lambda_j v_j$。令 $S = [v_1 \ \cdots \ v_n]$，则

$$S^{-1}AS = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

证明　按列计算得到 $AS = [Av_1 \ \cdots \ Av_n] = [\lambda_1 v_1 \ \cdots \ \lambda_n v_n] = S\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$。$S$ 的列是基，故 S 可逆；左乘 S^{-1} 即得结论。 ■

■ **例 27.3　相似到对角矩阵**　对 $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$，验证 $v_1 = (1,1)^{\mathsf{T}}$、$v_2 = (1,-1)^{\mathsf{T}}$ 给出一组特征向量基，并写出对角化。

解　相乘得 $Av_1 = (3,3) = 3v_1$、$Av_2 = (1,-1) = v_2$。两向量不成比例，所以构成一组基。令

$S=[v_1\ v_2]=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$，则

$$S^{-1}=\frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix},\qquad S^{-1}AS=\begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}.$$

这里相似矩阵是对角矩阵，表示沿两条特征方向分别乘 3 和 1。有 n 个互异特征值时，前一讲的独立性定理保证这些向量构成基；特征值重复时则要检查特征空间是否提供足够多的独立向量。 ♦

■ **例 27.4　两个不同特征值下的显式相似矩阵**　令 $D=\begin{bmatrix}3 & 0\\ 0 & 4\end{bmatrix}$、$M=\begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix}$。求 $M^{-1}DM$，并说明对角线以外的 1 是否改变了特征值。

解　先求 $M^{-1}=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$，再按定义相乘：

$$M^{-1}DM=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 4\end{bmatrix}\begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix}=\begin{bmatrix}3 & 1\\ 0 & 4\end{bmatrix}.$$

右上角从 0 变成 1，但特征多项式仍是 $(t-3)(t-4)$，与定理 27.5 相符。相似变换能改变矩阵条目，保留谱信息。 ♦

■ **例 27.5　同时交换一行和一列**　取 $A=\begin{bmatrix}1 & 2 & 0\\ 3 & 4 & 5\\ 6 & 7 & 8\end{bmatrix}$。先交换第 $1,2$ 行，再交换第 $1,2$ 列，求所得矩阵并写成相似变换。

解　令 $P=\begin{bmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}$。左乘 P 交换前两行；$P^{\mathsf T}=P$，所以右乘 P 再交换前两列。直接计算得

$$PAP=\begin{bmatrix}4 & 3 & 5\\ 2 & 1 & 0\\ 7 & 6 & 8\end{bmatrix}.$$

又 $P^2=I$，因此 $P^{-1}=P$，所得矩阵就是 $P^{-1}AP$，与 A 相似。单独交换行或单独交换列一般不是相似变换；成对使用同一个置换矩阵才得到这里的共轭形式。 ♦

27.3　Jordan 块记录特征向量不足

对四阶零特征值矩阵，秩和零空间维数还不够描述相似类。考虑

$$J=\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix},\qquad K=\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}.$$

两者都有特征值 0 的代数重数 4、秩 2，所以零空间维数都是 2，特征空间也都是二维。J 的非零作用是 $Je_2 = e_1$、$Je_4 = e_3$，所以它分成两条长度为 2 的链；K 则有 $Ke_3 = e_2$、$Ke_2 = e_1$，另有 $Ke_4 = 0$，分成长度为 3 和 1 的链。矩阵平方会保留这个差异：$J^2 = 0$，但 $K^2 e_3 = e_1$，故 $\operatorname{rank}(K^2) = 1$。由定理 27.5 的秩不变性，$J$ 与 K 不相似。

想一想　已知两个矩阵的特征值、秩、零空间维数都相同，是否足以判断相似？

答　不足。上面的 J, K 都只有特征值 0（代数重数 4），秩都是 2，零空间维数也都是 2，但 $\operatorname{rank}(J^2) = 0$、$\operatorname{rank}(K^2) = 1$，所以它们不相似。相似会保留每个幂的秩，因为 $(M^{-1}AM)^k = M^{-1}A^kM$；单看 A 本身的秩会丢掉链长度信息。

■ **例 27.6　同秩、同零空间维数仍可能不相似**　对上面的 $J = J_2(0) \oplus J_2(0)$ 与 $K = J_3(0) \oplus J_1(0)$，分别求零空间维数和平方的秩，并据此判断相似性。

解　直接作用可得 $Jx = (x_2, 0, x_4, 0)^{\mathsf{T}}$，故 $Jx = 0$ 等价于 $x_2 = x_4 = 0$，零空间维数为 2。同理 $Kx = (x_2, x_3, 0, 0)^{\mathsf{T}}$，所以 $Kx = 0$ 等价于 $x_2 = x_3 = 0$，零空间维数也为 2。而 $J^2 = 0$，所以 $\operatorname{rank}(J^2) = 0$；$K^2x = (x_3, 0, 0, 0)^{\mathsf{T}}$，因此 $\operatorname{rank}(K^2) = 1$。若 $K = M^{-1}JM$，则 $K^2 = M^{-1}J^2M = 0$，与 $K^2 \neq 0$ 矛盾。故两矩阵不相似；它们的块大小分别是 $2+2$ 与 $3+1$。 ♦

这种链结构由 Jordan 块表示。对正整数 r，定义

定义 27.8（Jordan 块）　数域上的 $r \times r$ 矩阵

$$J_r(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}$$

称为特征值为 λ、大小为 r 的 **Jordan 块**（Jordan block）。当 $r = 1$ 时，$J_1(\lambda) = [\lambda]$。

写 $J_r(\lambda) = \lambda I + N_r$，其中 N_r 只在主对角线上方紧邻的位置有 1。以 $r = 3$ 为例，标准基满足

$$N_3 e_1 = 0, \qquad N_3 e_2 = e_1, \qquad N_3 e_3 = e_2.$$

因此 e_1 是特征向量，e_2, e_3 是逐步映到前一向量的广义方向。一般地，若一组向量满足

$$(A - \lambda I)v_1 = 0, \qquad (A - \lambda I)v_j = v_{j-1} \quad (2 \le j \le r), \qquad v_1 \neq 0,$$

若 $v \neq 0$ 且存在正整数 k 使 $(A-\lambda I)^k v = 0$，则称 v 为 λ 的**广义特征向量**（generalized eigenvector）；上式中的每个 v_j 都是这样的向量。满足上式的一组向量称为一条长度 r 的 **Jordan 链**（Jordan chain）。按列排成基矩阵时，这条链给出 $J_r(\lambda)$ 的上三角对角块。

图 27.2 展开了 $J_3(\lambda)$ 中的三步关系。箭头表示先减去特征值的标量作用，即应用 $N_3 = J_3(\lambda) - \lambda I$；链的末端 e_1 已是特征向量。

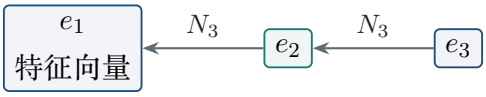


图 27.2　$J_3(\lambda)-\lambda I$ 把 e_3 送到 e_2、把 e_2 送到 e_1，而 e_1 被送到零。

同一块里只有一个独立特征向量：$(J_r(\lambda)-\lambda I)x=0$ 逐行给出 $x_2=\cdots=x_r=0$，故特征空间是 $\operatorname{span}\{e_1\}$。块的大小是 r，却只贡献一个独立特征向量；这正是 Jordan 块记录「缺少方向」的方式。

把若干 Jordan 块排在对角线上，其余位置填零，所得分块对角矩阵称为 Jordan 矩阵。例如上面的 J 是 $J_2(0)\oplus J_2(0)$，而 K 是 $J_3(0)\oplus J_1(0)$。符号 $\oplus$ 表示把两个矩阵放在对角线上组成分块对角矩阵。

特征多项式中因子 $(t-\lambda)$ 的指数称为 λ 的**代数重数**（algebraic multiplicity）；特征空间维数 $g_\lambda=\dim N(A-\lambda I)$ 称为 λ 的**几何重数**（geometric multiplicity）。

定理 27.9（Jordan 标准形定理）　设 $A\in\mathbb{C}^{n\times n}$。存在可逆复矩阵 S，使

$$S^{-1}AS=J=\operatorname{diag}\bigl(J_{r_1}(\lambda_1),\dots,J_{r_q}(\lambda_q)\bigr)$$

为 Jordan 矩阵；除去 Jordan 块的排列顺序后，这些块由 A 唯一确定。对每个特征值 λ，所有对应块的大小之和等于 λ 的代数重数，对应块的个数等于 $\dim N(A-\lambda I)$。A 可对角化当且仅当所有 Jordan 块大小都为 1。

因此每个特征值都满足 $1\le g_\lambda\le m_\lambda$；等号 $g_\lambda=m_\lambda$ 当且仅当该特征值对应的所有 Jordan 块大小都是 1。

只知道代数重数和特征空间维数，通常还不能恢复每个 Jordan 块的大小。若一个特征值的代数重数为 4、特征空间维数为 2，块大小既可能是 $3+1$，也可能是 $2+2$；两种排列都有两个块，块大小之和都是 4。要区分它们，还要看更细的链结构。

想一想　两矩阵的代数重数与特征空间维数都相同，能否断定它们的 Jordan 块大小相同？
答　不能。上面的 $J=J_2(0)\oplus J_2(0)$ 与 $K=J_3(0)\oplus J_1(0)$ 都有代数重数 4、特征空间维数 2，但块大小分别为 $2+2$ 与 $3+1$。计算 J^2、K^2 的秩可以区分它们；更一般的块结构需要比特征值和特征空间维数更细的信息。

本讲不证明一般存在性与唯一性，也不提供对任意矩阵求出 S 和全部块大小的通用算法。我们只从给定的块矩阵读出两项计数：一个大小为 r 的块贡献特征多项式因子 $(t-\lambda)^r$，并贡献一个独立特征向量；因此块的大小总和记录代数重数，块数记录特征空间维数。若 n 阶 Jordan 形共有 d 个块，大小分别为 $r_1,\dots,r_d$，则上超对角线共有 $\sum_i(r_i-1)=n-d$ 个 1。实矩阵若有非实特征值，以上相似关系要在复数范围内理解；若其特征值全为实数，可在实数范围内选取 Jordan 基。

当所有特征值互异时，每个特征空间维数为 1 且代数重数为 1，所以每个 Jordan 块只能是 1×1，Jordan 形就是对角形。重根若仍有足够多的独立特征向量，也会给出多个大小为 1 的块；

出现更大的块，才表示该特征值下的特征向量不足。

■ **例 27.7　从 Jordan 链写出相似矩阵**　设

$$B=\begin{bmatrix}0&1&7&0\\0&0&1&0\\0&0&0&0\\0&0&0&0\end{bmatrix}.$$

构造一组基，使 B 成为 Jordan 矩阵。

解　把 B 看成幂零矩阵，先从 e_3 向前追像：$Be_3=7e_1+e_2$，再作用一次得到 $B^2e_3=e_1$，第三次为零。因此取链

$$v_1=e_1,\qquad v_2=7e_1+e_2,\qquad v_3=e_3.$$

另有 $Be_4=0$，所以取 $v_4=e_4$ 作一个 1×1 的零块。逐个计算：$Bv_1=0$、$Bv_2=v_1$、$Bv_3=v_2$、$Bv_4=0$。这些向量排成

$$P=[v_1\ v_2\ v_3\ v_4]=\begin{bmatrix}1&7&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix},$$

是可逆矩阵，因为它是上三角且对角线全为 1，并且

$$P^{-1}=\begin{bmatrix}1&-7&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$$

按列的像关系，$BP=P(J_3(0)\oplus J_1(0))$，于是

$$P^{-1}BP=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0\end{bmatrix}.$$

矩阵中的 7 由链向量 v_2 吸收；换基后它不再出现。这个例子给出一个具体的 Jordan 基，不代表任意矩阵都能照同一短步骤处理。　♦

这里 B 的特征多项式为 t^4，而 $N(B)=\operatorname{span}\{e_1,e_4\}$，维数为 2。Jordan 形中的两个块大小为 3 和 1，其和为 4、块数为 2，分别对应代数重数与特征空间维数。由于有一个大小大于 1 的块，B 不可对角化。

27.4　重根为什么会分成不同的相似类

特征值为 $4,4$ 的两个 2×2 矩阵是最小的对照。标量矩阵 $4I$ 的每个非零向量都是特征向量；上三角矩阵

$$J_2(4)=\begin{bmatrix}4 & 1\\ 0 & 4\end{bmatrix}$$

只有一条特征方向。二者的特征多项式都是 $(t-4)^2$，迹都是 8，行列式都是 16，但特征空间维数分别为 2 和 1。由定理 27.5，它们不相似。

■ **例 27.8　判断两个重根矩阵能否相似**　比较 $4I_2$ 与 $J_2(4)$，并求二者属于特征值 4 的全部特征向量。

解　先求特征空间。对 $4I_2$，方程 $(4I_2-4I_2)x=0$ 对所有 x 成立，所以任意非零向量都是特征向量。对 $J_2(4)$，

$$(J_2(4)-4I_2)x=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=0$$

要求 $x_2=0$，故特征空间是 $\operatorname{span}\{(1,0)\}$。维数不同，矩阵不相似。这个例子说明：特征值重复时，代数重数 2 不等于特征空间维数；维数可以只有 1。 ♦

易错提醒　「特征值相同就相似」是错误的。$4I_2$ 与 $J_2(4)$ 有相同的特征值、迹和行列式，却因特征空间维数不同而不相似。比较相似矩阵时，特征值相同只是必要条件，不能替代相似变换或更强的不变量检查。

在二阶、特征值固定为 4 的情形，可以把差别说得更完整。

定理 27.10（二阶重根的两种相似类）　设 $A\in\mathbb{F}^{2\times 2}$，其中 $\mathbb{F}=\mathbb{R}$ 或 $\mathbb{C}$，且特征多项式为 $(t-4)^2$。则 A 与 $4I_2$ 相似，或与 $J_2(4)$ 相似；只有这两种可能，且二者互不同类。

证明　特征多项式为 $(t-4)^2$ 等价于 $\operatorname{tr}A=8$、$\det A=16$。令 $N=A-4I$。于是 $\operatorname{tr}N=0$，且 $\det N=\det(A-4I)=0$。写成 $N=\begin{bmatrix}a & b\\ c & -a\end{bmatrix}$，由行列式为零得 $a^2+bc=0$，从而

$$N^2=\begin{bmatrix}a^2+bc & 0\\ 0 & a^2+bc\end{bmatrix}=0.$$

若 $N=0$，则 $A=4I_2$。若 $N\neq 0$，取 v 使 $Nv\neq 0$，令 $w=Nv$。则 $Nw=N^2v=0$，而 v,w 线性无关：若 $w=cv$，再次作用 N 会得到 $0=Nw=cw$，与 $w\neq 0$ 矛盾（$c=0$ 时也与 $w=cv\neq 0$ 矛盾）。在基 (w,v) 下，$Aw=4w$、$Av=w+4v$，矩阵正是 $J_2(4)$。它与 $4I_2$ 的特征空间维数不同，故二者不相似。 ■

■ **例 27.9　把非标准矩阵化到 $J_2(4)$**　设 $A=\begin{bmatrix}5 & 1\\ -1 & 3\end{bmatrix}$。从定理 27.10 的证明中构造一组基，使

A 的矩阵成为 $J_2(4)$。

解　取 $N = A - 4I = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$，用 $v = e_2 = (0,1)^{\mathsf{T}}$ 开始链：$Nv = (1,-1)^{\mathsf{T}} =: w$，且 $Nw = 0$。所以 (w, v) 是基，因为两向量不成比例。令

$$M = [w\ v] = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \qquad M^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

按基向量的像，$Aw = 4w$、$Av = w + 4v$，故

$$M^{-1}AM = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix} = J_2(4).$$

这里第二列的 1 来自 $Nv = w$，表示 v 还不是特征向量，但经过 $A - 4I$ 一次就落到特征向量 w 上。 ♦

■ **例 27.10　Jordan 块里的非零系数可以缩放**　令 $J = J_2(4)$、$M = \mathrm{diag}(1,2)$。计算 $M^{-1}JM$，并解释右上角系数 2 是否产生了新的相似类。

解　$M^{-1} = \mathrm{diag}(1, \frac{1}{2})$，所以

$$M^{-1}JM = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 0 & 4 \end{bmatrix}.$$

矩阵仍与 $J_2(4)$ 相似，因此右上角的任意非零系数都可由对角换基缩放为 1；系数为 0 时则得到 $4I_2$，落入另一相似类。 ♦

检验相似的顺序　先找显式换基矩阵，或检查两边是否都有一组特征向量基；找不到时，先比特征多项式和特征空间维数。若这些不变量仍相同，不能据此宣布相似；Jordan 块大小才记录更细的链结构。

前后联系　相似变换 $M^{-1}AM$ 用同一组基改变输入和输出的坐标；下一讲的奇异值分解会分别选择定义域与值域的正交基，因此也适用于长方形矩阵。

本讲小结

- 对实矩阵 A，$x^{\mathsf T}A^{\mathsf T}Ax=\|Ax\|^2$；$A^{\mathsf T}A$ 正定当且仅当 A 的列线性无关。
- 若 A,B 是同一线性变换在两组基下的矩阵，则 $B=M^{-1}AM$，其中 M 可逆。
- 相似矩阵有相同的特征多项式、迹、行列式、秩，以及各特征值对应的特征空间维数；这些不变量相同不保证相似。
- 若 A 有特征向量基 $S=[v_1\ \cdots\ v_n]$，则 $S^{-1}AS=\operatorname{diag}(\lambda_1,\dots,\lambda_n)$。
- 复矩阵的 Jordan 形由块 $J_r(\lambda)$ 组成；对每个 λ，块大小之和是代数重数，块数是 $\dim N(A-\lambda I)$。
- $J_r(\lambda)$ 大小 r 的块只有一个独立特征向量；全部块大小为 1 当且仅当矩阵可对角化。

练习

27.1 ★ 判断真假并说明理由：(a) 若 $A\sim B$，则 $A^2\sim B^2$；(b) 若 $A^2\sim B^2$，则 $A\sim B$。假的命题给出反例。

27.2 ★★ 设 $D=\begin{bmatrix}3&0\\0&4\end{bmatrix}$、$M=\begin{bmatrix}1&-1\\0&1\end{bmatrix}$。计算 $M^{-1}DM$，并核对其特征值。

27.3 ★★ 证明 $\operatorname{diag}(3,3)$ 与 $\begin{bmatrix}3&1\\0&3\end{bmatrix}$ 不相似。指出只比较特征多项式为什么不够。

27.4 ★★ 对任意 $n\times n$ 矩阵 A，交换第 $1,2$ 行，再交换第 $1,2$ 列得到 B。证明 $A\sim B$。

27.5 ★★★ 令 $J=J_2(0)\oplus J_2(0)$、$K=J_3(0)\oplus J_1(0)$。设 $M=(m_{ij})$ 满足 $JM=MK$。写出两边的矩阵并证明 M 第一列全为零；由此证明 J 与 K 不相似。

27.6 ★ 令 $D=\operatorname{diag}(2,5)$，$M=\begin{bmatrix}1&1\\0&1\end{bmatrix}$。求 $M^{-1}DM$，并说明它与 D 的特征向量是否必须用同一坐标向量表示。

27.7 ★★ 设 $C=J_2(2)\oplus J_1(2)$。求特征多项式与 $\dim N(C-2I)$，判断 C 是否可对角化，并解释理由。

27.8 ★★ 用定理 27.10 说明：特征多项式为 $(t-4)^2$ 的实 2×2 矩阵有哪两种相似类；分别给出一个代表矩阵。

27.9 ★ 对 $J_3(2)$，验证 e_3 是广义特征向量，并求使 $(J_3(2)-2I)^k e_3=0$ 成立的最小正整数 k。

27.10 ★ 某特征值 λ 对应的 Jordan 块大小为 $4,2,1$。求它的代数重数和几何重数，并判断这些块是否给出对角矩阵。

练习答案与提示

27.1 (a) 真。若 $A=M^{-1}BM$，则

$$A^2=M^{-1}BMM^{-1}BM=M^{-1}B^2M,$$

所以 $A^2 \sim B^2$。(b) 假：取 $A = 0$、$B = J_2(0) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$，则 $A^2 = B^2 = 0$。但 $\operatorname{rank} A = 0$、$\operatorname{rank} B = 1$，相似矩阵秩相同，故 $A \not\sim B$。

27.2 $M^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$，所以

$$M^{-1}DM = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 0 & 4 \end{bmatrix}.$$

其特征多项式为 $(t-3)(t-4)$，特征值仍是 $3, 4$。

27.3 两矩阵的特征多项式都是 $(t-3)^2$。对 $3I_2$，特征空间为整个 $\mathbb{R}^2$，维数 2；对 $J_2(3)$，$J_2(3) - 3I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$，故特征空间是 $\operatorname{span}\{(1,0)\}$，维数 1。相似会保持特征空间维数，因此二者不相似。特征多项式只记录代数重数，未记录特征空间维数。

27.4 令 P 为交换前两个坐标的置换矩阵，则 $P^{-1} = P$。左乘 P 交换两行，右乘 P 交换两列，因此 $B = PAP = P^{-1}AP$，按定义 $A \sim B$。

27.5 由 $JM = MK$ 得

$$JM = \begin{bmatrix} m_{21} & m_{22} & m_{23} & m_{24} \\ 0 & 0 & 0 & 0 \\ m_{41} & m_{42} & m_{43} & m_{44} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad MK = \begin{bmatrix} 0 & m_{11} & m_{12} & 0 \\ 0 & m_{21} & m_{22} & 0 \\ 0 & m_{31} & m_{32} & 0 \\ 0 & m_{41} & m_{42} & 0 \end{bmatrix}.$$

比较第二行得 $m_{21} = m_{22} = 0$；比较第一行进而得 $m_{11} = 0$。比较第四行得 $m_{41} = m_{42} = 0$；比较第三行进而得 $m_{31} = 0$。因此 M 的第一列为零，不可逆。若 J, K 相似，就存在可逆 M 使 $JM = MK$，矛盾；所以不相似。

27.6 $M^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$，从而

$$M^{-1}DM = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ 0 & 5 \end{bmatrix}.$$

D 的 5-特征向量可取 e_2；对应新坐标为 $M^{-1}e_2 = (-1, 1)$，且新矩阵把该向量乘以 5。它不要求仍用 e_2 表示：$M^{-1}DMe_2 = (-3, 5)$ 不是 e_2 的倍数。

27.7 $C - 2I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$，故特征空间由 $x_2 = 0$ 给出，维数为 2。块的特征多项式相乘为 $(t-2)^2(t-2) = (t-2)^3$，代数重数为 3。特征空间维数小于 3，没有三条独立特征向量，因此 C 不可对角化；其 Jordan 块大小是 2 和 1。

27.8 由定理 27.10，只有 $4I_2$ 与 $J_2(4)$ 两种可能。二者都是该多项式的代表：$4I_2$ 的特征空间维数为 2，$J_2(4)$ 的特征空间维数为 1，所以它们属于不同相似类。非标量情形可在定理证明中取 $N = A - 4I$，选 v 使 $Nv \neq 0$，再取基 (Nv, v)，得到第二种代表。

27.9 令 $N = J_3(2) - 2I$，则 $Ne_3 = e_2$、$N^2e_3 = e_1$、$N^3e_3 = 0$。因 $N^2e_3 \neq 0$，最小的 k 是 3。

27.10 代数重数为 $4 + 2 + 1 = 7$，几何重数为 Jordan 块数 3。由于有大小大于 1 的块，这些块不组成对角矩阵。

第 28 讲

奇异值分解

Singular Value Decomposition

对应材料：MIT 18.06SC Part 7；ZoomNotes 7.1–7.3；习题 29.1、29.2。

本讲目标

(1) 会从 $A^{\mathsf T}A$ 的特征值与标准正交特征向量求奇异值和右奇异向量。

(2) 理解任意实矩阵都有奇异值分解，并能说清证明中为何只对 $A^{\mathsf T}A$ 使用谱定理。

(3) 会验证 $A = U\Sigma V^{\mathsf T}$，并从 $Av_i = \sigma_i u_i$ 读出输入、输出方向。

(4) 能用零奇异值描述零空间、左零空间，以及行空间和列空间。

(5) 会处理 Fibonacci 矩阵与正交列矩阵的两个原课程习题。

28.1 哪些方向会被送到正交方向?

先看

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}.$$

标准基向量的像是 $(3,4)^{\mathsf T}$ 和 $(0,5)^{\mathsf T}$，点积为 20，并不垂直。若能换一组标准正交的输入方向，使它们的像也互相垂直，矩阵的作用就会简单得多。

■ 例 28.1 **寻找两组正交方向** 对上面的矩阵，验证 $v_1 = (1,1)^{\mathsf T}/\sqrt{2}$、$v_2 = (-1,1)^{\mathsf T}/\sqrt{2}$ 是标准正交向量，并计算它们的像。

解 先检查输入：$v_1^{\mathsf T}v_1 = v_2^{\mathsf T}v_2 = 1$，且 $v_1^{\mathsf T}v_2 = 0$。逐列计算

$$Av_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 3 \\ 9 \end{bmatrix}, \qquad Av_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -3 \\ 1 \end{bmatrix}.$$

两像的点积是 $(-9+9)/2 = 0$。它们的长度分别为 $3\sqrt{5}$ 和 $\sqrt{5}$。因此令

$$u_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \qquad u_2 = \frac{1}{\sqrt{10}}\begin{bmatrix} -3 \\ 1 \end{bmatrix},$$

便有 $Av_1 = 3\sqrt{5}\,u_1$、$Av_2 = \sqrt{5}\,u_2$。本例已经给出一个 SVD 的全部列。 ♦

定义 28.1（奇异值与奇异向量）　设 $A \in \mathbb{R}^{m\times n}$。矩阵 $A^{\mathsf{T}}A$ 的特征值都非负；它们的非负平方根称为 A 的**奇异值**（singular value）。若单位向量 v 满足 $A^{\mathsf{T}}Av = \sigma^2 v$，就称 v 为对应的**右奇异向量**（right singular vector）。当 $\sigma > 0$ 时，令 $u = Av/\sigma$，称 u 为对应的**左奇异向量**（left singular vector）。

非负性可直接从特征方程验证：若 $A^{\mathsf{T}}Av = \lambda v$ 且 $v \neq 0$，则 $\lambda\|v\|^2 = \|Av\|^2 \geq 0$。因此平方根确实是实数。

对例 28.1，$A^{\mathsf{T}}A = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix}$，特征值是 $45, 5$，所以奇异值为 $3\sqrt{5}, \sqrt{5}$。上一例的 v_1, v_2 正是右奇异向量。相应的 u_1, u_2 也标准正交。

对 $\sigma > 0$，两个关系同时成立：

$$Av = \sigma u, \qquad A^{\mathsf{T}}u = \sigma v.$$

第二式由 $A^{\mathsf{T}}u = A^{\mathsf{T}}Av/\sigma = \sigma v$ 得到。因此 u 是 AA^{T} 对应于特征值 σ^2 的单位特征向量。零奇异值时 $Av = 0$；此时不能用 Av/σ 定义 u。

想一想　若 $\sigma = 0$，能否仍用 $u = Av/\sigma$ 找左奇异向量？
答　不能除以 0。零奇异方向分别由 $N(A)$ 和 $N(A^{\mathsf{T}})$ 补出；二者维数是 $n - r$、$m - r$，可能不同，所以不能总把左右零方向逐个配对。第 5 节会用秩亏例子具体说明。

一个边界例子是零矩阵：它的奇异值全为 0，任意标准正交基都可作右奇异向量；这时特征向量并不唯一。再看非零矩阵 $N = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$，单位向量 e_1 在 N 下被送到 0，所以它对应奇异值 0；e_2 对应奇异值 2。

易错提醒　奇异值不等于一般矩阵的特征值。上面的 N 满足 $N^2 = 0$，所以它的特征值全是 0；但 $N^{\mathsf{T}}N = \mathrm{diag}(0, 4)$，其奇异值为 $2, 0$。奇异值来自 $A^{\mathsf{T}}A$ 的谱，不能把 A 的特征分解条件直接套到任意矩阵上。

28.2　每个矩阵都有奇异值分解

无论 A 是方阵、长方形矩阵，还是不对称矩阵，$A^{\mathsf{T}}A$ 总是对称的，而且

$$x^{\mathsf{T}}A^{\mathsf{T}}Ax = \|Ax\|^2 \geq 0.$$

所以它可以使用第 24 讲的谱定理，特征值也不会为负。接下来的定理把这个观察变成对 A 本身的分解。

定理 28.2（奇异值分解定理）　设 $A \in \mathbb{R}^{m\times n}$，秩为 r。存在正交矩阵 $U \in \mathbb{R}^{m\times m}$、$V \in \mathbb{R}^{n\times n}$，以及 $m \times n$ 矩形对角矩阵 Σ，使

$$A = U\Sigma V^{\mathsf{T}}.$$

Σ 的对角元为 $\sigma_1 \geq \cdots \geq \sigma_r > 0$，其余位置为 0。这些正数就是 A 的全部正奇异值。

直观理解　先在输入空间选一组标准正交方向 v_i，使它们的像 Av_i 两两垂直。把非零像向量单位化，得到输出空间的标准正交方向 u_i。剩余输入方向被 A 消成零，可以用零对角元补齐。

思路　证明分三步：先用谱定理给 $A^{\mathsf{T}}A$ 找标准正交特征基；再将正特征值对应的像 Av_i 单位化，并验证它们正交；最后补全输出空间的标准正交基。特征值为零的方向需要直接证明 $Av_i = 0$，不能除以其奇异值。

证明　**第一步：选输入基并定义奇异值。**令 $G = A^{\mathsf{T}}A$。它是实对称矩阵，且对任意 x 有 $x^{\mathsf{T}}Gx = \|Ax\|^2 \geq 0$。由谱定理，取一组标准正交特征向量 $v_1, \ldots, v_n$，使

$$Gv_i = \lambda_i v_i, \qquad \lambda_1 \geq \cdots \geq \lambda_n \geq 0.$$

记 $\sigma_i = \sqrt{\lambda_i}$。若 $\lambda_i = 0$，则 $\|Av_i\|^2 = v_i^{\mathsf{T}}A^{\mathsf{T}}Av_i = 0$，所以 $Av_i = 0$。

零空间相同：$Ax = 0$ 会推出 $A^{\mathsf{T}}Ax = 0$；反过来，若 $A^{\mathsf{T}}Ax = 0$，则 $\|Ax\|^2 = x^{\mathsf{T}}A^{\mathsf{T}}Ax = 0$，故 $Ax = 0$。因此 $N(A^{\mathsf{T}}A) = N(A)$，秩—零度定理给出正特征值的个数 $r = \operatorname{rank} A$。

第二步：构造输出的标准正交向量。对 $1 \leq i \leq r$，定义 $u_i = Av_i/\sigma_i$。当 $1 \leq i, j \leq r$ 时，

$$u_i^{\mathsf{T}}u_j = \frac{v_i^{\mathsf{T}}A^{\mathsf{T}}Av_j}{\sigma_i\sigma_j} = \frac{\lambda_j v_i^{\mathsf{T}}v_j}{\sigma_i\sigma_j} = \delta_{ij}.$$

因此 $u_1, \ldots, u_r$ 标准正交。因为这些向量在 $\mathbb{R}^m$ 中，$r \leq m$；将它们补成标准正交基 $u_1, \ldots, u_m$。

第三步：写成矩阵等式。令 $V = [v_1 \ \cdots \ v_n]$、$U = [u_1 \ \cdots \ u_m]$，并令 $\Sigma_{ii} = \sigma_i$（$1 \leq i \leq r$），其余元为零。对 $i \leq r$，$Av_i = \sigma_i u_i$；对 $i > r$，$Av_i = 0$。所以两边第 i 列相等，得到 $AV = U\Sigma$。因为 V 正交，$V^{\mathsf{T}} = V^{-1}$；右乘 V^{T} 即得 $A = U\Sigma V^{\mathsf{T}}$。 ■

这段证明只对 $A^{\mathsf{T}}A$ 使用谱定理。它没有要求 A 对称、方阵或可对角化；A 的形状只决定 U, Σ, V 的尺寸。

$A^{\mathsf{T}}A$ 与 AA^{T} 的正特征值也完全相同，重数相同。若 $A^{\mathsf{T}}Av = \lambda v$ 且 $\lambda > 0$，令 $u = Av/\sqrt{\lambda}$，则 $AA^{\mathsf{T}}u = \lambda u$；反向从 $AA^{\mathsf{T}}u = \lambda u$ 出发，令 $v = A^{\mathsf{T}}u/\sqrt{\lambda}$ 即可。两个方向的零特征值数目分别是 $n - r$ 与 $m - r$，所以长方形矩阵的零特征值数目可能不同。

定理给出的分解称为**完整 SVD**（full SVD）：U 有 m 列，V 有 n 列。若只保留正奇异值对应的 r 列，则得到**紧 SVD**（reduced SVD）

$$A = U_r\Sigma_r V_r^{\mathsf{T}},$$

其中 $U_r \in \mathbb{R}^{m\times r}$、$\Sigma_r \in \mathbb{R}^{r\times r}$、$V_r \in \mathbb{R}^{n\times r}$。完整分解便于同时看见零空间和左零空间；紧分解省去全为零的部分。

完整分解的尺寸可以直接读出：U 是 $m\times m$，Σ 是 $m\times n$，V 是 $n\times n$。当 $m<n$ 时，V 仍有 n 列，而 U 只有 m 列；额外的 $v_{m+1},\dots,v_n$ 都在 $N(A)$ 中，Σ 对应的列为零。反过来，$m>n$ 时，额外的 $u_{n+1},\dots,u_m$ 属于 $N(A^{\mathsf T})$。

命题 28.3（正奇异值由矩阵唯一确定）　同一矩阵的奇异值及其重数不依赖于所选的 SVD；右、左奇异向量一般不唯一。

证明　任一 SVD 都给出 $A^{\mathsf T}A=V\Sigma^{\mathsf T}\Sigma V^{\mathsf T}$。因此 $A^{\mathsf T}A$ 的特征值是 σ_i^2 与零；按大小排列后取非负平方根，正奇异值及其重数就唯一确定。若某个特征值重复，对应特征空间可换标准正交基；而单个向量也可同时反号，故奇异向量未必唯一。 ■

例 28.2　秩二矩阵也有零奇异值　令

$$B=\begin{bmatrix}0&1&0\\0&0&8\\0&0&0\end{bmatrix}.$$

求它的奇异值，并给出一组完整的 U,Σ,V。

解　先算 $B^{\mathsf T}B=\operatorname{diag}(0,1,64)$。按特征值从大到小排列，取 $v_1=e_3$、$v_2=e_2$、$v_3=e_1$；于是奇异值是 $8,1,0$。前两个像为 $Be_3=8e_2$、$Be_2=e_1$，补上 $u_3=e_3$，可取

$$U=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},\quad \Sigma=\begin{bmatrix}8&0&0\\0&1&0\\0&0&0\end{bmatrix},\quad V=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}.$$

三者满足 $U^{\mathsf T}U=V^{\mathsf T}V=I$ 且 $B=U\Sigma V^{\mathsf T}$。B 的秩为 2，正奇异值也恰有 2 个。 ♦

命题 28.4（奇异向量与四个基本子空间）　对完整 SVD $A=U\Sigma V^{\mathsf T}$，有

$$C(A^{\mathsf T})=\operatorname{span}\{v_1,\dots,v_r\},\quad N(A)=\operatorname{span}\{v_{r+1},\dots,v_n\},$$

$$C(A)=\operatorname{span}\{u_1,\dots,u_r\},\quad N(A^{\mathsf T})=\operatorname{span}\{u_{r+1},\dots,u_m\}.$$

证明　若 $x=\sum_{j=1}^{n}c_jv_j$，则 $Ax=\sum_{i=1}^{r}\sigma_ic_iu_i$。由于各 $\sigma_i>0$ 且 u_i 标准正交，$Ax=0$ 当且仅当 $c_1=\cdots=c_r=0$，得到零空间公式。像 Ax 的全体线性组合恰好张成 $u_1,\dots,u_r$，得到列空间公式。

又 $A^{\mathsf T}y=\sum_{i=1}^{r}\sigma_i(u_i^{\mathsf T}y)v_i$。因此 $A^{\mathsf T}y=0$ 当且仅当 $u_i^{\mathsf T}y=0$（$1\le i\le r$），即 y 属于其余 $u_{r+1},\dots,u_m$ 张成的空间。并且 $v_i=A^{\mathsf T}u_i/\sigma_i$，所以行空间正是前 r 个 v_i 张成的空间。 ■

读分解的每一列　从 $AV=U\Sigma$ 逐列读取最直接：$Av_i=\sigma_iu_i$。正奇异值给出长度伸缩比例；零奇异值表示输入方向被 A 消去。

28.3　计算：对角化 $A^{\mathsf{T}}A$

存在性证明也给出了一套算法。它先在输入空间找 v_i，再计算并单位化 Av_i；不需要直接猜 U 和 V。

■ **例 28.3　满秩的 2×2 矩阵**　求

$$A=\begin{bmatrix}4 & 4\\ -3 & 3\end{bmatrix}$$

的一组奇异值分解。

解　先求右奇异向量：

$$A^{\mathsf{T}}A=\begin{bmatrix}25 & 7\\ 7 & 25\end{bmatrix}.$$

向量 $(1,1)^{\mathsf{T}}$ 的特征值是 32，$(1,-1)^{\mathsf{T}}$ 的特征值是 18。标准化并按特征值从大到小排列，

$$v_1=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix},\qquad v_2=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}.$$

因此 $\sigma_1=4\sqrt{2}$、$\sigma_2=3\sqrt{2}$。再算像：

$$Av_1=\begin{bmatrix}4\sqrt{2}\\0\end{bmatrix},\qquad Av_2=\begin{bmatrix}0\\-3\sqrt{2}\end{bmatrix}.$$

故 $u_1=e_1$、$u_2=-e_2$。把向量放进列组成

$$V=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix},\qquad U=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix},\qquad \Sigma=\begin{bmatrix}4\sqrt{2} & 0\\ 0 & 3\sqrt{2}\end{bmatrix}.$$

直接逐列核对 $AV=U\Sigma$；又因 $V^{\mathsf{T}}V=I$，右乘 V^{T} 得 $A=U\Sigma V^{\mathsf{T}}$。U,V 都正交，因此这就是完整 SVD。

将乘积按列展开，可得到

$$A=4\sqrt{2}\,u_1v_1^{\mathsf{T}}+3\sqrt{2}\,u_2v_2^{\mathsf{T}}=\begin{bmatrix}4 & 4\\ 0 & 0\end{bmatrix}+\begin{bmatrix}0 & 0\\ -3 & 3\end{bmatrix}.$$

每一项的秩为 1；两项相加恢复原矩阵。这也解释了秩为 2 的原因：这里有两个非零奇异值。 ♦

有时 $A^{\mathsf{T}}A$ 与 A 本身的特征向量没有直接关系。若 A 对称，则可取 V 为它的标准正交特征向量矩阵，奇异值是特征值绝对值；若 A 正定，可进一步取 $U=V$，Σ 就是特征值矩阵。一般矩阵要从 $A^{\mathsf{T}}A$ 开始。

■ **例 28.4　正定矩阵是 SVD 的特殊情形**　令 $S=\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}$。用它的特征向量写出一组 SVD，并核对奇异值。

解　S 的标准正交特征向量可取 $q_1=(1,1)^{\mathsf T}/\sqrt{2}$、$q_2=(1,-1)^{\mathsf T}/\sqrt{2}$，对应特征值 $3,1$。两者都为正，所以

$$U=V=Q=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix},\qquad \Sigma=\operatorname{diag}(3,1),\qquad S=Q\Sigma Q^{\mathsf T}.$$

同时 $S^{\mathsf T}S=Q\operatorname{diag}(9,1)Q^{\mathsf T}$，故奇异值确为 $3,1$。正定性保证了此处左右奇异向量可以取成同一组。 ♦

想一想　若 $A^{\mathsf T}A$ 有重复的正特征值，对应的右奇异向量唯一吗？

答　不唯一。重复特征值的特征空间中有许多标准正交基；在这个空间内换一组标准正交基，仍会得到同一批奇异值。即使奇异值不同，同时把 u_i,v_i 的符号都反转，乘积 $u_iv_i^{\mathsf T}$ 也不变。

从理论上说，计算 $A^{\mathsf T}A$ 的特征值足以得到奇异值；实际的浮点计算还要考虑误差。若 A 满列秩，最大与最小奇异值之比

$$\kappa_2(A)=\frac{\sigma_1}{\sigma_n}$$

衡量线性变换对相对误差的放大程度。此时 $A^{\mathsf T}A$ 的条件数是 $\kappa_2(A)^2$；当最小奇异值很小时，先形成 $A^{\mathsf T}A$ 会使小特征值更难准确求出。数值软件通常用正交变换把 A 化为双对角矩阵，再迭代求其奇异值；这类方法包括 Golub–Kahan双对角化（bidiagonalization）。本讲用 $A^{\mathsf T}A$ 建立存在性与手算步骤，数值计算则会采用更稳定的实现。

28.4　几何意义：正交变换、伸缩、正交变换

矩阵乘积 $A=U\Sigma V^{\mathsf T}$ 从右向左作用。$V^{\mathsf T}$ 先把输入改写到右奇异向量坐标；Σ 沿坐标轴分别伸缩；U 再把结果送到输出方向。正交矩阵保长度，所以真正改变长度的是 Σ。

对例 28.3，$V^{\mathsf T}$ 把圆上的半径方向转到 v_1,v_2 坐标，Σ 分别乘以 $4\sqrt{2}$ 和 $3\sqrt{2}$，再由 U 反射并对齐输出轴。图 28.1 把单位圆及其像画在同一平面上。

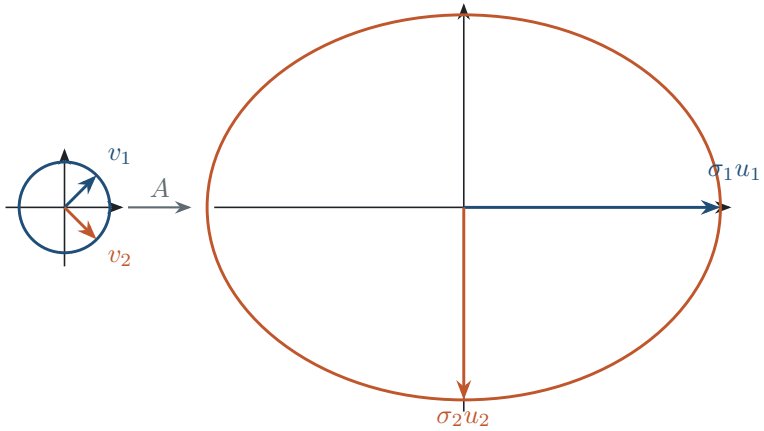


图 28.1　矩阵 A 把单位圆变成椭圆，半轴方向为 u_1,u_2，半轴长分别为奇异值 $4\sqrt{2},3\sqrt{2}$。

从图中看，输入方向 v_1,v_2 仍然互相垂直；它们的像分别沿椭圆的长轴和短轴。若输入 $x=$

av_1+bv_2 且 $\|x\|=1$，则 $a^2+b^2=1$，输出为 $Ax=4\sqrt{2}\,a\,u_1+3\sqrt{2}\,b\,u_2$，故椭圆方程是 $(y_1/(4\sqrt{2}))^2+(y_2/(3\sqrt{2}))^2=1$。

这个描述对任意维数都成立。若 $\|x\|=1$，令 $a_i=v_i^{\mathsf T}x$，则 $\sum_{i=1}^{n}a_i^2=1$，且 $Ax=\sum_{i=1}^{r}\sigma_i a_i u_i$。所以像落在 $C(A)$ 中，并沿 u_i 方向有半轴长 σ_i；若 $r<n$，其余输入系数不会出现在像里，椭球就压到 r 维子空间。再加上 $u_{r+1},\dots,u_m$ 方向便能看清值域中不被 A 触及的部分。

■ **例 28.5　一个输入向量的长度变化**　对例 28.3，令 $x=2v_1+3v_2$。求 $\|x\|$、Ax 和 $\|Ax\|$，并用奇异值解释结果。

解　因为 v_1,v_2 标准正交，$\|x\|^2=2^2+3^2=13$。由 $Av_1=4\sqrt{2}\,u_1$、$Av_2=3\sqrt{2}\,u_2$，

$$Ax=8\sqrt{2}\,u_1+9\sqrt{2}\,u_2=\begin{bmatrix}8\sqrt{2}\\-9\sqrt{2}\end{bmatrix}.$$

输出长度平方为 $(8\sqrt{2})^2+(-9\sqrt{2})^2=290$。直接乘原矩阵也得到 $Ax=(8\sqrt{2},-9\sqrt{2})^{\mathsf T}$，核对一致。不同奇异方向的系数不会在长度平方中产生交叉项。 ♦

对任意 $x=\sum_i(v_i^{\mathsf T}x)v_i$，有 $\|Ax\|^2=\sum_{i=1}^{r}\sigma_i^2(v_i^{\mathsf T}x)^2\le\sigma_1^2\|x\|^2$。所以 σ_1 是最大伸缩比例；若 $\sigma_1>\sigma_2$，等号在 x 沿 v_1 的方向取到。

这个上界也约束方阵的特征值。若 A 有特征值 λ 及非零特征向量 x，将实向量范数延伸到复向量的欧几里得范数，正交矩阵 U,V 仍保持长度，于是

$$|\lambda|\|x\|=\|Ax\|=\|U\Sigma V^{\mathsf T}x\|\le\sigma_1\|x\|,\qquad |\lambda|\le\sigma_1.$$

奇异值给出所有特征值模长的共同上界，但一般不等于特征值；第 1 节的幂零矩阵正好说明二者来源不同。

矩阵中所有元素的平方和也能由奇异值读取。令 $\|A\|_F=(\sum_{i,j}a_{ij}^2)^{1/2}$，称为**Frobenius 范数**（Frobenius norm）。由 $A^{\mathsf T}A=V\Sigma^{\mathsf T}\Sigma V^{\mathsf T}$ 及迹在相似变换下不变，

$$\|A\|_F^2=\operatorname{tr}(A^{\mathsf T}A)=\operatorname{tr}(\Sigma^{\mathsf T}\Sigma)=\sum_{i=1}^{r}\sigma_i^2.$$

对例 28.3，矩阵元素平方和为 $4^2+4^2+(-3)^2+3^2=50$，奇异值平方和也是 $32+18=50$。

完整 SVD 还把 A 写成互相正交的秩一矩阵之和：

$$A=\sum_{i=1}^{r}\sigma_i u_i v_i^{\mathsf T}.$$

这里「正交」指**Frobenius 内积**（Frobenius inner product）$\langle X,Y\rangle_F=\operatorname{tr}(X^{\mathsf T}Y)$：对 $i\ne j$，

$$\langle u_iv_i^{\mathsf T},u_jv_j^{\mathsf T}\rangle_F=(u_i^{\mathsf T}u_j)(v_i^{\mathsf T}v_j)=0.$$

因此各秩一项的平方范数分别为 σ_i^2，总平方误差会按被删去项的 σ_i^2 相加。若只保留最大的 k 项，就得到第 30 讲要证明的最佳秩 k 近似。

若 A 是可逆方阵，SVD 还直接给出

$$A^{-1} = V\Sigma^{-1}U^{\mathsf{T}}, \qquad |\det A| = \prod_{i=1}^{n} \sigma_i.$$

第一式交换变换次序并把各伸缩因子取倒数；第二式来自 $|\det U| = |\det V| = 1$。对例 28.3，$|\det A| = 24$，而奇异值乘积 $(4\sqrt{2})(3\sqrt{2}) = 24$。若方阵奇异，至少一个奇异值为零，行列式也为零。

易错提醒　「正交矩阵把所有向量长度变成 1」是错误的。正交矩阵保持的是长度：$\|Qx\| = \|x\|$。例如 $Q = I$、$x = (2,0)^{\mathsf{T}}$ 时，$\|Qx\| = 2$。在 SVD 中，长度变化来自 Σ 的奇异值。

28.5　零空间、正交列与两个原题

若 $\sigma_i = 0$，则 $Av_i = 0$，所以对应的 v_i 正好给出零空间的标准正交基方向。左零空间 $N(A^{\mathsf{T}})$ 由补齐 U 时加入的方向组成。定理 28.4 给出了这四个子空间的一般公式。

例 28.6　秩亏矩阵的完整 SVD　令

$$C = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}.$$

求其奇异值分解，并指出四个基本子空间在分解中的位置。

解　先算

$$C^{\mathsf{T}}C = \begin{bmatrix} 80 & 60 \\ 60 & 45 \end{bmatrix}.$$

它的特征值是 $125, 0$。取

$$v_1 = \frac{1}{5}\begin{bmatrix} 4 \\ 3 \end{bmatrix}, \qquad v_2 = \frac{1}{5}\begin{bmatrix} 3 \\ -4 \end{bmatrix};$$

前者对应 125，后者对应 0。因此 $\sigma_1 = 5\sqrt{5}$、$\sigma_2 = 0$。计算

$$Cv_1 = \begin{bmatrix} 5 \\ 10 \end{bmatrix} = 5\sqrt{5}\left(\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right).$$

取 $u_1 = (1,2)^{\mathsf{T}}/\sqrt{5}$，再取其正交单位向量 $u_2 = (2,-1)^{\mathsf{T}}/\sqrt{5}$。于是

$$U = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 5\sqrt{5} & 0 \\ 0 & 0 \end{bmatrix}, \quad V = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix},$$

满足 $C = U\Sigma V^{\mathsf{T}}$。

v_1 张成行空间，v_2 张成零空间；u_1 张成列空间，u_2 张成左零空间。每一对确实正交：$(4,3)\cdot(3,-4) = 0$，$(1,2)\cdot(2,-1) = 0$。 ♦

同一组关系还可以直接解相容的方程 $Cx = b$。把 b 在列空间基 $u_1, \dots, u_r$ 中展开为 $b = \sum_{i=1}^{r} \beta_i u_i$，则行空间中的特解为

$$x_{\text{row}} = \sum_{i=1}^{r} \frac{\beta_i}{\sigma_i} v_i, \qquad Cx_{\text{row}} = b.$$

所有解都是 $x_{\text{row}} + z$，其中 $z \in N(C)$。若 b 还有非零的左零空间分量，它不在 $C(C)$ 中，方程便不相容。

例 28.7　用 SVD 写出秩亏系统的全部解　对上例的矩阵 C，求 $Cx = b$ 在 $b = (3,6)^{\mathsf{T}}$ 时的全部解，并说明哪些 $b \in \mathbb{R}^2$ 能使方程相容。

解　上例给出 $u_1 = (1,2)^{\mathsf{T}}/\sqrt{5}$、$v_1 = (4,3)^{\mathsf{T}}/5$、$v_2 = (3,-4)^{\mathsf{T}}/5$ 及 $\sigma_1 = 5\sqrt{5}$。先算 $\beta_1 = u_1^{\mathsf{T}} b = 3\sqrt{5}$，所以

$$x_{\text{row}} = \frac{\beta_1}{\sigma_1} v_1 = \frac{3}{5} v_1 = \begin{bmatrix} 12/25 \\ 9/25 \end{bmatrix}.$$

直接相乘得 $Cx_{\text{row}} = (3,6)^{\mathsf{T}}$，而 $Cv_2 = 0$，故全部解为

$$x = \begin{bmatrix} 12/25 \\ 9/25 \end{bmatrix} + t \begin{bmatrix} 3/5 \\ -4/5 \end{bmatrix}, \qquad t \in \mathbb{R}.$$

对任意 $b \in \mathbb{R}^2$，$C(C) = \operatorname{span}\{u_1\}$，所以相容的充要条件是 $u_2^{\mathsf{T}} b = 0$，其中 $u_2 = (2,-1)^{\mathsf{T}}/\sqrt{5}$ 张成左零空间。非唯一性只来自零空间中的自由分量。 ♦

图 28.2 把这些方向放进定义域和值域两侧；蓝色框表示行、列方向，青色框表示零空间和左零空间方向。

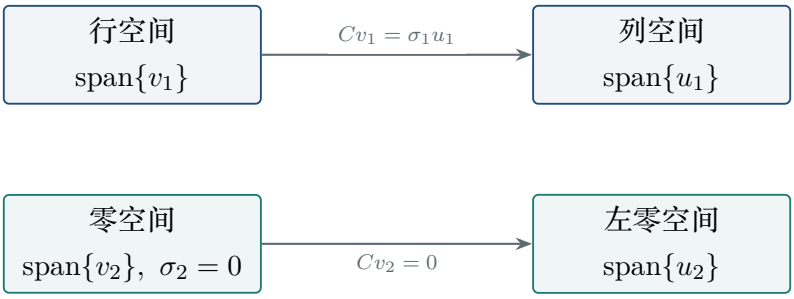


图 28.2　正奇异值把行空间方向送到列空间方向，零奇异值方向落入零空间；左零空间补足值域的正交基。

图 28.2 中的两条映射关系是 $Cv_i = \sigma_i u_i$；完整 U, V 的补齐步骤把四个基本子空间同时放入一个分解里。

例 28.8　宽矩阵的右零方向　令

$$W = \begin{bmatrix} 0 & 4 & 0 \\ 3 & 0 & 0 \end{bmatrix}.$$

求它的正奇异值和一组完整 SVD，并说明零空间与左零空间的维数。

解　$W^{\mathsf{T}}W=\operatorname{diag}(9,16,0)$。按特征值从大到小取 $v_1=e_2$、$v_2=e_1$、$v_3=e_3$；正奇异值是 $4,3$，而 $Wv_3=0$。取 $u_1=e_1$、$u_2=e_2$，便有

$$U=I_2,\qquad \Sigma=\begin{bmatrix}4&0&0\\0&3&0\end{bmatrix},\qquad V=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}.$$

相乘可得 $W=U\Sigma V^{\mathsf{T}}$。这里 $N(W)=\operatorname{span}\{e_3\}$，但 $\operatorname{rank}W=2$ 等于输出维数，所以 $N(W^{\mathsf{T}})=\{0\}$。多出的右零方向没有一个左零向量与它配对。 ♦

■ **例 28.9**　**正交列的矩阵**　设

$$D=\begin{bmatrix}3&0\\4&0\\0&2\end{bmatrix},$$

其列向量互相正交，长度分别为 5 和 2。写出 D 的完整 SVD。

解　正交列立即给出

$$D^{\mathsf{T}}D=\begin{bmatrix}25&0\\0&4\end{bmatrix}.$$

所以 $V=I_2$，Σ 的前两个对角元为 $5,2$。把单位列向量补成 $\mathbb{R}^3$ 的标准正交基，可取

$$U=\frac{1}{5}\begin{bmatrix}3&0&-4\\4&0&3\\0&5&0\end{bmatrix},\qquad \Sigma=\begin{bmatrix}5&0\\0&2\\0&0\end{bmatrix}.$$

因为 $D=U\Sigma I_2^{\mathsf{T}}$，这就是所求分解。一般地，若 $n\le m$ 且 A 的 n 列非零、两两正交，列长为 σ_i，则 $A^{\mathsf{T}}A=\operatorname{diag}(\sigma_1^2,\dots,\sigma_n^2)$，可取 $V=I$，U 的前 n 列为 w_i/σ_i。 ♦

前后联系　第 27 讲的相似变换 $M^{-1}AM$ 用同一组基改变方阵的输入与输出坐标；SVD 为输入、输出分别选择标准正交基，因此也适用于长方形矩阵。ZoomNotes 后续用 SVD 讨论 PCA、最佳低秩近似和图像压缩；本讲先建立分解，图像压缩在第 30 讲展开。

本讲小结

- 任意实矩阵 $A\in\mathbb{R}^{m\times n}$ 都有**完整 SVD**：$A=U\Sigma V^{\mathsf{T}}$，U,V 正交，Σ 为矩形对角矩阵。
- **奇异值**是 $A^{\mathsf{T}}A$ 的特征值平方根：$\sigma_i=\sqrt{\lambda_i(A^{\mathsf{T}}A)}\ge 0$；正值个数等于 $\operatorname{rank}A$。
- 正奇异值对应 $Av_i=\sigma_iu_i$ 与 $A^{\mathsf{T}}u_i=\sigma_iv_i$；$v_i$、$u_i$ 分别是输入和输出的标准正交方向。
- $v_i\in N(A)$ 当且仅当对应特征值为 0；完整的 V,U 分别补入零空间与左零空间方向。
- A 把单位球变成椭球；各半轴方向是 u_i，半轴长是 σ_i，零值会使椭球降维。
- $\|Ax\|^2=\sum_{i=1}^{r}\sigma_i^2(v_i^{\mathsf{T}}x)^2$，且 $\|A\|_F^2=\sum_{i=1}^{r}\sigma_i^2$。

练习

28.1 ★　（原题 29.1）对 Fibonacci 矩阵 $F=\begin{bmatrix}1&1\\1&0\end{bmatrix}$，验证奇异值为 $\frac{1+\sqrt{5}}{2}$ 与 $\frac{\sqrt{5}-1}{2}$，并给出一组 U,V。

28.2 ★　（原题 29.2）若 $A=[w_1\ \cdots\ w_n]$ 的列非零且两两正交，列长为 σ_i，求 $A^{\mathsf T}A$ 并给出一组 SVD。

28.3 ★　求 $A=\begin{bmatrix}0&3\\4&0\end{bmatrix}$ 的一组 SVD。

28.4 ★★　对 $C=\begin{bmatrix}1&2\\2&4\end{bmatrix}$ 求奇异值，并写出行空间、零空间、列空间、左零空间的基。

28.5 ★★　判断「方阵的奇异值就是它的特征值」是否正确；若不正确，给一个 2×2 反例。

28.6 ★★　对例 28.3 证明：对任意 $x\in\mathbb{R}^2$，$3\sqrt{2}\,\|x\|\le\|Ax\|\le 4\sqrt{2}\,\|x\|$，并找出取等方向。

28.7 ★★　证明 $\sum_{i,j}a_{ij}^2=\sum_k\sigma_k^2$，其中右侧遍历 A 的所有正奇异值；指出何时等号左、右都为 0。

28.8 ★★　证明 $A^{\mathsf T}A$ 与 $AA^{\mathsf T}$ 的正特征值（含重数）相同；它们的零特征值个数是否一定相同？

28.9 ★★　对 $E=\begin{bmatrix}3&0\\4&0\\0&0\end{bmatrix}$ 给出奇异值和一组完整 SVD。说明为什么不能对零列使用 w_i/σ_i。

28.10 ★★★　证明 SVD 一般不唯一。先同时反转一对 u_i,v_i 的符号，再说明重复正奇异值时还能怎样改变对应的基。

练习答案与提示

28.1 令 $\phi=(1+\sqrt{5})/2$、$\psi=(\sqrt{5}-1)/2$、$d=\sqrt{\phi^2+1}$。$F^{\mathsf T}F=\begin{bmatrix}2&1\\1&1\end{bmatrix}$，其特征值是 $(3+\sqrt{5})/2=\phi^2$、$(3-\sqrt{5})/2=\psi^2$。可取

$$V=\frac{1}{d}\begin{bmatrix}\phi&1\\1&-\phi\end{bmatrix},\quad U=\frac{1}{d}\begin{bmatrix}\phi&-1\\1&\phi\end{bmatrix},\quad \Sigma=\begin{bmatrix}\phi&0\\0&\psi\end{bmatrix}.$$

两矩阵 U,V 正交，且 $Fv_1=\phi u_1$、$Fv_2=\psi u_2$。因此 $F=U\Sigma V^{\mathsf T}$。这验证了原题给出的奇异值。

28.2 $A^{\mathsf T}A=\operatorname{diag}(\sigma_1^2,\dots,\sigma_n^2)$。若列长未按降序排列，取置换矩阵 P 将列按长度从大到小重排，记重排后的长度为 $\tau_1\ge\cdots\ge\tau_n>0$；若原来已排序则 $P=I_n$。取 U 的前 n 列为重排后的 w_i/τ_i，再补成 $\mathbb{R}^m$ 的标准正交基。令 Σ 的前 n 个对角元为 τ_i，其余为零，并取 $V=P$。因为 $AP=U\Sigma$，所以 $A=U\Sigma P^{\mathsf T}=U\Sigma V^{\mathsf T}$。非零列互相正交保证 $n\le m$。

28.3 $A^{\mathsf T}A=\operatorname{diag}(16,9)$，所以 $V=I_2$、$\Sigma=\operatorname{diag}(4,3)$。因为 $Ae_1=4e_2$、$Ae_2=3e_1$，可取 $U=\begin{bmatrix}0&1\\1&0\end{bmatrix}$。于是 $A=U\Sigma V^{\mathsf T}$。

28.4 $C^{\mathsf T}C=\begin{bmatrix}5&10\\10&20\end{bmatrix}$，特征值为 $25,0$，所以奇异值为 $5,0$。取

$$U=V=\frac{1}{\sqrt5}\begin{bmatrix}1&-2\\2&1\end{bmatrix},\qquad \Sigma=\operatorname{diag}(5,0).$$

行空间与列空间均为 $\operatorname{span}\{(1,2)^{\mathsf T}\}$；零空间与左零空间均为 $\operatorname{span}\{(-2,1)^{\mathsf T}\}$。

28.5 错。取 $N=\begin{bmatrix}0&2\\0&0\end{bmatrix}$。因为 $N^2=0$，其特征值是 $0,0$；而 $N^{\mathsf T}N=\operatorname{diag}(0,4)$，奇异值是 $2,0$。

28.6 写 $x=av_1+bv_2$。则 $\|x\|^2=a^2+b^2$，$\|Ax\|^2=32a^2+18b^2$。因此 $18(a^2+b^2)\le\|Ax\|^2\le32(a^2+b^2)$。上界在 $x\in\operatorname{span}\{v_1\}$ 取等，下界在 $x\in\operatorname{span}\{v_2\}$ 取等；$x=0$ 时两边也相等。

28.7 $\sum_{i,j}a_{ij}^2=\operatorname{tr}(A^{\mathsf T}A)=\operatorname{tr}(V\Sigma^{\mathsf T}\Sigma V^{\mathsf T})=\operatorname{tr}(\Sigma^{\mathsf T}\Sigma)=\sum_k\sigma_k^2$。两边都是非负数，等于 0 当且仅当 $A=0$，也当且仅当所有奇异值为 0。

28.8 若 $A^{\mathsf T}Av=\lambda v$ 且 $\lambda>0$，则 $Av\neq0$，并且 $AA^{\mathsf T}(Av)=\lambda Av$。反向若 $AA^{\mathsf T}u=\lambda u$，则 $A^{\mathsf T}u\neq0$ 且 $A^{\mathsf T}A(A^{\mathsf T}u)=\lambda A^{\mathsf T}u$。对应特征空间之间的映射 $v\mapsto Av/\sqrt\lambda$ 与 $u\mapsto A^{\mathsf T}u/\sqrt\lambda$ 互为逆映射，故维数相同，重数也相同。零特征值个数分别为 $n-r$、$m-r$，不一定相同。

28.9 $E^{\mathsf T}E=\operatorname{diag}(25,0)$，奇异值是 $5,0$。取

$$U=\frac15\begin{bmatrix}3&0&-4\\4&0&3\\0&5&0\end{bmatrix},\quad \Sigma=\begin{bmatrix}5&0\\0&0\\0&0\end{bmatrix},\quad V=I_2.$$

U 的列标准正交，且 $E=U\Sigma V^{\mathsf T}$。零列长度 $\sigma_2=0$，除法 w_2/σ_2 没有定义；应补选正交基向量。

28.10 若把 u_i,v_i 同时换成 $-u_i,-v_i$，则 $\sigma_iu_iv_i^{\mathsf T}$ 不变，所以分解不变。若一组重复奇异值等于 $\sigma>0$，记相应列块为 U_I,V_I，该部分为 $U_I(\sigma I)V_I^{\mathsf T}$。对任意正交矩阵 R，同时换成 U_IR,V_IR 后，这部分变为 $U_IR(\sigma I)R^{\mathsf T}V_I^{\mathsf T}=U_I(\sigma I)V_I^{\mathsf T}$，因为 $RR^{\mathsf T}=I$。所以重根对应的基可以一起旋转，分解仍不变。

第 29 讲

线性变换及其矩阵

Linear Transformations and Their Matrices

对应材料：MIT 18.06SC Session 3.6（Strang 第 30 讲视频 Linear Transformations and Their Matrices）课时摘要；Session 3.6 习题 30.1、30.2 及解答；Strang《Introduction to Linear Algebra》第 4 版 §7.1–7.2（第 5 版 §8.1–8.2）；ZoomNotes Part 8。

本讲目标

(1) 会用定义 $T(cv+dw)=cT(v)+dT(w)$ 判断一个变换是不是线性的，知道 $T(0)=0$ 只是必要条件，能给出平移、取长度、取绝对值等反例。

(2) 理解「线性变换把直线变成直线、平行四边形变成平行四边形」，会通过一个图形（房子）的变化看出一个变换在几何上做了什么。

(3) 知道 $\mathbb{R}^n \to \mathbb{R}^m$ 的线性变换都是乘以一个矩阵，矩阵的第 j 列就是 $T(e_j)$；会写出旋转、反射、投影的矩阵。

(4) 掌握「线性变换由它在一组基上的值完全决定」，会用坐标把任意线性变换写成矩阵：第 j 列是 $T(v_j)$ 在输出基下的坐标。

(5) 理解复合变换对应矩阵乘积 A_2A_1，逆变换对应 A^{-1}，并由此看清矩阵乘法规则从哪里来。

到目前为止，矩阵是一切的出发点：先有 A，再有 Ax。本讲换一个顺序，先不写坐标、不写矩阵，只问：一个把向量变成向量的规则 T，要满足什么条件才配叫「线性」？回答了这个问题，矩阵会自己出现：一旦选定一组基，每个线性变换就对应一个矩阵；换一组基，矩阵跟着变，变换本身不变。这个观点能把前面很多结论串起来：矩阵乘法为什么那样定义，特征向量为什么能把矩阵变成对角的，相似矩阵为什么有相同的特征值。为节省篇幅，正文中的列向量常写成 (x_1,x_2)。

29.1 不用坐标看变换

29.1.1 例子：投影到一条直线

在平面上取一条过原点的直线 L。对每个向量 v，令 $T(v)$ 是 v 在 L 上的正交投影。这里没有选坐标轴，T 只是一个几何规则：从 v 的终点向 L 作垂线，垂足就是 $T(v)$。

这个规则有两个性质。

数乘。把 v 拉长到 $2v$，垂线跟着按比例放大（相似三角形），所以 $T(2v)=2T(v)$；一般地 $T(cv)=cT(v)$，$c<0$ 时向量反向，投影也反向。

加法。$v-T(v)$ 垂直于 L，$w-T(w)$ 也垂直于 L，两者之和

$$(v+w)-\big(T(v)+T(w)\big)$$

仍垂直于 L。而 $T(v)+T(w)$ 在 L 上，从 $v+w$ 到它的差向量又垂直于 L，所以它就是 $v+w$ 的垂足：$T(v+w)=T(v)+T(w)$。图 29.1 画出了这件事：先相加再投影，与先投影再相加，落在同一点。

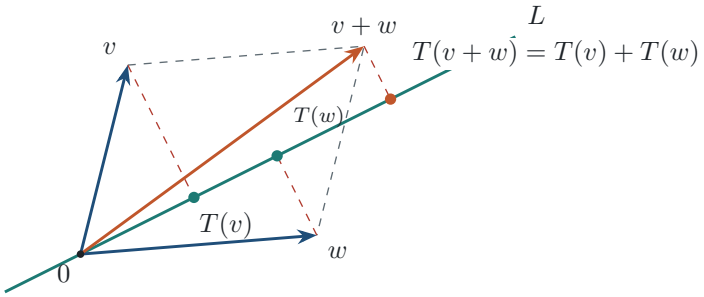


图 29.1　投影到直线 L 保持加法。红色虚线是三条垂线，$T(v)$ 与 $T(w)$ 沿 L 相加，正好到达 $v+w$ 的垂足。

29.1.2 例子：旋转 45°

再看把每个向量绕原点逆时针转 45° 的规则 T。v、w 和 $v+w$ 构成一个平行四边形，整个平行四边形一起旋转后仍是平行四边形，对角线 $v+w$ 转到对角线 $T(v)+T(w)$，所以 $T(v+w)=T(v)+T(w)$。先把 v 拉长 c 倍再转，与先转再拉长，结果一样，所以 $T(cv)=cT(v)$。

这两个例子的共同点是：规则本身与坐标无关，但它和加法、数乘「交换次序」。这正是线性的含义。

29.2 线性变换的定义

29.2.1 定义与直接推论

定义 29.1（线性变换）　设 V,W 是向量空间。映射 $T\colon V\to W$ 称为**线性变换**（linear transformation），如果对所有 $v,w\in V$ 和所有数 c，

$$T(v+w)=T(v)+T(w),\qquad T(cv)=cT(v).$$

两个条件可以合成一个：

$$\boxed{T(cv+dw)=cT(v)+dT(w)\qquad \text{对所有}v,w\in V\ \text{和所有数}c,d.}$$

由定义推出这一条：$T(cv+dw)=T(cv)+T(dw)=cT(v)+dT(w)$。反过来，取 $c=d=1$ 得加法条件；取 $d=0$，由 $0\cdot w=0$ 与 $0\cdot T(w)=0$ 得数乘条件。

由定义立刻得到下面几条。

(1) $T(0)=0$。在 $T(cv)=cT(v)$ 中取 $c=0$：$T(0)=T(0\cdot v)=0\cdot T(v)=0$。

(2) $T(-v)=-T(v)$，$T(v-w)=T(v)-T(w)$。取 $c=-1$ 即得。

(3) 对任意 n 个向量和 n 个数，

$$T(c_1v_1+c_2v_2+\cdots+c_nv_n)=c_1T(v_1)+c_2T(v_2)+\cdots+c_nT(v_n).$$

对 n 归纳：$n=2$ 是上面的合并条件；若对 $n-1$ 成立，把前 $n-1$ 项看成一个向量 u，则 $T(u+c_nv_n)=T(u)+c_nT(v_n)$，再对 $T(u)$ 用归纳假设。

第 (3) 条是线性变换最有用的形式：**线性变换把线性组合变成同样系数的线性组合**。

29.2.2 例子与反例

■ **例 29.1　哪些是线性的**　判断下列变换是否线性：(1) 平移 $T(v)=v+v_0$，$v_0\neq 0$ 是固定向量；(2) 取长度 $T(v)=\|v\|$，$T:\mathbb{R}^2\to\mathbb{R}$；(3) $T(x,y)=(x,|y|)$；(4) $T(x,y)=(x,y^2)$；(5) $T(x,y)=(2x+y,\ x-3y)$。

解　要说明「不是线性的」，举一个使定义失败的具体例子即可；要说明「是线性的」，必须对所有 v,w,c,d 验证。

(1) $T(0)=v_0\neq 0$，违反推论 (1)，不是线性的。也可以直接看数乘：$T(2v)=2v+v_0$，而 $2T(v)=2v+2v_0$，两者差 v_0。

(2) $T(0)=0$ 成立，但取 $c=-1$：$T(-v)=\|-v\|=\|v\|$，而 $-T(v)=-\|v\|$。只要 $v\neq 0$ 两者就不相等，不是线性的。加法也不成立：$v=(1,0)$，$w=(0,1)$ 时 $T(v+w)=\sqrt{2}$，$T(v)+T(w)=2$。

(3) $T(0)=0$，数乘对 $c\geq 0$ 也成立，但加法失败：

$$T\big((1,1)+(1,-1)\big)=T(2,0)=(2,0),\qquad T(1,1)+T(1,-1)=(1,1)+(1,1)=(2,2).$$

(4) $T(0)=0$，但数乘失败：$T\big(2\cdot(0,1)\big)=T(0,2)=(0,4)$，而 $2T(0,1)=(0,2)$。

(5) 是线性的。对任意 $v=(x,y)$，$w=(x',y')$ 和数 c,d，

$$\begin{aligned}T(cv+dw)&=\big(2(cx+dx')+(cy+dy'),\ (cx+dx')-3(cy+dy')\big)\\&=c\,(2x+y,\ x-3y)+d\,(2x'+y',\ x'-3y')=cT(v)+dT(w).\end{aligned}$$

每个输出分量都是输入分量的一次齐次式（没有常数项、没有乘积和乘方），这类变换都是线性的。下一节会看到它就是乘以矩阵 $\left[\begin{smallmatrix}2&1\\1&-3\end{smallmatrix}\right]$。 ♦

易错提醒　$T(0)=0$ **是线性的必要条件，不是充分条件**。例 29.1 的 (2)(3)(4) 都满足 $T(0)=0$，却都不是线性的。检验线性时，先看 $T(0)$ 能快速排除平移一类的变换，但最后必须验证加法和数乘两条。

另一个常见的混淆来自中学的「线性函数」$f(x)=kx+b$。按本讲的定义，它只在 $b=0$ 时才是线性变换；$b\neq 0$ 时它是线性变换再加一个平移，叫**仿射变换**（affine transformation）。

想一想　只要求数乘条件 $T(cv)=cT(v)$ 对所有实数 c 成立，能否推出 T 是线性的？

答　不能。令 $T:\mathbb{R}^2\to\mathbb{R}$，$T(x,y)=(x^3+y^3)^{1/3}$（取实立方根）。对任意实数 c，$T(cx,cy)=\big(c^3(x^3+y^3)\big)^{1/3}=cT(x,y)$，数乘条件成立。但 $T(1,0)+T(0,1)=1+1=2$，而 $T(1,1)=2^{1/3}\approx 1.26$，加法失败。两个条件缺一不可。

29.2.3　直线变直线：看房子怎样变

线性有一个直接的几何后果。连接 v 与 w 的线段上的点是 $(1-t)v+tw$，$0\le t\le 1$。由推论 (3)，

$$T\big((1-t)v+tw\big)=(1-t)\,T(v)+t\,T(w),$$

右边正是连接 $T(v)$ 与 $T(w)$ 的线段上参数为 t 的点。所以：

- 线段变成线段（$T(v)=T(w)$ 时退化成一个点），参数 t 不变，因此中点变成中点，等分点变成等分点；
- 过 v_0、方向为 u 的直线 v_0+tu 变成 $T(v_0)+tT(u)$，仍是直线，$T(u)=0$ 时退化成一点；
- 平行四边形 $0,v,w,v+w$ 变成 $0,T(v),T(w),T(v)+T(w)$，仍是平行四边形，可能被压扁成线段；方格纸变成由平行四边形组成的网格。

于是要知道 T 对一个由线段组成的图形做了什么，只需知道各顶点去了哪里。Strang 在课上画一座「房子」：看房子变成什么样，就看懂了这个变换。图 29.2 是同一座房子在四个线性变换下的像，门和窗故意画得不对称，以便看出图形是否被翻转。

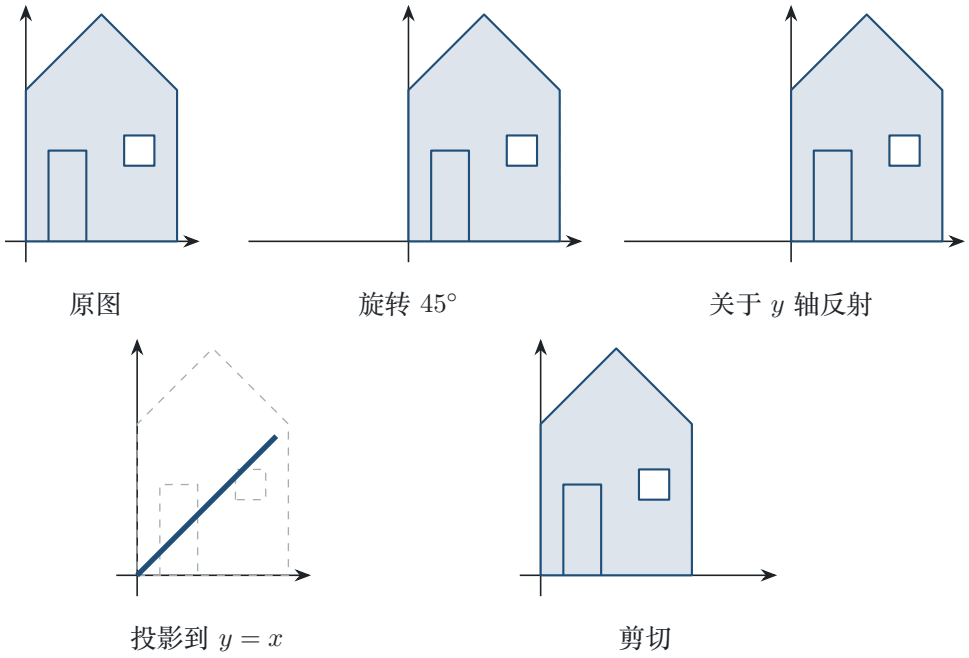


图 29.2　虚线为原房子，实线为旋转、反射、投影和剪切后的像。投影把房子压到从 0 到 $(2,2)$ 的线段上；关于 y 轴的反射交换门窗位置。

图中的剪切 $T(x,y)=(x+\tfrac{1}{2}y,\ y)$ 让每个点水平右移自身高度的一半：地面不动，屋顶向右歪，竖直的墙变斜，但仍是直线，平行的墙仍平行。只靠「直线变直线、平行保持平行」这两条，

就能从几个顶点画出整个像。

29.3　矩阵给出的线性变换

29.3.1　每个矩阵都是一个线性变换

给定 $m\times n$ 矩阵 A，令 $T(v)=Av$，$T\colon \mathbb{R}^n\to\mathbb{R}^m$。矩阵乘法满足分配律，数可以从乘积里提出来，所以

$$A(cv+dw)=A(cv)+A(dw)=c\,Av+d\,Aw,$$

T 是线性的。几个例子：

- $A=\left[\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right]$：$(x,y)\mapsto(x,-y)$，关于 x 轴反射。房子被上下翻转，倒挂在 x 轴下方。
- $A=\left[\begin{smallmatrix}2&0\\0&2\end{smallmatrix}\right]=2I$：每个向量拉长为 2 倍。习题 30.1 用极坐标描述同一个变换：$T(r,\theta)=(2r,\theta)$，长度加倍、方向不变，所以 $T(v)=2v$，是线性的。用直角坐标验证：$T(x,y)=(2x,2y)$，$\|T(x,y)\|=\sqrt{4x^2+4y^2}=2\|(x,y)\|$。
- 任意 2×3 矩阵给出 $\mathbb{R}^3\to\mathbb{R}^2$ 的线性变换。例如 $A=\left[\begin{smallmatrix}1&0&0\\0&1&0\end{smallmatrix}\right]$ 把 (x,y,z) 变成 (x,y)，丢掉高度，相当于把三维物体在 xy 平面上的影子画出来。

29.3.2　$\mathbb{R}^n\to\mathbb{R}^m$ 的线性变换都是矩阵

反过来也成立：从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性变换，没有一个能逃出「乘以矩阵」的范围。

> **定理 29.2（标准矩阵）**　设 $T\colon\mathbb{R}^n\to\mathbb{R}^m$ 是线性变换，$e_1,\dots,e_n$ 是 $\mathbb{R}^n$ 的标准基。令 A 是以 $T(e_1),\dots,T(e_n)$ 为列的 $m\times n$ 矩阵，则对所有 $v\in\mathbb{R}^n$ 有 $T(v)=Av$，并且这样的矩阵唯一。

证明　任意 $v=(x_1,\dots,x_n)$ 都可以写成 $v=x_1e_1+\cdots+x_ne_n$。由线性，

$$T(v)=x_1T(e_1)+\cdots+x_nT(e_n).$$

右边是 A 的各列以 $x_1,\dots,x_n$ 为系数的组合，按矩阵乘向量的列观点，它就是 Av。若另有矩阵 B 满足 $T(v)=Bv$，取 $v=e_j$ 得 B 的第 j 列 $Be_j=T(e_j)$，与 A 的第 j 列相同，所以 $B=A$。 ■

证明的关键一步是「v 是 e_j 的组合」加上「线性保持组合」。所以求矩阵只需回答 n 个问题：$e_1,\dots,e_n$ **分别去了哪里**。A 叫做 T 的**标准矩阵**（standard matrix）。

■ **例 29.2　旋转、投影与反射的矩阵**　求下列平面变换的标准矩阵：(1) 逆时针旋转角 θ；(2) 投影到直线 $y=x$；(3) 关于直线 $y=x$ 反射。

解　按定理 29.2，只需找 $T(e_1)$ 和 $T(e_2)$。

(1) $e_1=(1,0)$ 的方向角是 0，转后方向角为 θ，所以 $T(e_1)=(\cos\theta,\sin\theta)$。$e_2=(0,1)$ 的方向角是 $\pi/2$，转后为 $\pi/2+\theta$，所以 $T(e_2)=\big(\cos(\pi/2+\theta),\sin(\pi/2+\theta)\big)=(-\sin\theta,\cos\theta)$。于是

$$Q_\theta=\begin{bmatrix}\cos\theta&-\sin\theta\\ \sin\theta&\cos\theta\end{bmatrix},\qquad Q_{45^\circ}=\frac{1}{\sqrt{2}}\begin{bmatrix}1&-1\\1&1\end{bmatrix}.$$

验证图 29.2：房子的右下角 $(2,0)$ 转到 $Q_{45^\circ}(2,0)=(\sqrt{2},\sqrt{2})$，屋顶 $(1,3)$ 转到 $\frac{1}{\sqrt{2}}(1-3,\,1+3)=(-\sqrt{2},\,2\sqrt{2})$。

(2) 直线方向 $a=(1,1)$，$a^{\mathsf{T}}a=2$。向量 b 在 a 上的投影是 $\dfrac{a^{\mathsf{T}}b}{a^{\mathsf{T}}a}a$（第 14 讲）。$a^{\mathsf{T}}e_1=1$，$a^{\mathsf{T}}e_2=1$，所以 $T(e_1)=T(e_2)=\frac{1}{2}(1,1)$，

$$P=\begin{bmatrix}1/2 & 1/2\\ 1/2 & 1/2\end{bmatrix}=\frac{aa^{\mathsf{T}}}{a^{\mathsf{T}}a}.$$

检查：$P^2=P$（投影两次与投影一次一样），$Pa=a$（直线上的点不动），$P(1,-1)=0$（垂直方向被压成 0）。

(3) 反射交换两个坐标：$T(e_1)=e_2$，$T(e_2)=e_1$，矩阵为 $H=\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$。它与投影的关系是：从 v 走到垂足 Pv，再走同样远，到达镜像点 $v+2(Pv-v)=2Pv-v$。所以 $H=2P-I$，代入 P 正好得到 $\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$。 ♦

想一想　平移 $T(v)=v+v_0$ 不是线性的，所以不存在 2×2 矩阵 A 使 $Av=v+v_0$。可是计算机图形学里平移、旋转都用矩阵乘法做。这是怎么回事？

答　图形学用齐次坐标（homogeneous coordinates）：把平面上的点 (x,y) 写成 $\mathbb{R}^3$ 中的 $(x,y,1)$，则

$$\begin{bmatrix}1&0&a\\0&1&b\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}=\begin{bmatrix}x+a\\y+b\\1\end{bmatrix}.$$

这个 3×3 矩阵是 $\mathbb{R}^3$ 上的线性变换，它把平面 $z=1$ 平移了 (a,b)。平移在 $\mathbb{R}^2$ 里不线性，升一维后变成了线性变换的一部分；这与本讲的结论不矛盾。

29.4　基决定一切

29.4.1　在一组基上的值决定整个变换

定理 29.2 的证明只用到一件事：每个向量都能唯一地写成 $e_1,\dots,e_n$ 的组合。标准基并不特殊，任何一组基都有这个性质（第 9 讲）。这引出本讲的核心事实。

定理 29.3（线性变换由基上的值决定）　设 $v_1,\dots,v_n$ 是 V 的一组基。

(1)（唯一性）若线性变换 $T,S\colon V\to W$ 满足 $T(v_i)=S(v_i)$，$i=1,\dots,n$，则 $T=S$。

(2)（存在性）对 W 中任意给定的 n 个向量 $y_1,\dots,y_n$，存在唯一的线性变换 $T\colon V\to W$ 使 $T(v_i)=y_i$。

思路　任意 v 写成 $v=c_1v_1+\cdots+c_nv_n$，线性迫使 $T(v)=c_1T(v_1)+\cdots+c_nT(v_n)$。唯一性由此直接得到；存在性则反过来用这个公式定义 T，再验证它线性。

证明　(1) 任取 $v\in V$，写成 $v=c_1v_1+\cdots+c_nv_n$（基张成 V，这样的系数存在）。由线性，

$$T(v)=\sum_{i=1}^{n}c_iT(v_i)=\sum_{i=1}^{n}c_iS(v_i)=S(v).$$

(2) 基向量线性无关，所以 v 的系数 $c_1,\dots,c_n$ 唯一，定义

$$T(c_1v_1+\cdots+c_nv_n)=c_1y_1+\cdots+c_ny_n$$

不会产生歧义。把 v_i 写成基展开时取第 i 个系数为 1、其余为 0，定义式给出 $T(v_i)=y_i$。验证线性：若 $v=\sum c_iv_i$，$w=\sum d_iv_i$，则 $av+bw=\sum(ac_i+bd_i)v_i$，按定义

$$T(av+bw)=\sum_i(ac_i+bd_i)y_i=a\sum_i c_iy_i+b\sum_i d_iy_i=aT(v)+bT(w).$$

唯一性由 (1) 得到。 ■

两个条件各有用处。若 $v_1,\dots,v_n$ 线性相关，值就不能随意指定：比如 $v_2=2v_1$，线性迫使 $T(v_2)=2T(v_1)$，要求 $T(v_1)=e_1$、$T(v_2)=e_2$ 就自相矛盾。若它们不张成 V，指定这些值又不足以确定 T：在张成的子空间之外，T 可以任意取值。「基」恰好介于两者之间：值可以任意指定，而且指定后 T 完全确定。

29.4.2 坐标

上面的系数 $c_1,\dots,c_n$ 叫做 v 关于基 $v_1,\dots,v_n$ 的坐标（coordinates）。坐标总是相对于一组有顺序的基说的。在标准基下，$v=(x_1,\dots,x_n)$ 的坐标就是它的分量 $x_1,\dots,x_n$，所以平时不必区分「向量」和「坐标」；换一组基，同一个向量的坐标就变了。Strang 在课上说：坐标来自基，*coordinates come from a basis*。

图 29.3 中，$v_1=(1,1)$ 与 $v_2=(1,-1)$ 构成 $\mathbb{R}^2$ 的一组基，它们张出一张斜的网格。向量 $x=(4,2)$ 在标准基下的坐标是 $(4,2)$：向右走 4，向上走 2。在新基下，解 $c_1(1,1)+c_2(1,-1)=(4,2)$，即 $c_1+c_2=4$，$c_1-c_2=2$，得 $c_1=3$，$c_2=1$：沿 v_1 走 3 格，再沿 v_2 走 1 格。同一个箭头，两套坐标。

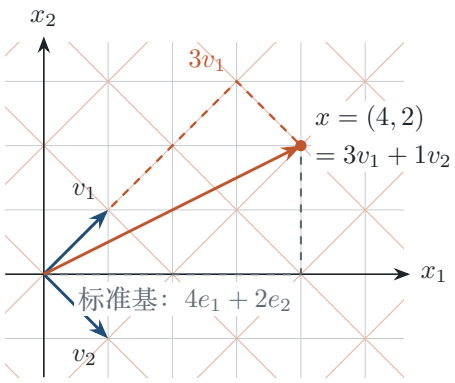


图 29.3　灰色方格是标准基的网格，橙色斜格是基 $v_1=(1,1)$、$v_2=(1,-1)$ 的网格。同一个向量 x 在标准基下的坐标是 $(4,2)$（灰色折线），在新基下的坐标是 $(3,1)$（橙色折线）。

■ **例 29.3　只知道基上的值**　线性变换 $T\colon\mathbb{R}^2\to\mathbb{R}^2$ 满足 $T(1,1)=(3,1)$，$T(1,-1)=(1,-1)$。求 $T(5,1)$ 和 T 的标准矩阵。

解　记 $v_1=(1,1)$，$v_2=(1,-1)$。先求 $(5,1)$ 的坐标：$c_1+c_2=5$，$c_1-c_2=1$，得 $c_1=3$，$c_2=2$。

检查：$3(1,1)+2(1,-1)=(5,1)$。由线性，

$$T(5,1)=3\,T(v_1)+2\,T(v_2)=3(3,1)+2(1,-1)=(11,1).$$

标准矩阵需要 $T(e_1)$ 和 $T(e_2)$。由 $e_1=\frac{1}{2}(v_1+v_2)$，$e_2=\frac{1}{2}(v_1-v_2)$，

$$T(e_1)=\tfrac{1}{2}\big((3,1)+(1,-1)\big)=(2,0),\qquad T(e_2)=\tfrac{1}{2}\big((3,1)-(1,-1)\big)=(1,1),$$

所以 $A=\left[\begin{smallmatrix}2&1\\0&1\end{smallmatrix}\right]$。另一种写法是把两个条件并成一个矩阵方程 $A\left[\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right]=\left[\begin{smallmatrix}3&1\\1&-1\end{smallmatrix}\right]$，右乘 $\left[\begin{smallmatrix}1&1\\1&-1\end{smallmatrix}\right]^{-1}=\left[\begin{smallmatrix}1/2&1/2\\1/2&-1/2\end{smallmatrix}\right]$ 得到同一个 A。验算：$A(1,1)=(3,1)$，$A(1,-1)=(1,-1)$，$A(5,1)=(11,1)$。 ♦

29.5　线性变换的矩阵

29.5.1　构造：第 j 列是 $T(v_j)$ 的坐标

现在 V 和 W 可以是任何有限维向量空间：多项式、矩阵、函数，不必是 $\mathbb{R}^n$。设 $T\colon V\to W$ 线性，V 有基 $v_1,\dots,v_n$（**输入基**（input basis）），W 有基 $w_1,\dots,w_m$（**输出基**（output basis））。每个 $T(v_j)$ 都在 W 中，可以唯一地写成 $w_1,\dots,w_m$ 的组合：

$$\boxed{T(v_j)=a_{1j}w_1+a_{2j}w_2+\cdots+a_{mj}w_m,\qquad j=1,\dots,n.}$$

把系数 $a_{1j},\dots,a_{mj}$ 排成第 j 列，得到 $m\times n$ 矩阵 $A=(a_{ij})$，叫做 T **关于这两组基的矩阵**。n 是输入空间的维数，所以 A 有 n 列；m 是输出空间的维数，所以 A 有 m 行。

> **定理 29.4（矩阵把输入坐标变成输出坐标）**　若 v 关于输入基的坐标是 $c=(c_1,\dots,c_n)$，则 $T(v)$ 关于输出基的坐标是 Ac。

证明　由 $v=\sum_j c_jv_j$ 和线性，

$$T(v)=\sum_{j=1}^{n}c_jT(v_j)=\sum_{j=1}^{n}c_j\sum_{i=1}^{m}a_{ij}w_i=\sum_{i=1}^{m}\Big(\sum_{j=1}^{n}a_{ij}c_j\Big)w_i.$$

w_i 的系数 $\sum_j a_{ij}c_j$ 正是 Ac 的第 i 个分量。坐标是唯一的，所以 $T(v)$ 的坐标就是 Ac。 ■

证明中交换了两个求和号的次序，这一步把「T 作用在组合上」变成了「矩阵的第 i 行乘 c」。图 29.4 用一个方块图概括这件事：上面一行是向量空间里的抽象变换，下面一行是坐标之间的矩阵乘法；先取坐标再乘 A，与先作用 T 再取坐标，结果相同。

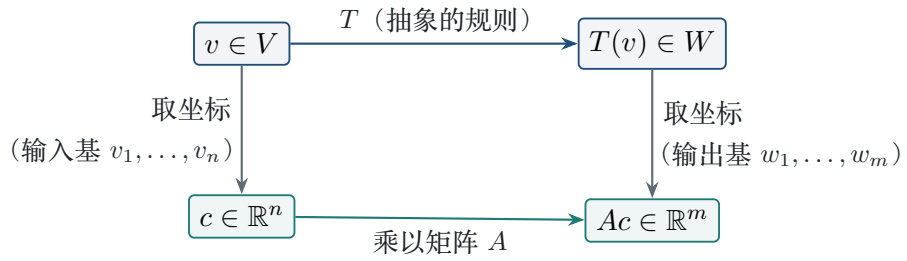


图 29.4　线性变换与它的矩阵。沿两条路径从左上走到右下，得到同一个结果：$T(v)$ 的输出坐标等于 A 乘 v 的输入坐标。

易错提醒

- $T(v_j)$ 的坐标写成 A 的第 j 列，不是第 j 行。写成行会得到 A^{T}。检验办法：A 乘 $c=e_j$（即 $v=v_j$）应当取出第 j 列。
- 坐标要用输出基来表示。$T(v_j)$ 在 W 里，只能用 $w_1,\dots,w_m$ 展开。
- 矩阵依赖两组基和它们的顺序。交换 v_1,v_2 会交换 A 的两列，交换 w_1,w_2 会交换 A 的两行。

以二次多项式 $p(x)=a+bx+cx^2$ 为输入，求导后得到 $p'(x)=b+2cx$。输入有三个系数，输出有两个系数；只要分别选好输入、输出基，求导也能写成矩阵乘法。

■ **例 29.4　导数的矩阵**　令 $T\colon P_2\to P_1$ 为求导变换，其中 P_2 是次数不超过 2 的实系数多项式空间，P_1 是次数不超过 1 的多项式空间。输入基取 $(1,x,x^2)$，输出基取 $(1,x)$，求 T 的矩阵，并计算 $T(2-3x+4x^2)$。

解　逐个对输入基向量求导，并用输出基写坐标：$T(1)=0=0\cdot 1+0\cdot x$，$T(x)=1=1\cdot 1+0\cdot x$，$T(x^2)=2x=0\cdot 1+2\cdot x$。所以三列依次是 $(0,0)$、$(1,0)$、$(0,2)$，

$$D=\begin{bmatrix}0&1&0\\0&0&2\end{bmatrix}.$$

多项式 $p=2-3x+4x^2$ 的输入坐标是 $(2,-3,4)$，故 $[T(p)]_{(1,x)}=D(2,-3,4)=(-3,8)$，即 $T(p)=-3+8x$。输入和输出基的维数不同，所以 D 是 2×3 矩阵。 ♦

这个矩阵也能直接读出核和像：求导会消去常数项，输出却能取到任意一次以下多项式。稍后我们把这两个集合写成一般定义。

■ **例 29.5　定积分的矩阵与左右复合**　令 $S\colon P_1\to P_2$ 为 $S(q)(x)=\int_0^x q(t)\,\mathrm{d}t$。仍用 P_1 的基 $(1,x)$ 和 P_2 的基 $(1,x,x^2)$，求 S 的矩阵，并计算求导与积分的两种复合。

解　把输出基向量逐个积分：$S(1)=x$，$S(x)=\frac{1}{2}x^2$。因此

$$E=\begin{bmatrix}0&0\\1&0\\0&1/2\end{bmatrix}.$$

先积分再求导对应 DE，先求导再积分对应 ED：

$$DE=\begin{bmatrix}1&0\\0&1\end{bmatrix},\qquad ED=\begin{bmatrix}0&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

第一式表示 $D(S(q))=q$ 对每个 $q\in P_1$ 成立。第二式把 $a+bx+cx^2$ 变为 $bx+cx^2$，常数项被去掉。积分下限固定为 0，所以积分算子不会补回原来的常数项。 ♦

29.5.2 复合变换与矩阵乘法

若 $T_1: V \to W$ 的坐标矩阵为 A，$T_2: W \to Z$ 的坐标矩阵为 B，且前一变换的输出基正是后一变换的输入基，那么先做 T_1 再做 T_2，坐标就先乘 A 再乘 B。矩阵乘法的次序因此是 BA。

定理 29.5（复合与逆变换的矩阵）　设 $T_1: V \to W$、$T_2: W \to Z$ 是线性变换，所选基在 W 中相同。若它们的矩阵分别为 A、B，则 $T_2 \circ T_1$ 的矩阵为 BA。若 $T: V \to W$ 是双射，矩阵为 A，则 T^{-1} 在相反两组基下的矩阵为 A^{-1}。

证明　对输入坐标列 c，T_1 的输出坐标是 Ac，再经 T_2 后变成 $B(Ac) = (BA)c$，故复合矩阵是 BA。若 T 可逆，$T^{-1} \circ T$ 与 $T \circ T^{-1}$ 都是恒等变换，矩阵分别为 $A^{-1}A$ 与 AA^{-1}；两者都等于相应空间的单位矩阵，所以逆变换的矩阵是 A^{-1}。 ■

易错提醒　先做 T_1、后做 T_2 时，矩阵次序写成 BA，不是 AB。向量坐标是列，复合时右边的矩阵先作用；若维数允许但基不匹配，直接相乘仍可能算出矩阵，却未必表示这两个变换的复合。

■ **例 29.6　剪切与拉伸的复合**　令 $T(x,y) = (x+y, y)$，$S(x,y) = (x, 2y)$。求先做 T 再做 S 的矩阵，并求 T^{-1}。

解　两变换的标准矩阵为 $A = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$、$B = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right]$。所以

$$BA = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \qquad (S \circ T)(x,y) = (x+y, 2y).$$

由 $T(x-y, y) = (x, y)$，可得 $T^{-1}(x,y) = (x-y, y)$，其矩阵为 $A^{-1} = \left[\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right]$。例如 $(2,-1)$ 经复合后为 $(1,-2)$；用矩阵计算也是 $BA(2,-1) = (1,-2)$。 ♦

29.6 同一变换，换基后矩阵怎样变

一个线性变换 T 不依赖坐标；它的矩阵要先说明输入、输出分别使用哪组基。换基只改变坐标描述，不改变 $T(v)$ 这个向量。

设 $T: \mathbb{R}^n \to \mathbb{R}^m$ 的标准矩阵是 A。把输入基向量按列排成矩阵 $P = [v_1 \ \cdots \ v_n]$，把输出基向量按列排成 $Q = [w_1 \ \cdots \ w_m]$。若 $c = [v]_{(v_1,\dots,v_n)}$，则 $v = Pc$；而 $T(v)$ 的标准坐标为 APc。再用 Q^{-1} 转为输出基坐标，得到

$$[T(v)]_{(w_1,\dots,w_m)} = Q^{-1}APc.$$

定理 29.6（换基后的矩阵）　沿用上述记号，T 关于输入基 $(v_1,\dots,v_n)$ 和输出基 $(w_1,\dots,w_m)$ 的矩阵为 $Q^{-1}AP$。特别地，当 $V = W$ 且输入、输出都用同一组基矩阵 P 时，新矩阵为 $P^{-1}AP$。

证明　输入坐标 c 先由 P 变成标准坐标 Pc。作用 T 后得到标准输出坐标 APc。输出基矩阵 Q

把输出基坐标 d 变成标准坐标 Qd，所以 $Qd = APc$，即 $d = Q^{-1}APc$。这对每个 c 都成立，故矩阵为 $Q^{-1}AP$。 ■

易错提醒 同一变换换基后，矩阵一般会变；这不表示变换本身变了。只有输入、输出都换成同一组基时，公式才是 $P^{-1}AP$；若两边基不同，应使用 $Q^{-1}AP$，不能只在一侧乘逆矩阵。

■ **例 29.7　同一剪切的两种矩阵**　仍令 $T(x,y) = (2x+y, y)$，标准基矩阵为 $A = \left[\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right]$。改用基 $v_1 = (1,1)$、$v_2 = (1,-1)$ 表示输入和输出，求新矩阵，并用它计算 $T(2,4)$。

解　基矩阵及其逆为

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad P^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

因此新矩阵是

$$M = P^{-1}AP = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}.$$

向量 $(2,4)$ 的新坐标为 $(3,-1)$，因为 $3v_1 - v_2 = (2,4)$。用 M 计算，新输出坐标为 $M(3,-1) = (6,2)$；换回原坐标，$6v_1 + 2v_2 = (8,4)$。直接代入标准公式也得 $T(2,4) = (8,4)$。两个矩阵不同，表示的仍是同一个 T。 ♦

若一组基由 T 的特征向量组成，矩阵的每一列只在对应位置有一个特征值，因而成为对角矩阵。矩阵 $P^{-1}AP$ 正是把基向量作为列组成 P 后得到的表示；第 20 讲已经用这种选择来简化矩阵的幂。

29.7　核与像：变换压缩了什么

正交投影把垂直于投影线的方向压到零，同时保留投影线上的方向。要描述这种「哪些输入被压成零、哪些输出能够出现」，需要两个由线性变换自然产生的子空间。

定义 29.7（核与像）　设 $T\colon V \to W$ 是线性变换。T 的**核**（kernel）与**像**（image）分别定义为

$$\ker T = \{v \in V : T(v) = 0\}, \qquad \operatorname{im} T = \{T(v) : v \in V\} \subseteq W.$$

例如，恒等变换的核只有零向量、像是整个空间；零变换的核是整个定义域、像只有零向量。若 $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ 是投影到 $y = x$，则 $T(1,1) = (1,1)$、$T(1,-1) = (0,0)$：投影线上的向量保留，垂直方向被压到零。

定理 29.8（线性变换的核和像是子空间）　对任意线性变换 $T\colon V \to W$，$\ker T$ 是 V 的子空间，$\operatorname{im} T$ 是 W 的子空间。

证明　先看核。$T(0)=0$，所以 $0\in\ker T$。若 $u,v\in\ker T$，且 a,b 是标量，则

$$T(au+bv)=aT(u)+bT(v)=a\cdot 0+b\cdot 0=0,$$

所以 $au+bv\in\ker T$。

再看像。$T(0)=0$，所以 $0\in\operatorname{im}T$。若 $y_1=T(u)$、$y_2=T(v)$ 是像中的向量，则

$$ay_1+by_2=aT(u)+bT(v)=T(au+bv)\in\operatorname{im}T.$$

两者都含零向量且对线性组合封闭，故都是子空间。 ■

若 A 是 T 关于输入基、输出基的矩阵，那么输入坐标列 c 被送到 Ac。因此，核中的向量坐标恰好满足 $Ac=0$；像中的向量坐标恰好是 A 的列向量的线性组合。换句话说，$\ker T$ 对应矩阵零空间 $N(A)$，$\operatorname{im}T$ 对应列空间 $\operatorname{Col}(A)$。

■ **例 29.8**　**投影的核和像**　投影到 $y=x$ 的标准矩阵是 $P=\frac{1}{2}\left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]$。求 $\ker T$、$\operatorname{im}T$，并解释向量 $(2,0)$ 的像。

解　对 (x,y)，$T(x,y)=\left(\frac{x+y}{2},\frac{x+y}{2}\right)$。$T(x,y)=0$ 当且仅当 $x+y=0$，所以

$$\ker T=\operatorname{span}\{(1,-1)\}.$$

每个输出都是 (s,s) 的形式；任意 (s,s) 也可由输入 $(2s,0)$ 得到，所以

$$\operatorname{im}T=\operatorname{span}\{(1,1)\}.$$

特别地，$T(2,0)=(1,1)$。核给出被压掉的方向，像给出所有可能的输出方向。 ♦

图 29.5 用两条过原点的直线区分核与像：橙色的核方向被投影到零，蓝色的像方向则包含全部投影结果。

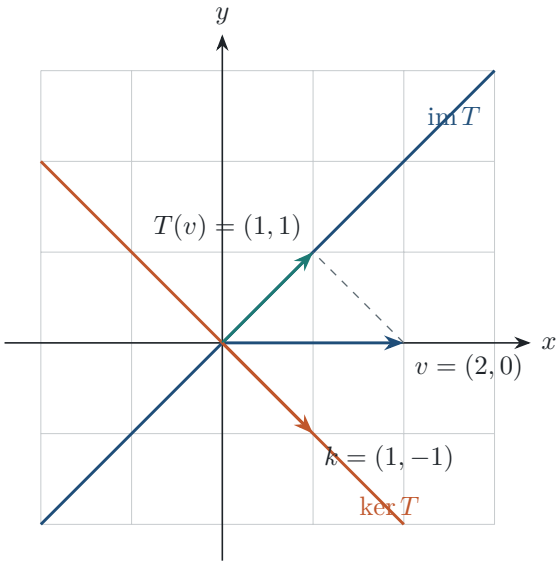


图 29.5　投影到 $y=x$ 的像是蓝色直线，核是橙色直线；$(2,0)$ 投影到 $(1,1)$，而 $(1,-1)$ 被投影到零。

同一结论可从多项式求导矩阵读出。若 $p=a+bx+cx^2$，则 $D(a,b,c)=(b,2c)$；$D(a,b,c)=0$ 恰好要求 $b=c=0$，所以 $\ker D=\operatorname{span}\{1\}$。给定任意 $d+ex\in P_1$，选 $p=dx+\frac{e}{2}x^2$ 就有 $p'=d+ex$，所以 $\operatorname{im} D=P_1$。求导丢掉常数项，但没有丢掉任何一次以下多项式作为输出的可能性。

命题 29.9（单射、满射与核、像）　线性变换 $T\colon V\to W$ 是单射，当且仅当 $\ker T=\{0\}$；T 是满射，当且仅当 $\operatorname{im} T=W$。

证明　若 T 是单射且 $T(v)=0=T(0)$，则 $v=0$。反过来，若核只有 0，而 $T(u)=T(v)$，则 $T(u-v)=0$，所以 $u-v=0$，即 $u=v$。满射的定义正是每个 $w\in W$ 都能写成 $T(v)$，这等价于 $\operatorname{im} T=W$。 ■

想一想　核只有零向量能保证变换满射吗？像等于整个陪域能保证变换单射吗？

答　都不能。$T\colon\mathbb{R}^2\to\mathbb{R}^3$，$T(x,y)=(x,y,0)$ 的核只有零向量，但像是 xy 平面，不是整个 $\mathbb{R}^3$。$S\colon\mathbb{R}^3\to\mathbb{R}^2$，$S(x,y,z)=(x,y)$ 的像是整个 $\mathbb{R}^2$，但核含有非零向量 $(0,0,1)$。双射须同时单射、满射。

本讲小结

- 线性变换保持线性组合：$T(cv+dw)=cT(v)+dT(w)$；$T(0)=0$ 必要但不充分。
- T 在一组基上的值唯一决定整个变换；矩阵第 j 列是 $T(v_j)$ 在输出基下的坐标。
- 输入坐标 c 经矩阵 A 变为输出坐标 Ac；标准矩阵的列为 $T(e_j)$。
- 换输入、输出基矩阵 P,Q 后，矩阵变为 $Q^{-1}AP$；同一基时是 $P^{-1}AP$。
- 先做矩阵 A、再做矩阵 B，复合矩阵为 BA；双射的逆变换对应 A^{-1}。
- $\ker T=\{v:T(v)=0\}$、$\operatorname{im} T=\{T(v):v\in V\}$ 都是子空间；矩阵坐标下分别对应 $N(A)$ 和 $\operatorname{Col}(A)$。

练习

29.1 ★　判断真假，假的给出反例：(a) $T(0)=0$ 就能推出 T 线性；(b) 线性变换的核是子空间；(c) 若 $\ker T=\{0\}$，则 T 是单射；(d) $T\colon\mathbb{R}^3\to\mathbb{R}^2$ 的矩阵有 3 行、2 列。

29.2 ★　设 $T(x,y,z)=(2x-y+z,\,x+2z)$。写出标准矩阵并计算 $T(1,-1,2)$。

29.3 ★★　线性变换 $T\colon\mathbb{R}^2\to\mathbb{R}^2$ 满足 $T(1,1)=(2,1)$、$T(1,-1)=(0,3)$。求标准矩阵及 $T(2,0)$。

29.4 ★★　$A=\left[\begin{smallmatrix}1&1\\0&2\end{smallmatrix}\right]$ 是某变换的标准矩阵。输入、输出都改用基 $(1,1),(1,-1)$，求新矩阵。

29.5 ★★　求投影矩阵 $P=\frac{1}{2}\left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]$ 对应变换的核和像，并计算 $P(4,2)$。

29.6 ★★　对导数矩阵 $D=\left[\begin{smallmatrix}0&1&0\\0&0&2\end{smallmatrix}\right]$ 和零常数项积分矩阵 $E=\left[\begin{smallmatrix}0&0\\1&0\\0&1/2\end{smallmatrix}\right]$，计算 DE、ED，说明它们各自作用在哪个空间上。

29.7 ★★　证明：线性变换 T 是单射当且仅当 $\ker T=\{0\}$。

29.8 ★★★ T、S 的标准矩阵分别为 $A=\begin{bmatrix}1&1\\0&1\end{bmatrix}$、$B=\begin{bmatrix}1&0\\1&2\end{bmatrix}$。求 $S\circ T$ 的矩阵，并计算 $(S\circ T)(2,-1)$。

29.9 ★★★ 对 $T(x,y)=(x+2y,y)$，求逆变换及其矩阵，并验证两个复合都等于恒等变换。

29.10 ★ 求所有次数不超过 2 的实系数多项式 p，使 $p'(x)=3-4x$。说明这些解如何反映导数变换的核。

练习答案与提示

29.1 (a) 假。$T(x,y)=(x,y^2)$ 满足 $T(0,0)=0$，但 $T(0,2)=(0,4)\neq 2T(0,1)=(0,2)$。(b) 真，由定理 29.8。(c) 真，由命题 29.9。(d) 假：输入维数给列数，输出维数给行数，矩阵应为 2×3。

29.2 标准矩阵是 $A=\begin{bmatrix}2&-1&1\\1&0&2\end{bmatrix}$，所以 $T(1,-1,2)=(2+1+2,\,1+4)=(5,5)$。

29.3 令 $v_1=(1,1)$、$v_2=(1,-1)$。由 $e_1=\tfrac12(v_1+v_2)$、$e_2=\tfrac12(v_1-v_2)$，得 $T(e_1)=(1,2)$、$T(e_2)=(1,-1)$，故标准矩阵为 $\left[\begin{smallmatrix}1&1\\2&-1\end{smallmatrix}\right]$。又 $(2,0)=v_1+v_2$，所以 $T(2,0)=(2,1)+(0,3)=(2,4)$。

29.4 $P=\begin{bmatrix}1&1\\1&-1\end{bmatrix}$，$P^{-1}=\tfrac12\begin{bmatrix}1&1\\1&-1\end{bmatrix}$。因此 $P^{-1}AP=\begin{bmatrix}2&-1\\0&1\end{bmatrix}$。

29.5 $P(x,y)=\left(\frac{x+y}{2},\frac{x+y}{2}\right)$，故 $x+y=0$ 时且仅当输出为零，$\ker P=\operatorname{span}\{(1,-1)\}$；所有输出都在 $y=x$ 上且整条线都能取到，$\operatorname{im}P=\operatorname{span}\{(1,1)\}$。$P(4,2)=(3,3)$。

29.6 逐项相乘得

$$DE=\begin{bmatrix}1&0\\0&1\end{bmatrix}=I_2,\qquad ED=\begin{bmatrix}0&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

DE 表示先积分再求导，作用在 P_1 上，是恒等变换；ED 表示先求导再积分，作用在 P_2 上，把 $a+bx+cx^2$ 变为 $bx+cx^2$，即去掉常数项。

29.7 若 T 单射且 $T(v)=0=T(0)$，则 $v=0$。反过来，若 $\ker T=\{0\}$ 且 $T(u)=T(v)$，则 $T(u-v)=0$，于是 $u-v=0$，即 $u=v$。两方向都成立。

29.8 先做 T 再做 S，矩阵次序是 BA：

$$BA=\begin{bmatrix}1&0\\1&2\end{bmatrix}\begin{bmatrix}1&1\\0&1\end{bmatrix}=\begin{bmatrix}1&1\\1&3\end{bmatrix}.$$

逐步计算：$T(2,-1)=(1,-1)$，$S(1,-1)=(1,-1)$，所以 $(S\circ T)(2,-1)=(1,-1)$，也与 $BA(2,-1)$ 相同。

29.9 解 $u+2v=x$、$v=y$ 得 $T^{-1}(x,y)=(x-2y,y)$，其矩阵为 $A^{-1}=\begin{bmatrix}1&-2\\0&1\end{bmatrix}$。矩阵乘积验证为

$$AA^{-1}=\begin{bmatrix}1&-2+2\\0&1\end{bmatrix}=I,\qquad A^{-1}A=\begin{bmatrix}1&2-2\\0&1\end{bmatrix}=I.$$

因此两个复合都等于恒等变换；直接代入也得 $T^{-1}(T(x,y))=T(T^{-1}(x,y))=(x,y)$。

29.10 写 $p(x)=a+bx+cx^2$，则 $p'(x)=b+2cx$。比较系数得 $b=3$、$c=-2$，而 a 任意；全部解为 $p(x)=a+3x-2x^2$。自由的常数 a 正是导数变换的核 $\operatorname{span}\{1\}$。

第 30 讲

基变换与图像压缩

Change of Basis and Image Compression

对应材料：Session 3.7; Problem Set 31.1-31.2; ZoomNotes Part 8

本讲目标

(1) 会把向量从标准坐标换成任意有序基的坐标，并能由基矩阵解释 $x = Pc$。

(2) 会推导同一线性变换在新输入、输出基下的矩阵 $Q^{-1}AP$，并检验相似变换。

(3) 会验证 Haar 向量正交、求单位化基与信号系数，说明 DCT 和小波基为何有利于压缩。

(4) 会按奇异值降序截断 SVD，计算重建误差，并区分精确恢复、低秩近似与零系数量化。

30.1 坐标换了，向量变了吗?

向量 $x = (3, 1)$ 在标准基 (standard basis) 下写成 $3e_1 + e_2$。若改用 $b_1 = (1, 1)$、$b_2 = (1, -1)$，则

$$x = 2b_1 + b_2.$$

向量的几何位置没有移动，变的是描述它的两个系数：标准坐标是 $(3, 1)$，新坐标是 $(2, 1)$。图 30.1 把这两种描述放在同一张坐标图里。

定义 30.1 (有序基的坐标与基矩阵) 设 $\mathcal{B} = (b_1, \dots, b_n)$ 是 $\mathbb{R}^n$ 的一组有序基 (ordered basis)。唯一满足 $x = c_1b_1 + \cdots + c_nb_n$ 的列向量 $c = (c_1, \dots, c_n)^{\mathsf{T}}$ 称为 x 在 $\mathcal{B}$ 下的坐标向量 (coordinate vector)，记作 $[x]_{\mathcal{B}}$。以 $b_1, \dots, b_n$ 为列的矩阵 $P_{\mathcal{B}}$ 称为基矩阵 (basis matrix)。

按矩阵乘法的列观点，$P_{\mathcal{B}}c = c_1b_1 + \cdots + c_nb_n = x$。例如上面的基矩阵是 $P_{\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$，其行列式为 $-2 \neq 0$，确实可逆；而 $P_{\mathcal{B}} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$。零向量的坐标仍是零；在标准基下，基矩阵是 I，所以坐标与原分量相同。

「每个向量都有坐标」来自基张成整个空间；「坐标唯一」来自基向量线性无关。两条合在一起说明 $P_{\mathcal{B}}$ 可逆。对当前的二维基，

$$P_{\mathcal{B}}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad [(u, v)]_{\mathcal{B}} = \left(\frac{u+v}{2}, \frac{u-v}{2} \right)^T.$$

例如 $u=v$ 时第二坐标为零，向量沿 b_1；$u=-v$ 时第一坐标为零，向量沿 b_2。这两个边界情形可以直接检查坐标公式的正负号。

易错提醒　坐标向量不是原向量的另一个名字。把 $x=(3,1)$ 的新坐标误写成 $P_{\mathcal{B}}x=(4,2)$，就会把一次「从坐标回到向量」的计算当成「从向量求坐标」。正确方向是 $x=P_{\mathcal{B}}[x]_{\mathcal{B}}$，因此 $[x]_{\mathcal{B}}=P_{\mathcal{B}}^{-1}x$。

■ **例 30.1**　**求一组新坐标**　求 $z=(5,-1)$ 在 $\mathcal{B}=((1,1),(1,-1))$ 下的坐标，并代回验证。

解　思路是解线性组合 $z=c_1b_1+c_2b_2$。比较两个分量得 $c_1+c_2=5$、$c_1-c_2=-1$；相加得 $2c_1=4$，所以 $c_1=2$，再得 $c_2=3$。代回：$2(1,1)+3(1,-1)=(5,-1)$。因此 $[z]_{\mathcal{B}}=(2,3)^{\mathsf{T}}$。 ♦

如图 30.1 所示，实线先走 $2b_1$，再沿 b_2 的方向走一步，抵达 x。

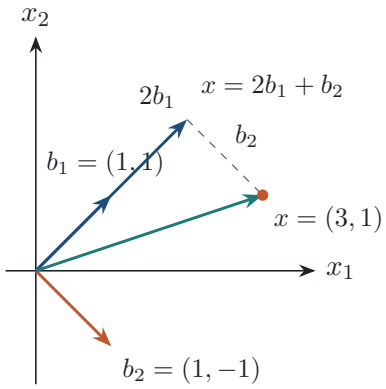


图 30.1　同一向量 $x=(3,1)$ 在标准基下的坐标为 $(3,1)$，在 $\mathcal{B}$ 下的坐标为 $(2,1)$。

橙点只有一个；两种坐标都指向它。蓝色实箭头给出 $2b_1$，灰色虚线把 b_2 平移到 $2b_1$ 的终点，最后得到 x。矩阵 $P_{\mathcal{B}}$ 的列就是基向量，因此一次矩阵乘法把新坐标合成为原坐标。

想一想　如果把基的顺序改为 $\mathcal{B}'=(b_2,b_1)$，向量 $x=(3,1)$ 的位置会变吗？坐标呢？

答　位置不变，因为 $x=2b_1+b_2=b_2+2b_1$。新顺序下第一项系数属于 b_2，所以 $[x]_{\mathcal{B}'}=(1,2)^{\mathsf{T}}$；列顺序和坐标顺序必须一起交换。

30.2　同一个变换为何换了矩阵？

上一节只换了向量的写法。现在固定线性变换 T，问它的矩阵为何也会改变。设旧输入基为 $\mathcal{V}$、旧输出基为 $\mathcal{Y}$，矩阵为 A；新输入基为 $\mathcal{V}'$、新输出基为 $\mathcal{Y}'$。若 P 把新输入坐标合成为旧输入坐标，Q 把新输出坐标合成为旧输出坐标，则同一向量经过两条路径：

$$\text{新输入坐标}c \xrightarrow{P} \text{旧输入坐标}Pc \xrightarrow{A} \text{旧输出坐标}APc \xrightarrow{Q^{-1}} \text{新输出坐标}Q^{-1}APc.$$

定理 30.2（输入、输出同时换基的矩阵公式）　设 $T: V \to W$ 在旧基下的矩阵为 A。若 P 是从新输入坐标到旧输入坐标的基矩阵，Q 是从新输出坐标到旧输出坐标的基矩阵，则 T 在新基下的矩阵为

$$B = Q^{-1}AP.$$

当输入、输出是同一空间且使用同一基矩阵 P 时，公式退化为相似变换 $B = P^{-1}AP$。

在任意正交单位基中，删去系数集合 J 造成的平方重建误差是 $\sum_{j\in J}|c_j|^2$。所以只看每个系数是否「小」还不够：若有 100 个系数都为 $1/10$，删去它们的总误差范数是 $\sqrt{100(1/10)^2} = 1$。SVD 的尾部公式把所有舍弃方向的影响合在一起，给出可比较的误差预算。

直观理解　沿图 30.2 的四个节点走一遍：先把新坐标翻译成旧坐标，应用原来的 A，再把输出翻译回新坐标。中间的几何变换没有改变，只有两次「翻译」改变了坐标标签。

证明　设新输入坐标为 c。由基矩阵定义，旧输入坐标是 Pc。在旧基下应用 T 后，旧输出坐标是 APc。另一方面，若新输出坐标为 d，则旧输出向量也可写成 Qd。因此 $Qd = APc$。基矩阵的列是一组基向量，线性无关，所以 Q 可逆；左乘 Q^{-1} 得 $d = Q^{-1}APc$。这对每个 c 都成立，故新矩阵就是 $B = Q^{-1}AP$。 ■

条件「P, Q 是基矩阵」不能去掉：若 $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$，它把两个坐标压到一条轴上，不可逆；此时 $Q^{-1}AP$ 不能表示所有输入坐标的同一变换。基必须由线性无关且张成空间的向量组成。

■ **例 30.2　输入、输出基都改变**　令 $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$，输入新基的基矩阵为 $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$，输出新基的基矩阵为 $Q = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$。求新矩阵 B，并用新坐标 $c = (2,1)^{\mathsf{T}}$ 检查两条路径一致。

解　先按公式计算。$Q^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$，所以

$$B = Q^{-1}AP = \begin{bmatrix} 2 & 2 \\ 1 & -1 \end{bmatrix}.$$

新坐标路径给出 $Bc = (6,1)^{\mathsf{T}}$。另一条路径先算 $Pc = (3,1)^{\mathsf{T}}$，再算 $A(Pc) = (7,1)^{\mathsf{T}}$；最后 $Q^{-1}(7,1)^{\mathsf{T}} = (6,1)^{\mathsf{T}}$。两者相同，说明公式把每一步坐标翻译都记录下来了。 ♦

■ **例 30.3　输入、输出使用同一新基**　仍取 $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$，输入、输出均用基 $\mathcal{B} = ((1,1),(1,-1))$。求新矩阵及坐标 $c = (2,1)^T$ 的像。

解　两端的基矩阵相同，故 $B = P^{-1}AP$，其中 $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$、$P^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$。相乘得

$$B = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \qquad Bc = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

换回标准坐标：$P(4,3)^T = (7,1)^T$；直接计算 $AP(2,1)^T = A(3,1)^T = (7,1)^T$，结果相同。这说明相似矩阵的条目改了，变换作用于向量的结果没有改。 ♦

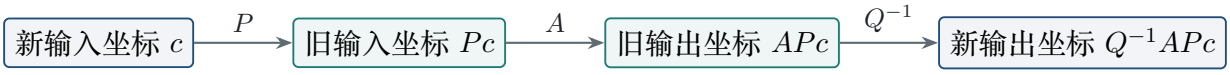


图 30.2　换基公式把新输入坐标依次映到新输出坐标。

方法：先标坐标，再乘矩阵　遇到换基题，先写清每个坐标属于哪一组基；输入端乘 P，做变换乘 A，输出端乘 Q^{-1}。同一基时两端的翻译合成 $P^{-1}AP$。

想一想　若只改变输入基而不改变输出基，公式是什么？

答　此时 $Q = I$，所以新矩阵是 AP。若只改变输出基，则 $P = I$，新矩阵是 $Q^{-1}A$。两个方向都改才得到完整的 $Q^{-1}AP$。

30.3　图像为什么能换一种坐标记录？

一帧 512×512 的黑白视频有 $512^2 = 262144$ 个亮度值，因此可看成 $\mathbb{R}^{262144}$ 中的一个向量。像素标准基逐像素存值；若整帧都是亮度 a，就要重复记录 262144 次 a。若新基含全 1 向量，常亮画面只需一个非零系数，其他坐标为零。压缩的关键是让常见图像在新基下有许多零或很小的系数。

设 $W = [w_1 \ \cdots \ w_n]$ 的列是一组基。把向量写成 $x = c_1w_1 + \cdots + c_nw_n$，就是 $x = Wc$，所以 $c = W^{-1}x$。若这些基向量两两正交且单位长度，则 $W^TW = I$，从而 $W^{-1} = W^T$，只需一次转置乘法就能分析或重建。不同数据适合不同基：整幅画面、照片纹理和字幕边缘的结构并不相同。

JPEG 对每个 8×8 像素块做实数**离散余弦变换**（discrete cosine transform, DCT）；一块有 64 个像素值和 64 个余弦系数。它与复**傅里叶基**（Fourier basis）有相近的频率分解思想，但 JPEG 使用实余弦基。量化会把一些幅度较小的系数舍入为零或较粗的数，从而减小存储量；被改动的系数会造成像素误差，所以这是**有损压缩**（lossy compression）。512×512 画面按不重叠的 8×8 块划分，共有 4096 块。

为具体计算一个块，采用正交归一的 DCT-II 基。令 $j = 0, \dots, 7$，定义

$$\phi_0(j) = \frac{1}{\sqrt{8}}, \qquad \phi_u(j) = \frac{1}{2}\cos\left(\frac{\pi(2j+1)u}{16}\right) \quad (1 \le u \le 7).$$

二维基图案是 $\phi_u(j)\phi_v(k)$，相应系数为

$$C_{uv} = \sum_{j=0}^{7}\sum_{k=0}^{7} x_{jk}\phi_u(j)\phi_v(k), \qquad 0 \le u, v \le 7.$$

这 64 个图案两两正交且长度为 1，因此完整变换保留全部像素信息。

■ 例 30.4　**常亮块经过 DCT 和量化**　一个 8×8 块的每个像素都是 3。求其 DCT 系数，并在简化的统一量化步长 $q=5$ 下重建。这里用统一步长便于手算；JPEG 实际按系数位置使用量化表。

解　先用正交基计算系数。常量列满足 $\sum_{j=0}^{7}\phi_0(j)=\sqrt{8}$，而每个 $u>0$ 的余弦列和为零，所以只有 C_{00} 非零，且

$$C_{00}=3\left(\sum_{j=0}^{7}\phi_0(j)\right)^2=24.$$

按 $q=5$ 量化，24 舍入到 25，其余系数仍为零。逆变换得到每个像素 $25\phi_0(j)\phi_0(k)=25/8=3.125$，与原值相差 $1/8$；因此整块的 Frobenius 误差为 $\sqrt{64(1/8)^2}=1$。误差来自量化，完整 DCT 本身没有丢失信息。 ♦

每个 DCT 向量描述一种水平、垂直方向的余弦起伏。缓慢变化的亮度通常集中在低频系数，细边缘和快速变化需要较高频率；自然图像常有大量很小的高频系数。JPEG 量化表会让部分高频系数用更粗的步长表示，但并非把所有高频都删掉。若图像里有清晰文字或细线，它们恰好可能需要较大的高频系数。

复 Fourier 基给出另一种频率坐标。令 $\omega=e^{2\pi i/8}$，并对 $k=0,\dots,7$ 定义

$$f_k=(1,\omega^k,\omega^{2k},\dots,\omega^{7k})^T.$$

若 $j\neq k$，则 $f_j^*f_k=\sum_{r=0}^{7}(\omega^{k-j})^r=0$；若 $j=k$，内积为 8。所以 $F=[f_0\ \cdots\ f_7]$ 满足 $F^*F=8I$，其单位化列构成 $\mathbb{C}^8$ 的正交单位基。

■ 例 30.5　**单频信号的 Fourier 坐标**　令信号 $x=f_3$。求它在标准坐标和 Fourier 基下的坐标，并比较非零项数。

解　标准坐标就是 $x=(1,\omega^3,\omega^6,\dots,\omega^{21})^T$，八个分量的模都为 1。由 $F^{-1}=\frac{1}{8}F^*$，Fourier 坐标为 $F^{-1}x=e_4$，只有频率 $k=3$ 的系数非零。信号与基中的振荡模式吻合时，坐标会变稀疏；像素向量本身并没有因此改变。 ♦

只改变基并不丢失信息：若完整保留所有系数，接收方可用逆变换重建原像素。文件还要说明哪些系数被保留、采用什么量化尺度；稀疏坐标为编码提供空间，但实际字节数取决于非零位置和取值怎样表示。把「基变换可逆」与「截断或量化有损」分开，才能准确描述压缩流程。

DCT 的每个振荡模式覆盖整个 8×8 块；Haar 的支撑从整段逐层缩到相邻像素。前者适合描述平滑的周期起伏，后者能把局部突变定位到某一小段。没有一组基能对所有图像都产生稀疏坐标，这也是 JPEG 与小波编码各有适用场景的原因。

另一个局部化的选择是 $\mathbb{R}^8$ 的**Haar 小波基**（Haar wavelet basis）。把八个基向量作为列排成

矩阵：

$$H=\begin{bmatrix}1&1&1&0&1&0&0&0\\1&1&1&0&-1&0&0&0\\1&1&-1&0&0&1&0&0\\1&1&-1&0&0&-1&0&0\\1&-1&0&1&0&0&1&0\\1&-1&0&1&0&0&-1&0\\1&-1&0&-1&0&0&0&1\\1&-1&0&-1&0&0&0&-1\end{bmatrix}.$$

若两列的支撑互不相交，内积为零；若一列支撑在另一列的常值区间内，正负项成对抵消。例：$h_1^Th_2=4-4=0$，而 h_2 在 h_5 的两个非零位置都等于 1，所以 $h_2^Th_5=1-1=0$。同一层的左右局部差也有不相交支撑。逐列计算得到

$$H^TH=\operatorname{diag}(8,8,4,4,2,2,2,2).$$

因此八列非零且两两正交，构成一组基；各列长度依次为 $\sqrt{8},\sqrt{8},2,2,\sqrt{2},\sqrt{2},\sqrt{2},\sqrt{2}$。把每一列除以自己的长度，就得到正交单位基。这正是习题 31.1 要求的正交性验证与单位化。

令 $N=\operatorname{diag}(1/\sqrt{8},1/\sqrt{8},1/2,1/2,1/\sqrt{2},1/\sqrt{2},1/\sqrt{2},1/\sqrt{2})$，则单位化后的基矩阵为 $U=HN$，并且 $U^TU=N^TH^THN=I_8$。对任意八像素信号，单位基坐标可直接算作 $\hat{c}=U^Tx$；未单位化时则需用 $H^{-1}=\operatorname{diag}(1/8,1/8,1/4,1/4,1/2,1/2,1/2,1/2)H^T$。这一处比例差异很容易漏掉：把单位基的点积公式用于未归一化的列，会把系数放大或缩小。

正交单位基还保证能量守恒：$\|x\|_2^2=\|U\hat{c}\|_2^2=\hat{c}^TU^TU\hat{c}=\|\hat{c}\|_2^2$。删去一些系数后，像素平方误差恰好等于被删系数的平方和。对未归一化基，这个结论要先把各方向按长度标准化；否则直接比较系数大小会把不同长度的方向混在一起。

图 30.3 只画出 Haar 基中的三个层次：全局常量、左右两半的差、最局部的相邻像素差。蓝格表示系数 $+1$，橙格表示 -1，空格表示 0。

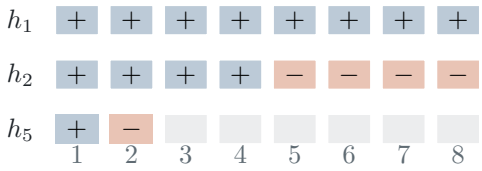


图 30.3　Haar 向量从全局平均 h_1 到局部差分 h_5，支撑范围逐渐缩小。

从图中看，h_1 描述整体亮度，h_2 描述左半与右半的明暗差，h_5 只比较前两个像素。若图像大块区域平滑，许多细尺度差分会很小；保存少数大系数并丢弃其余系数，就能用较少数据近似重建。

例 30.6　把两级亮度写成 Haar 系数　令八个像素的亮度向量为 $x=(9,9,9,9,1,1,1,1)^T$。求它在未归一化 Haar 基 H 下的系数，并重建 x。

解　对未归一化的正交列，系数为 $c_i=(h_i^Tx)/(h_i^Th_i)$。全局列和左右差列分别给出 $c_1=(4\cdot9+4\cdot1)/8=5$、$c_2=(4\cdot9-4\cdot1)/8=4$。其余局部差分只比较常值小块内相等的亮度，内积均为零。

因此 $c=(5,4,0,0,0,0,0,0)^T$，且 $Hc=5h_1+4h_2=x$。若单位化前两列，系数改为 $5\sqrt{8}$、$4\sqrt{8}$，重建出的像素仍不变。 ♦

换基本身不压缩数据：可逆的 W 会把 n 个数换成另 n 个数。压缩来自系数有结构，量化后许多值能用零或少量比特表示。视频还可利用时间上的相似性，先预测下一帧，再编码帧差；若差图更稀疏，所需数据会进一步减少。

想一想　只要使用正交基，就一定能压缩图像吗?
答　不一定。正交只让系数可用 W^Tx 快速算出，并保持长度；若系数仍然都很大，仍要保存很多数据。压缩还依赖基与图像结构相配，并依赖系数截断或量化带来的可接受误差。

30.4　保留重要方向：低秩截断与误差

一个 $m\times n$ 像素矩阵可能有很多条目，但主要变化常集中在少数方向上。利用奇异值分解 (singular value decomposition, SVD)，把这些方向按重要程度排序：

$$A=\sum_{i=1}^{r}\sigma_i u_i v_i^T,\qquad \sigma_1\ge\sigma_2\ge\cdots\ge\sigma_r>0.$$

每个 $u_iv_i^T$ 是一个秩 1 的图案，σ_i 是它的权重。保留前 k 项得到 $A_k=\sum_{i=1}^{k}\sigma_iu_iv_i^T$，其秩至多为 k。k 越小，存储的方向越少，通常误差越大。

用**Frobenius 范数**（Frobenius norm）衡量像素误差：$\|E\|_F^2=\sum_{i,j}e_{ij}^2$。平方根反映整体误差；除以 $\sqrt{mn}$ 后是每个像素的均方根误差。SVD 截断有精确误差公式，而且在所有秩不超过 k 的矩阵中误差最小。

若关心相对误差，则

$$\frac{\|A-A_k\|_F^2}{\|A\|_F^2}=\frac{\sum_{i>k}\sigma_i^2}{\sum_{i=1}^{r}\sigma_i^2}.$$

它表示被舍弃的总平方能量占原矩阵平方能量的比例。较大的低频或粗尺度方向通常先保留，尾部比例给出全局误差预算；它不表示误差会均匀散布在图像上，实际视觉影响还取决于残差出现在平坦区域还是关键边缘。

定理 30.3（最佳秩 k 近似与尾部能量）　设 $A=\sum_{i=1}^{r}\sigma_iu_iv_i^T$ 是奇异值按降序排列的实矩阵 SVD，令 $A_k=\sum_{i=1}^{k}\sigma_iu_iv_i^T$，其中 $0\le k\le r$。则对任意满足 $\operatorname{rank}B\le k$ 的矩阵 B，

$$\|A-A_k\|_F^2=\sum_{i=k+1}^{r}\sigma_i^2\le\|A-B\|_F^2.$$

直观理解　SVD 把矩阵拆成彼此正交的秩 1 图案。删掉某些图案时，误差平方会把被删权重的平方相加；若只能留 k 个方向，应先留权重最大的 k 个。

思路　先证明截断后的残差平方等于尾部平方和，这只需用秩 1 图案的正交性。再任取秩至多 k 的 B，把 $A-B$ 按 B 的列空间正交分解，并用它的维数至多 k 证明至多能保留前 k 个方向的能

量。

证明　**截断误差。**Frobenius 内积为 $\langle X,Y\rangle_F=\operatorname{tr}(X^TY)$。对 $i\neq j$，

$$\langle u_iv_i^T,u_jv_j^T\rangle_F=(u_i^Tu_j)(v_i^Tv_j)=0.$$

所以这些秩 1 项两两正交，且 $\|u_iv_i^T\|_F=1$。从 A 中删掉 $i\le k$ 的项后，残差的平方范数正是 $\sum_{i>k}\sigma_i^2$。

任意秩至多 k 的近似。令 $L=\mathrm{Col}(B)$，P_L 为到 L 的正交投影。P_LA-B 的每列在 L 中，$(I-P_L)A$ 的每列在 $L^\perp$ 中，故两项正交：

$$\|A-B\|_F^2=\|(I-P_L)A\|_F^2+\|P_LA-B\|_F^2\ge\|A\|_F^2-\|P_LA\|_F^2.$$

以下先设 $1\le k<r$；$k=0$ 与 $k=r$ 的边界在本证明末尾处理。令 $\alpha_i=\|P_Lu_i\|^2$。由 SVD 和 v_i 两两正交，

$$\|P_LA\|_F^2=\sum_{i=1}^{r}\sigma_i^2\alpha_i.$$

每个 α_i 在 $[0,1]$ 内。取 L 的一组标准正交基 $q_1,\dots,q_d$，则 $\sum_i\alpha_i=\sum_{j=1}^{d}\sum_i|q_j^Tu_i|^2\le d=\dim L\le k$，内层不等式来自 $u_1,\dots,u_r$ 正交归一。设 $s_i=\sigma_i^2$，则 $s_1\ge\cdots\ge s_r$。因此 $\sum_{i>k}\alpha_i\le\sum_{i\le k}(1-\alpha_i)$；尾部每个 $s_i\le s_k$，前 k 个 $s_i\ge s_k$，

$$\sum_{i>k}s_i\alpha_i\le s_k\sum_{i>k}\alpha_i\le s_k\sum_{i\le k}(1-\alpha_i)\le\sum_{i\le k}s_i(1-\alpha_i).$$

整理得 $\|P_LA\|_F^2\le\sum_{i=1}^{k}\sigma_i^2$。代回前式，$\|A-B\|_F^2\ge\sum_{i>k}\sigma_i^2=\|A-A_k\|_F^2$。若 $k=0$，唯一可行矩阵是 $B=0$，而 $A_0=0$；若 $k=r$，$A_r=A$ 且误差为零。两种边界也满足结论。 ■

奇异值降序是定理的必要约定。若 $A=\mathrm{diag}(1,4)$ 却保留第一列方向，只留下 $\mathrm{diag}(1,0)$，误差为 4；保留第二列方向得到 $\mathrm{diag}(0,4)$，误差为 1。相同的秩预算下，应保留较大的奇异值。

■ **例 30.7**　**一个 2×2 灰度块的秩一近似**　把矩阵 $A=\begin{bmatrix}3&1\\1&3\end{bmatrix}$ 看成四个像素。求奇异值、最佳秩一近似 A_1、Frobenius 误差和相对误差。

解　矩阵对称，向量 $v_1=u_1=(1,1)^T/\sqrt{2}$ 对应特征值 4，$v_2=u_2=(1,-1)^T/\sqrt{2}$ 对应特征值 2。两者为正，所以奇异值按降序为 $4,2$。保留第一项：

$$A_1=4u_1v_1^T=\begin{bmatrix}2&2\\2&2\end{bmatrix},\qquad A-A_1=\begin{bmatrix}1&-1\\-1&1\end{bmatrix}.$$

残差有四个大小为 1 的像素误差，故 $\|A-A_1\|_F=2$。原矩阵平方范数为 $3^2+1^2+1^2+3^2=20$，相对误差为 $2/\sqrt{20}=1/\sqrt{5}\approx0.447$；保留的奇异值平方占总能量 $16/20=80\%$。这项近似把对角较亮、非对角较暗的块改成统一亮度 2，每个像素的均方根误差为 1。 ♦

这个 2×2 例子用于看清误差图案，并不节省存储：原矩阵有 4 个像素，秩一因子需要 $2+2+1=5$ 个数。图像行列都很大而秩较低时，因子表示才可能减少参数量；实际编码还要计入量化和索

引开销。如图 30.4 所示，原图的两个较亮像素位于对角线上；秩一近似只保留整体亮度，因而把四个位置都改为 2。

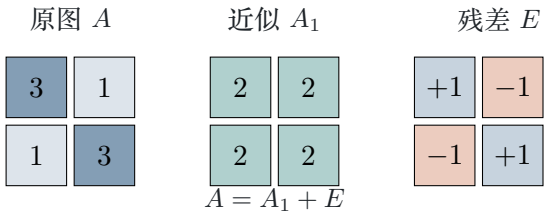


图 30.4　矩阵 A 的最佳秩一近似为全 2 矩阵，四个像素误差为 $+1, -1, -1, +1$。

蓝色深浅标原亮度，青色格表示近似亮度，残差中的蓝、橙分别表示正误差和负误差。Frobenius 范数把四个误差平方相加；它优化的是平方像素误差，并不直接衡量人眼对边缘或纹理的敏感程度。该矩阵用于说明 SVD 低秩近似，JPEG 的 DCT 系数阈值是另一种操作。

想一想　为何这里的平方误差是 4，而不是 1？

答　每个像素差为 1 或 -1，共有四个位置，所以 $\|E\|_F^2 = 1^2 + (-1)^2 + (-1)^2 + 1^2 = 4$；误差范数是平方根 2。除以四个像素数再开方，均方根误差才是 1。

30.5　怎样挑一组合适的基？

所有 2×2 实矩阵组成四维向量空间 $M_2(\mathbb{R})$。最直接的一组基由四个标准矩阵组成：

$$E_{11},\ E_{12},\ E_{21},\ E_{22}.$$

其中 E_{ij} 只有第 (i,j) 个位置为 1。任意矩阵可唯一写成

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = aE_{11} + bE_{12} + cE_{21} + dE_{22}.$$

也可以选择

$$D_+ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\quad D_- = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},\quad S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\quad K = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

在 Frobenius 内积 $\langle X, Y \rangle_F = \sum_{i,j} x_{ij} y_{ij}$ 下，这四个矩阵两两正交，平方长度都为 2，所以线性无关；四维空间中四个线性无关向量构成一组基。

D_+ 记录两个对角元的共同部分，D_- 记录对角差；S 和 K 分别记录对称与反对称的非对角部分。因而一个矩阵的系数直接显出它为何对称、对角或三角。

■ **例 30.8　用结构化基表示矩阵**　把 $M = \begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix}$ 写成 D_+, D_-, S, K 的线性组合，并比较两组基对特殊矩阵的描述。

解　令 $M = aD_+ + dD_- + sS + kK$。逐项比较得到 $a+d=4$、$a-d=2$、$s+k=2$、$s-k=0$，所以 $a=3$、$d=1$、$s=1$、$k=1$。因此 $M = 3D_+ + D_- + S + K$，四个系数都非零。一般矩阵 $\begin{bmatrix} p & q \\ r & t \end{bmatrix}$ 的四个系数为 $\frac{p+t}{2}, \frac{p-t}{2}, \frac{q+r}{2}, \frac{q-r}{2}$，可直接看出矩阵的结构如何变成坐标里的零。 ♦

■ **例 30.9**　**对称矩阵的两种坐标**　令 $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$。分别求它作为 $M_2(\mathbb{R})$ 中向量时在标准基、结构化基 (D_+, D_-, S, K) 下的坐标，再求它作为线性变换在基 $((1,1),(1,-1))$ 下的矩阵。

解　逐项比较可得 $A = 3D_+ + S$，所以在结构化基下坐标为 $(3,0,1,0)^T$；在标准基下坐标为 $(3,1,1,3)^T$。另一方面，$A(1,1)^T = 4(1,1)^T$、$A(1,-1)^T = 2(1,-1)^T$，故以这两个向量为列的 P 满足

$$P^{-1}AP = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}.$$

前两组坐标把矩阵当作四维空间中的一个向量，最后一个矩阵表示线性变换在新输入、输出基下的作用；它们描述的是不同的坐标对象。 ♦

对角矩阵在两组基下都只需两个基矩阵；标准基更适合上三角或下三角矩阵，因为相应地少用 E_{21} 或 E_{12}。结构化基更适合对称矩阵：若 $q=r$，则 K 的系数 $(q-r)/2=0$，只需 D_+, D_-, S 三项；标准基通常需要四个矩阵项，其中 E_{12}, E_{21} 系数相等。原习题 31.2 的判断取决于所选的第二组基；这里这组基把对称条件变成了一个坐标为零。

图像压缩也遵循同一取舍：若已知数据常含平滑块，选 DCT 或小波基能让系数稀疏；若需要压缩许多相关行列，SVD 会找出最重要的低秩方向。SVD 适合图像矩阵的全局相关结构，但计算完整分解有成本。对 $m \times n$ 矩阵的秩 k 表示，存储 k 个奇异值和两组奇异向量需约 $k(m+n+1)$ 个数；只有这个数量明显小于 mn，低秩表示才节省条目数。特征向量基则适合把线性变换对角化，计算该基的代价可能高于直接采用快速 Fourier 或小波变换。

■ **例 30.10**　**低秩表示的条目数临界值**　对 512×512 画面，用秩 k 的 SVD 因子表示需要 $k(512+512+1)$ 个数。求它少于逐像素表示时允许的最大整数 k。

解　因子表示需 $1025k$ 个数，逐像素表示需 262144 个。$k=255$ 时需 261375 个，少 769 个；$k=256$ 时需 262400 个，反而多 256 个。数量随 k 增长，因此最大可节省参数的秩是 255。这个门槛比较的是实数个数，实际文件大小还取决于量化和编码。 ♦

方法：先问数据有什么结构　挑基前先问哪些模式反复出现：整体亮度、局部边缘、矩阵对称性或变换的特征方向。好的基让常见对象的坐标更稀疏，且坐标变换本身也能高效计算。

想一想　若某组基能让一个矩阵的坐标很稀疏，是否就一定是最好的压缩基？

答　还要看变换和重建成本、数据是否普遍具有这种结构，以及丢弃小系数后产生的误差。基若难以快速计算，或对常见图像并不稀疏，少量非零系数也未必降低实际存储量。好的选择同时考虑表示稀疏、计算速度和误差。

本讲小结

- $x = P_{\mathcal{B}}[x]_{\mathcal{B}}$，所以 $[x]_{\mathcal{B}} = P_{\mathcal{B}}^{-1}x$；基列描述新坐标怎样合成原向量。
- 同一变换在新输入、输出基下的矩阵为 $Q^{-1}AP$；同一基时为 $P^{-1}AP$。
- 正交单位基矩阵满足 $W^{-1} = W^T$；未单位化 Haar 基的 Gram 矩阵为 $H^TH = \mathrm{diag}(8,8,4,4,2,2,2,2)$。
- JPEG 将每个 8×8 块换成 64 个 DCT 系数；量化小系数会产生误差，稀疏坐标才提供压缩空间。
- 若 $A = \sum_i \sigma_i u_i v_i^T$ 且奇异值降序，保留前 k 项的平方误差是 $\sum_{i>k}\sigma_i^2$，并在秩不超过 k 的矩阵中最小。
- $M_2(\mathbb{R})$ 的基可按结构选择；好基兼顾稀疏表示、计算速度和重建误差。

练习

30.1 ★　判断真假，假的给出反例：(a) 新坐标 c 与原坐标 x 总满足 $c = Px$；(b) 若 P, Q 分别把新输入、输出坐标变回旧坐标，则新矩阵是 $Q^{-1}AP$。

30.2 ★　在基 $\mathcal{B} = ((1,1),(1,-1))$ 下求 $(5,-1)$ 的坐标，并代回检查。

30.3 ★★　令 $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$，输入基矩阵 $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$，输出基矩阵 $Q = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$。求新矩阵并计算新输入坐标 $(2,1)^T$ 的输出坐标。

30.4 ★★　证明本讲列出的八个 Haar 向量两两正交，并把它们各自除以长度，得到一组正交单位基（原习题 31.1）。

30.5 ★★　用 Haar 基表示 $x = (7,7,7,7,3,3,3,3)^T$，并重建它。

30.6 ★★　给出 $M_2(\mathbb{R})$ 的标准基和本讲的结构化基；说明哪组对角矩阵、三角矩阵、对称矩阵更方便（原习题 31.2）。

30.7 ★★　对 $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ 求最佳秩一近似和 Frobenius 误差，并解释每个近似像素。

30.8 ★★　若误把 $\mathrm{diag}(1,4)$ 的第一列当成最大奇异方向，求截断误差；再与保留最大奇异方向比较。

30.9 ★★★　证明：若 u_i 两两正交归一、v_i 两两正交归一，则 $u_iv_i^T$ 两两正交；据此说明在只保留 k 个给定 SVD 项时，保留最大的 k 个奇异值会使 Frobenius 误差最小。

30.10 ★★★ 一张 512×512 灰度图以秩 20 的 SVD 因子存储，需要保存多少个实数？与逐像素保存相比，条目数减少了多少？忽略量化、文件头和索引。

练习答案与提示

30.1 (a) 假。$P(3,1)^T=(4,2)^T$，而 $(3,1)$ 在该基下的坐标是 $(2,1)^T$。(b) 真：新输入坐标先经 P 回到旧坐标，经 A 后再经 Q^{-1} 回到新输出坐标。

30.2 设坐标为 $(c_1,c_2)^T$。由 $c_1+c_2=5$、$c_1-c_2=-1$ 得 $c_1=2,c_2=3$；代回 $2(1,1)+3(1,-1)=(5,-1)$。

30.3 $B=Q^{-1}AP=\begin{bmatrix}2&2\\1&-1\end{bmatrix}$，所以 $B(2,1)^T=(6,1)^T$。直接算 $P(2,1)^T=(3,1)^T$、$A(3,1)^T=(7,1)^T$，再用 Q^{-1} 得 $(6,1)^T$。

30.4 h_1 与 h_2 的内积是 $4-4=0$；其余任意一列要么支撑与另一列不相交，要么落在较粗列的常值区间内且自身正负项成对抵消，所以所有不同列内积均为零。由 $H^TH=\operatorname{diag}(8,8,4,4,2,2,2,2)$，列长依次是 $\sqrt{8},\sqrt{8},2,2,\sqrt{2},\sqrt{2},\sqrt{2},\sqrt{2}$；逐列除以对应长度后，Gram 矩阵变成 I_8。

30.5 $c_i=(h_i^Tx)/(h_i^Th_i)$，故 $c=(5,2,0,0,0,0,0,0)^T$。$Hc=5h_1+2h_2=(7,7,7,7,3,3,3,3)^T$。

30.6 标准基为 $E_{11},E_{12},E_{21},E_{22}$；结构化基为 D_+,D_-,S,K。对角矩阵在两组基下都只用两个系数；三角矩阵在标准基下少一个非对角矩阵；对称矩阵满足 $q=r$，结构化基中的 K 系数 $(q-r)/2$ 为零，因此结构化基更方便。

30.7 $A_1=\begin{bmatrix}2&2\\2&2\end{bmatrix}$，残差为 $\begin{bmatrix}1&-1\\-1&1\end{bmatrix}$，故 $\|A-A_1\|_F=2$。近似矩阵把四个亮度都设为 2，误差平方和为 4。

30.8 保留第一方向会留下 $\operatorname{diag}(1,0)$，残差为 $\operatorname{diag}(0,4)$，误差是 4。最大奇异方向是第二列，保留 $\operatorname{diag}(0,4)$ 后误差为 1，所以方向必须按奇异值降序。

30.9 令 $R_i=u_iv_i^T$。则 $\langle R_i,R_j\rangle_F=(u_i^Tu_j)(v_i^Tv_j)=0$（$i\neq j$），且 $\|R_i\|_F=1$。保留指标集合 J 后，误差平方为 $\sum_{i\notin J}\sigma_i^2$；在 $|J|=k$ 时，按降序选前 k 项使保留平方和最大、舍弃平方和最小。

30.10 因子表示需要 $20(512+512+1)=20500$ 个实数。逐像素需 $512^2=262144$ 个；因子表示约为原来的 7.82%，条目数约减少 92.18%。这是未计编码开销的参数个数比较。

第 31 讲

左逆、右逆与伪逆

Left and Right Inverses; Pseudoinverse

对应材料：MIT 18.06SC Session 3.8（Lecture 33）讲义；Problem Set 32.1–32.2。

本讲目标

(1) 会按秩判断矩阵是否有左逆、右逆或双侧逆，并能写出标准构造。

(2) 会从满列秩推出 $A^{\mathsf{T}}A$ 可逆，并用左逆求最小二乘解与列空间投影。

(3) 会解释满行秩矩阵为何对每个 b 都有解，并构造右逆与行空间投影。

(4) 理解 A 在行空间到列空间之间可逆，并会用 SVD 求伪逆。

(5) 会用 $A^{+}b$ 求一般最小二乘问题的最小范数解，并区分投影与精确解。

31.1 三种逆：方阵、左逆与右逆

方阵可逆时，$A^{-1}A=I$ 且 $AA^{-1}=I$。矩形矩阵的两个乘积尺寸不同，所以我们要分别问：能不能从左边乘回单位阵，或从右边乘回单位阵？

定义 31.1（双侧逆与单侧逆） 设 $A\in\mathbb{R}^{m\times n}$。若 $B\in\mathbb{R}^{n\times m}$ 满足 $BA=I_n$，称 B 为 A 的**左逆**（left inverse）；若满足 $AB=I_m$，称 B 为 A 的**右逆**（right inverse）。当 $m=n$ 且两个等式都成立时，B 是 A 的**双侧逆**（two-sided inverse），记作 A^{-1}。

维数先告诉我们两侧逆各在做什么：BA 作用在输入空间 $\mathbb{R}^n$，AB 作用在输出空间 $\mathbb{R}^m$。若 A 有左逆，$Ax=0$ 就会推出 $x=BAx=0$；若有右逆，$AB=I_m$ 表示 A 的列能组合出每个 $b\in\mathbb{R}^m$。

例如 $A=\begin{bmatrix}1\\1\end{bmatrix}$ 的列非零，故可有左逆；矩阵 $C=\begin{bmatrix}1 & 2\end{bmatrix}$ 的列空间是 $\mathbb{R}$，故可有右逆。零矩阵 $[0]$ 既没有左逆也没有右逆，因为任何乘积仍是 $[0]$。

定理 31.2（单侧逆的秩判据） 设 $A\in\mathbb{R}^{m\times n}$，$r=\operatorname{rank}(A)$。则

(1) A 有左逆，当且仅当 $r=n$（满列秩）；

(2) A 有右逆，当且仅当 $r=m$（满行秩）；

(3) A 有双侧逆，当且仅当 $m=n=r$。

思路　左逆等式先约束零空间：它只能是 $\{0\}$，这等价于列独立。反过来，列独立使 Gram 矩阵 $A^{\mathsf{T}}A$ 正定，因而可逆；右逆把同一论证用于 A^{T}。方阵时，满列秩和满行秩合在一起就是 $r=m=n$。

证明　**左逆。**若 $LA=I_n$，且 $Ax=0$，则 $x=LAx=0$，所以 $N(A)=\{0\}$，即列独立且 $r=n$。

反过来，若 $r=n$，则对任意非零 $x\in\mathbb{R}^n$，$Ax\neq 0$，从而

$$x^{\mathsf{T}}A^{\mathsf{T}}Ax=\|Ax\|^2>0.$$

所以 $A^{\mathsf{T}}A$ 正定且可逆。取 $L=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$，便有 $LA=I_n$。

右逆。将左逆结论用于 A^{T}：A^{T} 有左逆当且仅当 $\operatorname{rank}(A^{\mathsf{T}})=m$。转置该左逆可得 A 的右逆；具体地，$R=A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}$ 满足 $AR=I_m$。

若 A 有双侧逆，则同时满列秩、满行秩，故 $m=n=r$。反之，若 $m=n=r$，左逆公式给出 $LA=I$，而方阵单侧逆也是双侧逆：A 可逆且 $L=A^{-1}$。 ■

这些秩判据也直接回答 $Ax=b$ 有几个解。满列秩时 $N(A)=\{0\}$，所以方程至多有一个解；是否存在取决于 b 是否落在 $\operatorname{Col}(A)$。满行秩时 $\operatorname{Col}(A)=\mathbb{R}^m$，每个 b 都能达到；若 $n>m$，解集是一个维数为 $n-m$ 的仿射空间。方阵满秩时，这两个条件同时成立，解对每个 b 都存在且唯一。

满秩条件不能省略：$A=\begin{bmatrix}1&0\\0&0\end{bmatrix}$ 的秩为 $1<2$，$e_2\in N(A)$，所以不可能有左逆；它的第二行是零行，所以列空间不是 $\mathbb{R}^2$，也不可能有右逆。

这两种单侧逆分别对应两种不同的要求。左逆要从 Ax 唯一恢复 x，所以输入方向不能被压到零；即使 $m>n$，只要列独立，A 在自己的列空间上仍可逆。标准公式还把列空间以外的分量先正交投影掉，再恢复系数。右逆则要为每个输出 b 选出一个输入；满行秩保证每个 b 可达，但当 $n>m$ 时，输出一般有整族原像。

■ **例 31.1**　对 $A=\begin{bmatrix}1\\1\end{bmatrix}$，求一个左逆，并检查它能否同时作为右逆。

解　按公式计算 $A^{\mathsf{T}}A=[2]$，所以 $L=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}=\begin{bmatrix}\frac12&\frac12\end{bmatrix}$。直接相乘得 $LA=[1]=I_1$。反向乘积为

$$AL=\begin{bmatrix}\frac12&\frac12\\[2pt]\frac12&\frac12\end{bmatrix}\neq I_2.$$

因此 L 是左逆，不能是右逆。这个例子也显示 A 把一个数映成沿 $(1,1)^{\mathsf{T}}$ 方向的向量，而左逆把该方向上的输出还原成输入。 ♦

这个 A 的左逆并不唯一：所有 $\begin{bmatrix}t&1-t\end{bmatrix}$ 都满足 $LA=1$。若输入输出 $y=(s,s)+(d,-d)$，这些左逆给出 $s+(2t-1)d$；只有 $t=1/2$ 的标准公式会把垂直于列空间的残差 $(d,-d)$ 完全忽略。这说明标准左逆还带有一个自然的正交投影选择。

想一想　若 A 是满列秩的 $m\times n$ 矩阵且 $m>n$，它能有右逆吗?

答　不能。右逆要求 $\operatorname{rank}(A)=m$，但 $\operatorname{rank}(A)\le n<m$。左逆可以存在，正如上例中 $m=2, n=1$ 的情形。

31.2　左逆、最小二乘与列空间投影

对满列秩矩阵，方程 $Ax=b$ 至多有一个解；但 b 可能不在 A 的列空间里。此时最接近 b 的可达向量由正交投影给出，而左逆公式正好能算出对应的系数。

定理 31.3（满列秩下的最小二乘公式）　设 $A\in\mathbb{R}^{m\times n}$ 满列秩。对任意 $b\in\mathbb{R}^m$，唯一使 $\|Ax-b\|$ 最小的向量为

$$\widehat{x}=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}b=A_{\text{left}}b,\qquad A_{\text{left}}=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}.$$

残差 $e=b-A\widehat{x}$ 满足 $A^{\mathsf{T}}e=0$，而 $P=AA_{\text{left}}$ 是到 $\operatorname{Col}(A)$ 的正交投影矩阵。

思路　最小化 $\|Ax-b\|^2$ 时，Ax 必须是 $\operatorname{Col}(A)$ 中离 b 最近的向量。正交投影的残差垂直于每一列，故 $A^{\mathsf{T}}(b-Ax)=0$；满列秩保证 $A^{\mathsf{T}}A$ 可逆，从而解唯一。

证明　令 $\widehat{x}=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}b$，则 $A^{\mathsf{T}}(b-A\widehat{x})=0$。所以 $e=b-A\widehat{x}$ 垂直于 $\operatorname{Col}(A)$。

对任意 $x=\widehat{x}+h$，有 $Ax-b=Ah-e$，且 $Ah\in\operatorname{Col}(A)$，因此 $Ah\perp e$。于是

$$\|Ax-b\|^2=\|Ah\|^2+\|e\|^2\ge\|e\|^2.$$

等号要求 $Ah=0$；因 A 的列独立，$h=0$，所以极小解唯一。

矩阵 $P=A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$ 满足

$$P^{\mathsf{T}}=P,\qquad P^2=A(A^{\mathsf{T}}A)^{-1}(A^{\mathsf{T}}A)(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}=P.$$

其输出总在 $\operatorname{Col}(A)$ 中，且对 $y=Ax$ 有 $Py=y$。因此 P 是到 $\operatorname{Col}(A)$ 的正交投影。 ■

命题 31.4（满列秩矩阵的所有左逆）　若 A 满列秩，令 $L_0=(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$、$P=AL_0$。所有左逆恰为

$$L=L_0+K(I-P),\qquad K\in\mathbb{R}^{n\times m}.$$

其中 L_0 把 $\operatorname{Col}(A)^{\perp}$ 中的分量映为零；其他左逆可以把这些分量送到任意方向，但在 $\operatorname{Col}(A)$ 上都必须与 L_0 相同。

证明　因为 $(I-P)A=0$，每个上述 L 都满足 $LA=L_0A=I_n$。反过来，若 L 是左逆，令 $D=L-L_0$，则 $DA=0$，并且 $DP=DAL_0=0$，所以 $D=D(I-P)$。取 $K=D$ 即得到所列形式。 ■

法方程不仅给出系数，也显出拟合误差如何分配。因为 $p=A\widehat{x}\in\operatorname{Col}(A)$、$e\perp\operatorname{Col}(A)$，有

$$\|b\|^2=\|p+e\|^2=\|p\|^2+\|e\|^2.$$

也可从平方误差的变化看出同一个正交条件。若把当前系数改成 $x+h$，则

$$\|A(x+h)-b\|^2=\|Ax-b\|^2+2h^{\mathsf T}A^{\mathsf T}(Ax-b)+h^{\mathsf T}A^{\mathsf T}Ah.$$

要使 x 成为极小点，任意方向 h 上的一次变化都必须为零，于是 $A^{\mathsf T}(Ax-b)=0$，即残差垂直于每一列。满列秩时 $A^{\mathsf T}A$ 正定，二次项对非零 h 为正，故这个驻点也是唯一极小点。因此每个最小二乘问题都可以分成可拟合的部分 p 和无法由列组合消去的部分 e。若 $b\in\mathrm{Col}(A)$，则 $e=0$，最小二乘公式退化为精确解；若 b 不在列空间，残差则记录了数据中无法拟合的部分。

■ **例 31.2**　用 $A=\begin{bmatrix}1\\1\end{bmatrix}$ 拟合 $b=\begin{bmatrix}2\\0\end{bmatrix}$。求最小二乘系数、拟合向量和残差。

解　套用法方程：$A^{\mathsf T}A=[2]$，$A^{\mathsf T}b=[2]$，所以 $\widehat{x}=1$。拟合向量为 $p=A\widehat{x}=(1,1)^{\mathsf T}$，残差为 $e=b-p=(1,-1)^{\mathsf T}$。

检查正交条件：$A^{\mathsf T}e=1-1=0$。所以 p 是直线 $\mathrm{Col}(A)=\mathrm{span}\{(1,1)^{\mathsf T}\}$ 上离 b 最近的点。图 31.1 把这次投影和垂直残差画在同一坐标系里。 ♦

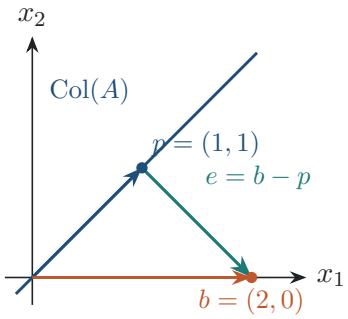


图 31.1　投影点 $p=(1,1)$ 落在列空间直线上，残差 $e=(1,-1)$ 与该直线垂直。

从图中看，橙色的 b 不在蓝色直线上，因而方程 $Ax=b$ 无解。蓝点 p 是唯一最近点；青色残差与蓝线垂直，这正是法方程 $A^{\mathsf T}(b-p)=0$ 的几何意思。

易错提醒　满列秩时，不能把 $A^{\mathsf T}$ 直接当成左逆。对 $A=(1,1)^{\mathsf T}$，$A^{\mathsf T}A=[2]$，所以 $A^{\mathsf T}A\neq I_1$；正确的左逆要乘上 $(A^{\mathsf T}A)^{-1}$。

31.3　右逆与行空间投影

满行秩意味着 $A^{\mathsf T}$ 的列独立，也就是 $N(A^{\mathsf T})=\{0\}$。这等价于 A 的左零空间只有零向量；消元时不会出现矛盾行，因此 $Ax=b$ 对每个 $b\in\mathbb{R}^m$ 都有解。若 $n>m$，未知数比主元多，解通常不止一个。

命题 31.5（满行秩下的右逆）　若 $A \in \mathbb{R}^{m\times n}$ 满行秩，则

$$A_{\text{right}} = A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}$$

是一个右逆。矩阵 $Q = A_{\text{right}}A$ 是到 $\mathrm{Row}(A)$ 的正交投影；对任意 b，$A_{\text{right}}b$ 是方程 $Ax = b$ 的一个解。

证明　满行秩使 AA^{T} 可逆。于是 $AA_{\text{right}} = AA^{\mathsf{T}}(AA^{\mathsf{T}})^{-1} = I_m$，所以 $A_{\text{right}}b$ 确实满足 $Ax = b$。

又 $Q = A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}A$ 对称，且

$$Q^2 = A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}(AA^{\mathsf{T}})(AA^{\mathsf{T}})^{-1}A = Q.$$

它的列属于 $\mathrm{Row}(A)$，并且对 $x = A^{\mathsf{T}}y$ 有 $Qx = x$，所以 Q 是到行空间的正交投影。 ■

命题 31.6（右逆给出的最小范数解）　设 A 满行秩，$R = A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}$。对每个 b，方程 $Ax = b$ 的全部解为

$$x = Rb + z, \qquad z \in N(A).$$

其中 Rb 是唯一的最小范数解；若 $n > m$，解集有 $n - m$ 个自由参数。

证明　若 $Ax = b$，则 $A(x - Rb) = b - b = 0$，所以 $x - Rb \in N(A)$；反过来，$ARb = b$ 且 $Az = 0$，故上述每个向量都是解。由于 $Rb \in \mathrm{Row}(A) = N(A)^{\perp}$，有 $\|Rb + z\|^2 = \|Rb\|^2 + \|z\|^2$，最小值且仅在 $z = 0$ 取得。满行秩使 $\dim N(A) = n - m$。 ■

若 $n > m$，右逆也通常不唯一。令 $R_0 = A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}$、$Q = R_0A$，则所有右逆可写为

$$R = R_0 + (I - Q)K, \qquad K \in \mathbb{R}^{n\times m}.$$

因为 $A(I - Q) = 0$，新增项只落在零空间中，不改变 $AR = I_m$。标准右逆取 $K = 0$，其每个输出都在行空间内，因而与零空间垂直，是所有右逆中给出最小范数解的那个。

右逆公式给出一组解，但没有消掉零空间里的自由分量。下面用一个方程把「总有解」和「解不唯一」分开看。

■ **例 31.3**　对 $A = \begin{bmatrix}1 & 2\end{bmatrix}$ 和 $b = 3$，求标准右逆给出的解，并写出全部解。

解　先算 $AA^{\mathsf{T}} = [5]$，所以 $A_{\text{right}} = \frac{1}{5}\begin{bmatrix}1\\2\end{bmatrix}$。因此一组解为 $x_0 = A_{\text{right}}b = (\frac{3}{5}, \frac{6}{5})^{\mathsf{T}}$。

$$Ax_0 = \frac{3}{5} + 2\cdot\frac{6}{5} = 3.$$

零空间由 $x_1 + 2x_2 = 0$ 给出，方向向量是 $(-2, 1)^{\mathsf{T}}$。所以全部解为

$$x = x_0 + t\begin{bmatrix}-2\\1\end{bmatrix}, \qquad t \in \mathbb{R}.$$

此外 $x_0^{\mathsf T}(-2,1)=0$，故右逆选出的解垂直于零空间，是所有解中范数最小的一组。这说明右逆解决了「有解」，但不保证「唯一解」。 ♦

乘积顺序在这里很重要：$AA_{\text{right}}=I_m$，而 $A_{\text{right}}A$ 一般不是 I_n。它只把 $\mathbb{R}^n$ 中的向量投影到行空间；被投影掉的部分正是 $N(A)$。

想一想　上例的 A 有右逆，为什么没有左逆？
答　A 有 2 列却只有秩 1，所以 $N(A)$ 含有非零向量 $(-2,1)^{\mathsf T}$。左逆会把 $A(-2,1)^{\mathsf T}=0$ 还原成非零向量，与 $LA=I_2$ 矛盾。

满行秩也可以出现在有两条独立方程的系统中。此时右逆的每一列给出一个单位右端项的特解，而右逆乘回去的投影保留行空间分量。

■ **例 31.4**　令 $A=\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}$。求标准右逆及其行空间投影，并用右逆求解 $Ax=(1,2)^{\mathsf T}$。

解　有 $AA^{\mathsf T}=\begin{bmatrix}2&1\\1&2\end{bmatrix}$，所以

$$A_{\text{right}}=A^{\mathsf T}(AA^{\mathsf T})^{-1}=\begin{bmatrix}\frac{2}{3}&-\frac{1}{3}\\-\frac{1}{3}&\frac{2}{3}\\\frac{1}{3}&\frac{1}{3}\end{bmatrix}.$$

直接相乘得 $AA_{\text{right}}=I_2$。而

$$Q=A_{\text{right}}A=\begin{bmatrix}\frac{2}{3}&-\frac{1}{3}&\frac{1}{3}\\-\frac{1}{3}&\frac{2}{3}&\frac{1}{3}\\\frac{1}{3}&\frac{1}{3}&\frac{2}{3}\end{bmatrix}$$

是到 $\mathrm{Row}(A)$ 的正交投影。对 $b=(1,2)^{\mathsf T}$，右逆给出

$$x_0=A_{\text{right}}b=(0,1,1)^{\mathsf T},\qquad Ax_0=(1,2)^{\mathsf T}.$$

其他解还可加上 $N(A)=\mathrm{span}\{(-1,-1,1)^{\mathsf T}\}$ 中的向量。 ♦

31.4　伪逆：把行空间送回去

若 A 既不满列秩也不满行秩，零空间和左零空间都非零，双侧逆无从谈起。仍有一块没有损失信息：A 把 $\mathrm{Row}(A)$ 一一对应到 $\mathrm{Col}(A)$。伪逆就是这段对应的逆映射。

命题 31.7（行空间到列空间的双射）　对任意 $A\in\mathbb{R}^{m\times n}$，限制映射

$$A\colon\ \mathrm{Row}(A)\longrightarrow\mathrm{Col}(A)$$

是一一对应。两个空间的维数都等于 $r=\mathrm{rank}(A)$。

证明　若 $x \in \mathrm{Row}(A)$ 且 $Ax = 0$，则 $x \in N(A)$。而 $\mathrm{Row}(A) = N(A)^{\perp}$，所以 x 同时垂直于自身，只能是 0。因此限制映射是单射。

它的像包含在 $\mathrm{Col}(A)$ 中；又像的维数等于定义域维数 r，而 $\dim \mathrm{Col}(A) = r$，所以像就是整个列空间。故该限制映射也是满射。 ■

定义 31.8（伪逆）　设 $A \in \mathbb{R}^{m\times n}$。$A$ 在行空间到列空间之间的逆映射，扩展为在 $N(A^{\mathsf{T}})$ 上取零的线性映射，称为 A 的伪逆（pseudoinverse），记作 A^{+}。于是 $A^{+}Ax = x$ 对每个 $x \in \mathrm{Row}(A)$ 成立，且 $A^{+}y$ 对每个 $y \in \mathrm{Col}(A)$ 是唯一的行空间原像。

定义中的「扩展为零」需要说明：任意 $x \in \mathbb{R}^n$ 都能唯一分解为 $x = x_r + x_n$，其中 $x_r \in \mathrm{Row}(A)$、$x_n \in N(A)$；而任意 $y \in \mathbb{R}^m$ 都能唯一分解为 $y = y_c + y_l$，其中 $y_c \in \mathrm{Col}(A)$、$y_l \in N(A^{\mathsf{T}})$。伪逆先去掉左零空间分量，再还原剩下的列空间分量。

定理 31.9（用奇异值分解构造伪逆）　设 $A = U\Sigma V^{\mathsf{T}}$，其中 $U \in \mathbb{R}^{m\times m}$、$V \in \mathbb{R}^{n\times n}$ 正交，$\Sigma \in \mathbb{R}^{m\times n}$ 的前 $r = \mathrm{rank}(A)$ 个对角元为正奇异值 $\sigma_1, \dots, \sigma_r$。令 $\Sigma^{+} \in \mathbb{R}^{n\times m}$ 的对应对角元为 $1/\sigma_1, \dots, 1/\sigma_r$，其余元素为零，则

$$A^{+} = V\Sigma^{+}U^{\mathsf{T}}.$$

并且 AA^{+} 是到 $\mathrm{Col}(A)$ 的正交投影，$A^{+}A$ 是到 $\mathrm{Row}(A)$ 的正交投影；$N(A^{+}) = N(A^{\mathsf{T}})$。

思路　在 SVD 坐标中，U 和 V 只是换成两组正交单位基。真正需要处理的是对角矩阵 Σ：正奇异值方向可以逐一取倒数，零奇异值方向则只能送到零。最后将两个坐标变换乘回去。

证明　正交矩阵只需转置即可求逆，因此只需考察对角矩阵 Σ。乘积 $\Sigma\Sigma^{+}$ 在 $\mathbb{R}^m$ 中把前 r 个奇异向量方向保留、其余方向变为零；$\Sigma^{+}\Sigma$ 在 $\mathbb{R}^n$ 中保留前 r 个右奇异向量方向。于是

$$AA^{+} = U(\Sigma\Sigma^{+})U^{\mathsf{T}}, \qquad A^{+}A = V(\Sigma^{+}\Sigma)V^{\mathsf{T}},$$

分别是列空间和行空间的正交投影。又 $A^{+} = V\Sigma^{+}U^{\mathsf{T}}$ 在 U 坐标中恰好消去左零空间的分量，所以 $N(A^{+}) = N(A^{\mathsf{T}})$。在 $\mathrm{Row}(A)$ 上，$A^{+}A$ 是恒等映射；在 $N(A)$ 上它为零，正好符合定义中的伪逆。 ■

命题 31.10（伪逆在满秩情形下的公式）　若 A 满列秩，则 $A^{+} = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$；若 A 满行秩，则 $A^{+} = A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}$。方阵可逆时，两式都化为 A^{-1}。

证明　满列秩时，$\mathrm{Row}(A) = \mathbb{R}^n$，$A^{+}$ 在 $\mathrm{Col}(A)$ 上把 Ax 还原为 x，并在 $\mathrm{Col}(A)^{\perp}$ 上取零。公式 $(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$ 也有相同作用：对 Ax 得到 x，对与列空间垂直的向量 e，因 $A^{\mathsf{T}}e = 0$ 而得到零。因此两者相同。满行秩时，$\mathrm{Col}(A) = \mathbb{R}^m$；$A^{+}b$ 是行空间中的唯一解，而 $A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}b$ 也是行空间中的解，故二者相同。方阵可逆时，左、右逆都是通常的逆矩阵。 ■

秩为零时也能直接读出边界：若 $A = 0$ 且 $m, n > 0$，它没有左逆或右逆，SVD 中没有正奇

异值，所以 $A^{+}=0$。两个投影 AA^{+}、$A^{+}A$ 也都是零矩阵。对任意 b，拟合向量只能是 $p=0$，残差为 $e=b$；所有 $x\in\mathbb{R}^{n}$ 都是最小二乘解，而最小范数解是 $x_{*}=0$。

这个边界也提醒我们：伪逆存在，不代表原方程 $Ax=b$ 对每个 b 都有解。例如零矩阵遇到 $b\neq 0$ 时，只能把 b 投影到零列空间，留下完整残差。一般情况下，AA^{+} 先确定唯一的最佳拟合，$A^{+}A$ 再标出输入中可辨识的部分。

例 31.5　秩一矩阵的伪逆　求 $A=\begin{bmatrix}4&3\\8&6\end{bmatrix}$ 的伪逆，并写出两个空间投影。

解　一组奇异向量为 $u_1=(1,2)^{\mathsf{T}}/\sqrt{5}$、$u_2=(2,-1)^{\mathsf{T}}/\sqrt{5}$，$v_1=(4,3)^{\mathsf{T}}/5$、$v_2=(3,-4)^{\mathsf{T}}/5$；唯一非零奇异值为 $\sigma_1=5\sqrt{5}$。因此

$$U=\frac{1}{\sqrt{5}}\begin{bmatrix}1&2\\2&-1\end{bmatrix},\quad \Sigma=\begin{bmatrix}5\sqrt{5}&0\\0&0\end{bmatrix},\quad V=\frac{1}{5}\begin{bmatrix}4&3\\3&-4\end{bmatrix},\quad \Sigma^{+}=\begin{bmatrix}\frac{1}{5\sqrt{5}}&0\\0&0\end{bmatrix}.$$

于是

$$A^{+}=V\Sigma^{+}U^{\mathsf{T}}=v_1\sigma_1^{-1}u_1^{\mathsf{T}}=\frac{1}{125}\begin{bmatrix}4&8\\3&6\end{bmatrix}.$$

相乘得到

$$AA^{+}=\begin{bmatrix}\frac{1}{5}&\frac{2}{5}\\\frac{2}{5}&\frac{4}{5}\end{bmatrix},\qquad A^{+}A=\frac{1}{25}\begin{bmatrix}16&12\\12&9\end{bmatrix}.$$

前者投影到 $\mathrm{Col}(A)=\mathrm{span}\{(1,2)^{\mathsf{T}}\}$，后者投影到 $\mathrm{Row}(A)=\mathrm{span}\{(4,3)^{\mathsf{T}}\}$。二者都对称且幂等；它们之所以不是整个平面的单位矩阵，是因为 A 的秩只有 1。 ♦

图 31.2 把这个秩一矩阵的四个基本子空间放在同一张图里：A 把行空间方向送到列空间方向，却把零空间方向压为零；左零空间与列空间垂直，不能由 Ax 产生。

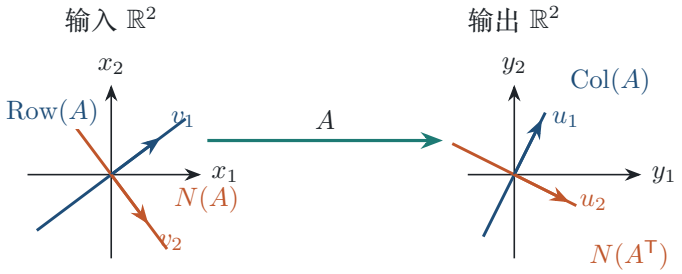


图 31.2　伪逆只在 $\mathrm{Row}(A)$ 与 $\mathrm{Col}(A)$ 之间反转 A；$N(A)$ 被消去，$N(A^{\mathsf{T}})$ 被忽略。

源平面里的 v_1 是可逆方向，满足 $Av_1=\sigma_1u_1$；v_2 是 A 无法分辨的方向，满足 $Av_2=0$。伪逆先把 $N(A^{\mathsf{T}})$ 分量投掉，再沿 u_1 方向除以奇异值并返回 v_1 方向。因此两个乘积分别恢复输出端的可达部分和输入端的可辨识部分。

想一想　若把 Σ^+ 中某个原本为零的位置改成非零数，仍能使 $\Sigma\Sigma^+$ 和 $\Sigma^+\Sigma$ 保留相同的投影作用吗？

答　不能。非零奇异值以外的零位置若被改成非零，相关乘积会多出非零方向，便不再是投到列空间或行空间的投影。伪逆只倒转可逆的奇异方向。

31.5　一般最小二乘与最小范数解

满列秩公式 $(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$ 需要列独立。若列相关，$A^{\mathsf{T}}A$ 不可逆，普通法方程无法选出唯一系数；但最接近 b 的列空间向量仍然唯一。伪逆给出其中范数最小的系数。

定理 31.11（伪逆给出最小二乘的最小范数解）　对任意 $A\in\mathbb{R}^{m\times n}$ 和 $b\in\mathbb{R}^m$，$x_*=A^+b$ 是使 $\|Ax-b\|$ 最小的最小范数解。拟合向量 $p=AA^+b$ 是 b 在 $\mathrm{Col}(A)$ 上的正交投影，残差 $e=b-p$ 属于 $N(A^{\mathsf{T}})$。所有最小二乘解为

$$x=A^+b+z,\qquad z\in N(A).$$

思路　先把 b 分成列空间投影 p 与左零空间残差 e；任何 Ax 都在列空间，因此 e 是不可避免的误差。等号条件给出全部最小二乘解，再利用行空间与零空间互相垂直，选出其中范数最小的一组。

证明　由定理 31.9，$p=AA^+b\in\mathrm{Col}(A)$，而 $e=(I-AA^+)b\in N(A^{\mathsf{T}})=\mathrm{Col}(A)^\perp$。对任意 x，$Ax-p\in\mathrm{Col}(A)$，所以

$$\|Ax-b\|^2=\|(Ax-p)-e\|^2=\|Ax-p\|^2+\|e\|^2\geq\|e\|^2.$$

等号当且仅当 $Ax=p$。因为 $A(A^+b)=p$，所有达到最小值的向量恰为 A^+b+z，其中 $z\in N(A)$。又 $A^+b\in\mathrm{Row}(A)=N(A)^\perp$，故

$$\|A^+b+z\|^2=\|A^+b\|^2+\|z\|^2.$$

范数最小且唯一的解对应 $z=0$。 ■

■ **例 31.6　秩亏时拟合与选解**　对 $A=\begin{bmatrix}4&3\\8&6\end{bmatrix}$ 和 $b=(1,3)^{\mathsf{T}}$，求所有最小二乘解，并指出最小范数解。

解　由前述 SVD 算例的伪逆 $A^+=\frac{1}{125}\begin{bmatrix}4&8\\3&6\end{bmatrix}$，得

$$x_*=A^+b=\begin{bmatrix}\frac{28}{125}\\[2pt]\frac{21}{125}\end{bmatrix},\quad p=Ax_*=\begin{bmatrix}\frac{7}{5}\\[2pt]\frac{14}{5}\end{bmatrix},\quad e=b-p=\begin{bmatrix}-\frac{2}{5}\\[2pt]\frac{1}{5}\end{bmatrix}.$$

可直接检查 $A^{\mathsf{T}}e=0$，所以 p 是 b 在列空间 $\mathrm{span}\{(1,2)^{\mathsf{T}}\}$ 上的投影。

零空间为 $N(A)=\mathrm{span}\{(-3,4)^{\mathsf{T}}\}$。

全部最小二乘解为

$$x = x_* + t\begin{bmatrix} -3 \\ 4 \end{bmatrix}, \qquad t \in \mathbb{R}.$$

由于 $x_* \cdot (-3, 4) = 0$，其中 $t = 0$ 的解范数最小。列相关使系数不唯一，但投影向量 p 唯一；伪逆选出垂直于零空间的那组系数。 ♦

■ **例 31.7　用最小二乘拟合一条直线**　在时刻 $t = 0, 1, 2$ 观测到 $y = 1, 2, 2$。用直线 $y = c_0 + c_1 t$ 拟合三次观测，求截距、斜率和残差平方和。

解　把截距和斜率作为未知数，设计矩阵与观测列为

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad y = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

X 两列独立，故法方程有唯一解。计算得

$$X^{\mathsf{T}}X = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}, \quad X^{\mathsf{T}}y = \begin{bmatrix} 5 \\ 6 \end{bmatrix}, \quad \widehat{c} = (X^{\mathsf{T}}X)^{-1}X^{\mathsf{T}}y = \begin{bmatrix} \frac{7}{6} \\ \frac{1}{2} \end{bmatrix}.$$

拟合直线为 $y = \frac{7}{6} + \frac{1}{2}t$，三个拟合值为 $(\frac{7}{6}, \frac{5}{3}, \frac{13}{6})^{\mathsf{T}}$，残差为 $(-\frac{1}{6}, \frac{1}{3}, -\frac{1}{6})^{\mathsf{T}}$。残差与常数列、时间列都正交：$X^{\mathsf{T}}e = 0$；残差平方和为 $\|e\|^2 = \frac{1}{6}$。这说明最小二乘直线并不必经过每个观测点，而是让残差向量垂直于由模型列张成的空间。 ♦

在统计回归中，A 的列可代表观测变量，b 是观测结果。满列秩时，法方程给出唯一系数；变量列相关或变量多于有效观测方向时，$A^{+}b$ 仍给出最小二乘拟合，并在所有等价系数中选取最小范数的一组。

本讲小结

- A 有左逆当且仅当满列秩；有右逆当且仅当满行秩；双侧逆要求方阵且满秩。
- 满列秩时 $A_{\text{left}} = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$，且 AA_{left} 是到 $\mathrm{Col}(A)$ 的正交投影。
- 满行秩时 $A_{\text{right}} = A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}$，且 $A_{\text{right}}A$ 是到 $\mathrm{Row}(A)$ 的正交投影；$Ax = b$ 对每个 b 都有解，$A_{\text{right}}b$ 是其中的最小范数解。若 $n > m$，解集有 $n - m$ 个自由度，这些自由方向属于 $N(A)$。
- A 将 $\mathrm{Row}(A)$ 一一映到 $\mathrm{Col}(A)$；伪逆在这两个空间上反转该映射，并在 $N(A^{\mathsf{T}})$ 上取零。
- 若 $A = U\Sigma V^{\mathsf{T}}$，则 $A^{+} = V\Sigma^{+}U^{\mathsf{T}}$；$AA^{+}$、$A^{+}A$ 分别投影到列空间、行空间。
- 对任意 b，$AA^{+}b$ 是最小二乘拟合，$A^{+}b$ 是最小范数系数；其余最小二乘解只差一个零空间向量。

练习

31.1 ★　令 $A = (1, 1, 0)^{\mathsf{T}}$，求一个左逆 L 和投影 AL，并说明 LA 与 AL 为什么不同。

31.2 ★ 判断真假并说明理由：(a) A 有左逆当且仅当其行向量线性无关；(b) 满行秩矩阵 A 对每个 b 都有解 $Ax=b$；(c) 若 A 是 3×2 且满列秩，则有右逆。

31.3 ★★ 原习题 32.1：对 $A=\begin{bmatrix}1&0&1\\0&1&0\end{bmatrix}$ 求一个右逆，并检查乘积。

31.4 ★★ 原习题 32.2：对 $A=\begin{bmatrix}4&3\\8&6\end{bmatrix}$ 判断是否存在左逆、右逆、伪逆；若存在，求伪逆并验证它在行空间和列空间上的作用。

31.5 ★★ 对同一个矩阵 A 与 $b=(1,0)^{\mathsf T}$ 求最小范数最小二乘解、拟合向量和残差，并检查法方程。

31.6 ★★ 对 $A=(1,2)^{\mathsf T}$、$b=(1,0)^{\mathsf T}$ 求最小二乘系数、拟合向量及残差，验证残差垂直于列空间。

31.7 ★★ 求 $A=\begin{bmatrix}1&1\\2&2\end{bmatrix}$ 的伪逆及 AA^{+}、$A^{+}A$，并说明两个投影各自的像空间。

31.8 ★★★ 证明：AA^{+} 和 $A^{+}A$ 对称且幂等，并分别固定 $\mathrm{Col}(A)$、$\mathrm{Row}(A)$ 中的每个向量。

31.9 ★★ 对 $A=\begin{bmatrix}1&2\end{bmatrix}$、$b=3$，求范数最小的解，说明它与右逆给出的解有何关系。

练习答案与提示

31.1 $A^{\mathsf T}A=[2]$，所以 $L=\frac12\begin{bmatrix}1&1&0\end{bmatrix}$。$AL=\begin{bmatrix}\frac12&\frac12&0\\\frac12&\frac12&0\\0&0&0\end{bmatrix}$，投影到 $\mathrm{span}\{(1,1,0)^{\mathsf T}\}$；$LA=I_1$ 而 AL 是 3×3 投影。

31.2 (a) 假。左逆等价于列向量线性无关；反例可取 $\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}$，其行相关但满列秩。(b) 真，满行秩等价于 $\mathrm{Col}(A)=\mathbb{R}^m$。(c) 假，$3\times 2$ 满列秩时秩为 $2<3$，不可能满行秩。

31.3 可取 $R=\begin{bmatrix}\frac12&0\\0&1\\\frac12&0\end{bmatrix}$。相乘得 $AR=I_2$。

31.4 $\mathrm{rank}(A)=1<2$，所以没有左逆或右逆。其 SVD 可取

$$U=\frac{1}{\sqrt5}\begin{bmatrix}1&2\\2&-1\end{bmatrix},\quad \Sigma=\begin{bmatrix}5\sqrt5&0\\0&0\end{bmatrix},\quad V=\frac15\begin{bmatrix}4&3\\3&-4\end{bmatrix}.$$

因而 $A^{+}=\frac{1}{125}\begin{bmatrix}4&8\\3&6\end{bmatrix}$。$AA^{+}=\frac15\begin{bmatrix}1&2\\2&4\end{bmatrix}$ 投影到 $\mathrm{Col}(A)=\mathrm{span}\{(1,2)^{\mathsf T}\}$；$A^{+}A=\frac{1}{25}\begin{bmatrix}16&12\\12&9\end{bmatrix}$ 投影到 $\mathrm{Row}(A)=\mathrm{span}\{(4,3)^{\mathsf T}\}$。两者均对称且幂等。具体地，令 $v_1=(4,3)^{\mathsf T}/5$、$u_1=(1,2)^{\mathsf T}/\sqrt5$、$u_2=(2,-1)^{\mathsf T}/\sqrt5$，则 $Av_1=5\sqrt5u_1$、$A^{+}u_1=(5\sqrt5)^{-1}v_1$，而 $A^{+}u_2=0$。伪逆在可达方向上反转 A，并消去左零空间方向。

31.5 由 A^{+} 得 $x_*=A^{+}b=(\frac{4}{125},\frac{3}{125})^{\mathsf T}$，$p=Ax_*=(\frac15,\frac25)^{\mathsf T}$，$e=b-p=(\frac45,-\frac25)^{\mathsf T}$。直接计算 $A^{\mathsf T}e=(0,0)^{\mathsf T}$；其他最小二乘解为 $x_*+t(-3,4)^{\mathsf T}$，其中 x_* 与零空间垂直，故范数最小。

31.6 $A^{\mathsf T}A=[5]$、$A^{\mathsf T}b=[1]$，所以 $\widehat{x}=1/5$；$p=(1/5,2/5)^{\mathsf T}$、$e=(4/5,-2/5)^{\mathsf T}$，且 $A^{\mathsf T}e=0$。

31.7 $A^{+}=\frac{1}{10}\begin{bmatrix}1&2\\1&2\end{bmatrix}$，$AA^{+}=\frac{1}{5}\begin{bmatrix}1&2\\2&4\end{bmatrix}$ 投影到 $\operatorname{span}\{(1,2)^{\mathsf{T}}\}$，$A^{+}A=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}$ 投影到 $\operatorname{span}\{(1,1)^{\mathsf{T}}\}$。

31.8 由 SVD 定理，$AA^{+}=U(\Sigma\Sigma^{+})U^{\mathsf{T}}$、$A^{+}A=V(\Sigma^{+}\Sigma)V^{\mathsf{T}}$。括号内是对角矩阵，只有前 r 个对角元为 1，所以两者对称、幂等；前者固定列空间，后者固定行空间。

31.9 方程为 $x_1+2x_2=3$。右逆给出的解是 $(3/5,6/5)^{\mathsf{T}}$；零空间方向为 $(-2,1)^{\mathsf{T}}$，且该右逆解已与零空间垂直，因此它就是范数最小解。所有解为 $(3/5,6/5)^{\mathsf{T}}+t(-2,1)^{\mathsf{T}}$。

复习课 3

第三单元与全课程复习

Unit 3 and Cumulative Review

对应材料：依据第 24–31 讲讲义、课时摘要与习题；参考 2011 Unit 3 Exam（第 1–3 题）及解答、Final Review，以及 2011 Final Exam（第 1、8 题）及解答。例题与自测题在题后标注来源或改编范围。

本讲目标

(1) 能把实对称谱定理、复数 Fourier 变换、Jordan 形与 SVD 放进同一张知识图。

(2) 能在换基、最小二乘、低秩近似和伪逆题中先辨认空间与秩条件。

(3) 能用 Unit 3 原题练习谱分解、矩阵函数和正定性，并说明容易漏掉的条件。

(4) 能把三单元主线连回 $Ax=b$、四个基本子空间、正交投影与特征值。

R3.1 知识地图

本单元从对称矩阵的正交特征基，走到任意矩阵的奇异向量，再用基变换简化表示、用伪逆处理秩亏。图 R3.1 标出几条依赖；遮住表格中间一列，试着说出各概念的定义条件。

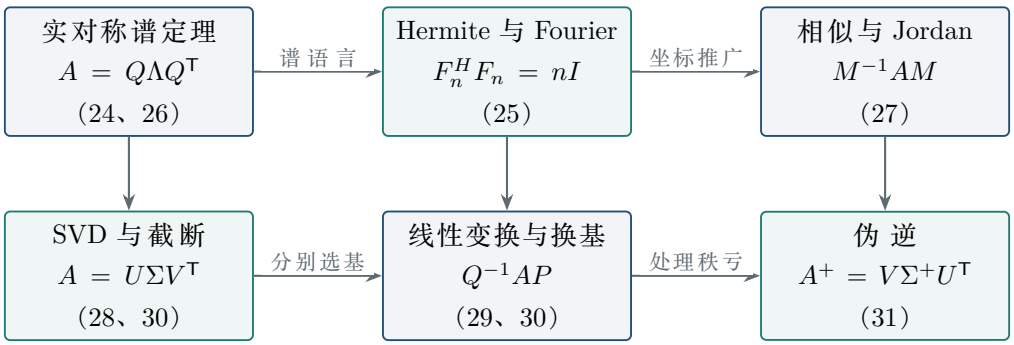


图 R3.1　谱分解处理对称算子，SVD 处理一般矩阵，伪逆在非零奇异方向上反转矩阵。

概念	定义要点	典型用法
谱定理	实对称矩阵可写成 $Q\Lambda Q^{\mathsf{T}}$，Q 正交，Λ 实对角	读出特征方向、计算 $p(A)$ 与二次型
Hermite / Fourier	$A=A^H$；$F_n^H F_n = nI$，复内积使用共轭	复矩阵的实谱、周期数据的频率坐标
相似 / Jordan	$B=M^{-1}AM$；Jordan 块记录特征向量链的长度	判断同一线性变换的坐标表示，区分重根情形
SVD	$A=U\Sigma V^{\mathsf{T}}$，$\sigma_i\ge 0$，正值个数是秩	四基本子空间、范数、最佳低秩近似
换基	输入基矩阵 P、输出基矩阵 Q 时为 $Q^{-1}AP$	让同一个线性变换在合适坐标中变简单
伪逆	$A^+=V\Sigma^+U^{\mathsf{T}}$，只对正奇异值取倒数	秩亏方程的最小二乘、最小范数解

R3.2　核心结论回顾

R3.2.1　对称、Hermite 与正定（第 24–26 讲）

定理（实对称谱定理）　对实矩阵 A，以下两件事等价：$A=A^{\mathsf{T}}$；存在正交矩阵 Q 和实对角矩阵 Λ，使

$$A = Q\Lambda Q^{\mathsf{T}}.$$

其中 Λ 的对角元是 A 的特征值，Q 的列是标准正交特征向量。若特征值重复，仍可在对应特征空间内正交化。

关键　不同特征值的特征向量先由对称性得到正交；重复特征值的特征空间是普通子空间，再用 Gram–Schmidt 取标准正交基。完整证明见第 24 讲。

用法　看到「实对称」就优先使用正交谱分解，而不是任意的 $S\Lambda S^{-1}$。随后 $p(A)=Qp(\Lambda)Q^{\mathsf{T}}$，例如逆矩阵把每个非零特征值换成倒数。

定理（Hermite 谱与 Fourier 正交性）　复矩阵的 Hermite 条件是 $A=A^H$（H 表示共轭转置）；Hermite 矩阵的特征值为实数，并可由酉矩阵正交对角化。令

$$F_n = (\omega^{jk})_{0\le j,k<n}, \qquad \omega = e^{2\pi i/n},$$

则 $F_n^H F_n = nI$，所以 $n^{-1/2}F_n$ 是酉矩阵。

关键　若 $Av=\lambda v$，比较 v^HAv 与 $(Av)^Hv$，Hermite 条件迫使 $\lambda=\bar{\lambda}$。Fourier 内积则化为等比数列求和：非零频率差的和为零。

用法　复题先把转置换成共轭转置；Fourier 坐标要记住归一化因子。未归一化的 F_n 满足 $F_n^{-1}=n^{-1}F_n^H$，不能直接当作酉矩阵。

定理（正定的三个等价判据）　实对称矩阵 A 正定，当且仅当下列任一条件成立：所有特征值为正；$x^{\mathsf{T}}Ax>0$ 对每个 $x\neq 0$；所有顺序主子式 $D_k=\det A_{1:k,1:k}$ 都为正。半正定只要求二次型非负，允许有非零零方向。

关键　谱分解把二次型写成 $x^{\mathsf{T}}Ax=(Q^{\mathsf{T}}x)^{\mathsf{T}}\Lambda(Q^{\mathsf{T}}x)$；顺序主子式判据则把正定性逐步传给对称消元的主元。完整证明见第 24、26 讲。

用法　判断小矩阵时先看对称性，再选最短判据。只看对角元或只看总行列式都不够；例如 $\begin{bmatrix}2&6\\6&7\end{bmatrix}$ 对角元为正但不是正定。

R3.2.2 相似、特征基与 Jordan（第 27 讲）

定义（相似与换坐标）　方阵 A,B 相似，是指存在可逆 M 使 $B=M^{-1}AM$。这是同一线性变换在同一输入、输出基变化下的矩阵关系。

关键　先写 $AS=S\Lambda$，再左乘 S^{-1}；可逆性保证「基」与「坐标」可以互相转换。相似变换的因子顺序不能交换。

用法　相似矩阵有相同的特征多项式、特征值、迹、行列式和秩，但这些必要条件合在一起仍可能不足以判定相似。重根时要比较几何重数或 Jordan 块。

定理（Jordan 形与可对角化）　复矩阵 A 可相似变换为 Jordan 矩阵 J。对每个特征值 λ，对应 Jordan 块的大小之和是代数重数，块的个数是 $\dim N(A-\lambda I)$；所有块大小为 1 当且仅当 A 可对角化。

关键　一个大小为 r 的块 $J_r(\lambda)$ 只有一个独立特征向量，超对角线的 1 记录不能被普通特征基消掉的链。完整定理见第 27 讲。

用法　遇到重复特征值，先算 $\dim N(A-\lambda I)$，再决定是对角块还是更长的 Jordan 块。$4I_2$ 与 $\begin{bmatrix}4&1\\0&4\end{bmatrix}$ 是最小反例：谱数据相同而不相似。

R3.2.3 SVD、四个基本子空间与截断（第 28、30 讲）

定理（SVD 与四个基本子空间）　任意 $A\in\mathbb{R}^{m\times n}$ 都有完整 SVD

$$A=U\Sigma V^{\mathsf{T}},$$

其中 U,V 正交，Σ 为矩形对角矩阵。正奇异值的个数为 $r=\operatorname{rank}A$；对应的右奇异向量张成 $C(A^{\mathsf{T}})$，其余右向量张成 $N(A)$，对应的左向量分别张成 $C(A)$ 与 $N(A^{\mathsf{T}})$。

关键　先对半正定矩阵 $A^{\mathsf{T}}A$ 求正交谱分解，令 $\sigma_i=\sqrt{\lambda_i}$，再用 $u_i=Av_i/\sigma_i$；零奇异值方向不能

除法，必须补成正交基。

用法　先数秩再写四个空间的维数：$r, n-r, r, m-r$。还可直接得到 $\|A\|_F^2 = \sum_i \sigma_i^2$ 和 $\|Ax\|^2 = \sum_i \sigma_i^2 (v_i^{\mathsf{T}} x)^2$。

定理（最佳低秩近似）　若 $A = \sum_{i=1}^{r} \sigma_i u_i v_i^{\mathsf{T}}$ 且 $\sigma_1 \geq \cdots \geq \sigma_r > 0$，保留前 k 项的

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^{\mathsf{T}}$$

在所有秩不超过 k 的矩阵中使 Frobenius 误差最小，并且

$$\|A - A_k\|_F^2 = \sum_{i>k} \sigma_i^2.$$

关键　秩一项 $u_i v_i^{\mathsf{T}}$ 两两正交，故删去项的平方误差就是权重平方和；奇异值必须按降序排列，才能在固定预算下保留最大能量。

用法　压缩题先计算尾部能量和，而不是只看最后一个奇异值。$k = 0$ 给出零矩阵，$k = r$ 精确恢复；相对误差的平方是尾部能量除以总能量。

R3.2.4 线性变换、换基与伪逆（第 29、31 讲）

定理（输入、输出坐标与换基）　设线性变换在旧基下的矩阵为 A。若输入端旧基到新基的矩阵为 P，输出端旧基到新基的矩阵为 Q，则新坐标矩阵为

$$[T]_{\text{new}} = Q^{-1} A P.$$

当输入、输出使用同一组基时才简化为 $P^{-1}AP$。

关键　从新输入坐标出发，先乘 P 还原旧输入，再乘 A，最后乘 Q^{-1} 翻译成新输出；按这三个动作写乘积最不易反序。

用法　题目出现两组不同的基时，绝不能直接套 $P^{-1}AP$。先在每个向量上标记「属于输入还是输出」，再检查矩阵尺寸。

定理（SVD 伪逆与最小二乘）　设 $A = U\Sigma V^{\mathsf{T}}$，把 Σ 的正对角元 σ_i 换成 $1/\sigma_i$、零位置仍置零，得到 Σ^{+}。则

$$A^{+} = V\Sigma^{+}U^{\mathsf{T}}, qquad AA^{+} = \text{proj}_{C(A)}, \qquad A^{+}A = \text{proj}_{C(A^{\mathsf{T}})}.$$

对任意 b，$A^{+}b$ 是最小二乘解中范数最小者，所有最小二乘解为 $A^{+}b + z$（$z \in N(A)$）。

关键　在 SVD 坐标中，Σ 只在正奇异方向上可逆；零方向送到零。两项投影分别保留列空间和行

空间的坐标。

用法　满列秩时可改写为 $(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$，满行秩时可改写为 $A^{\mathsf{T}}(AA^{\mathsf{T}})^{-1}$。秩亏时不要强行写普通逆，直接用 SVD 或四个基本子空间。

R3.3　题型与例题

R3.3.1　题型一：谱分解、正定性与矩阵函数

■ **例 R3.1　Unit 3 Exam 第 1 题改编：正交谱分解与分类**　给定

$$H=\frac{1}{2}\begin{bmatrix}1&1&-1&-1\\1&-1&-1&1\\1&1&1&1\\1&-1&1&-1\end{bmatrix},\qquad D=\operatorname{diag}(0,1,2,3),$$

令 $B=HDH^{\mathsf{T}}$，$C=H\operatorname{diag}(1,\frac{1}{2},\frac{1}{3},\frac{1}{4})H^{\mathsf{T}}$。求 B,C，并判断它们的对称性、正定性和奇异性。

解　每行与每列的内积都给出单位向量或零，因此 $H^{\mathsf{T}}H=I$。正交谱分解直接给出

$$B=\begin{bmatrix}\frac{3}{2}&-\frac{1}{2}&-1&0\\-\frac{1}{2}&\frac{3}{2}&0&-1\\-1&0&\frac{3}{2}&-\frac{1}{2}\\0&-1&-\frac{1}{2}&\frac{3}{2}\end{bmatrix},\quad C=\frac{1}{48}\begin{bmatrix}25&7&11&5\\7&25&5&11\\11&5&25&7\\5&11&7&25\end{bmatrix}.$$

B 的特征值是 $0,1,2,3$，所以它对称、半正定但奇异，不是正定。C 的特征值为 $1,1/2,1/3,1/4$，全部严格为正，故 C 对称正定且可逆。题目中的 H 只是正交矩阵；奇异值全为 1 并不能推出对称或正定。 ♦

点评　看到 $H^{\mathsf{T}}H=I$ 就把计算转到对角坐标；分类只看对角谱的符号。最后一句是第 24 讲的边界提醒：正交、对称、正定是三种不同条件。

■ **例 R3.2　Unit 3 Exam 第 2 题改编：特征方向与矩阵函数**　令

$$A=\begin{bmatrix}-1&2&4\\0&0&5\\0&0&1\end{bmatrix}.$$

求特征值和一组特征向量，判断 A^{1001} 是否等于 A，并写出 e^{At} 的特征坐标形式。另求 $A^{\mathsf{T}}A$ 的秩与正负性。

解　A 是上三角矩阵，特征值为 $-1,0,1$。对应向量可取

$$q_{-1}=(1,0,0)T,quadq_0=(2,1,0)T,quadq_1=(7,5,1)T.$$

三向量线性无关，所以 $A=S\operatorname{diag}(-1,0,1)S^{-1}$，其中 $S=[q_{-1}\ q_0\ q_1]$。于是

$$A^{1001}=S\operatorname{diag}(-1,0,1)^{1001}S^{-1}=A,$$

而 $A^{1000}\neq I$，因为零特征值仍给出零。矩阵指数为

$$e^{At}=S\,\mathrm{diag}(e^{-t},1,e^{t})S^{-1}.$$

直接乘得

$$A^{\mathsf T}A=\begin{bmatrix}1&-2&-4\\-2&4&8\\-4&8&42\end{bmatrix}.$$

它是 $A^{\mathsf T}A$，所以半正定；$\operatorname{rank}A=2$，故秩为 2，有两个正特征值和一个零特征值。一般不应把 $A^{\mathsf T}A$ 与 A 的特征向量混为一谈。 ♦

点评　上三角形先读谱，再检查特征向量是否足够多；有特征基才能逐项计算幂和指数。正负性用 $x^{\mathsf T}A^{\mathsf T}Ax=\|Ax\|^2$，不必展开三次特征多项式。

■ **例 R3.3　Unit 3 Exam 第 3 题改编：正交特征基下的逆**　设 $Aq_i=\lambda_i q_i$，q_i 是正交标准基且 $\lambda_i>0$。若 $b=\sum_i c_iq_i$，求 $A^{-1}b$ 在这组基下的系数，并特别写出第一系数。

解　将 b 展开后逐个方向作用：$A^{-1}q_i=\lambda_i^{-1}q_i$，所以

$$A^{-1}b=\sum_i\frac{c_i}{\lambda_i}q_i,\qquad c_i=q_i^{\mathsf T}b.$$

因此第一系数为

$$d_1=\frac{c_1}{\lambda_1}=\frac{q_1^{\mathsf T}b}{\lambda_1}.$$

这里正交归一性使坐标直接由内积读出；若向量只正交而未归一，则应除以 $q_i^{\mathsf T}q_i$。 ♦

点评　先把问题换到特征坐标，再把矩阵运算变成标量运算。最容易漏掉的是「正交归一」条件，它决定了 $c_i=q_i^{\mathsf T}b$ 这个简式。

R3.3.2　题型二：投影、四空间与秩亏

■ **例 R3.4　Final Exam 第 8 题改编：最小二乘投影**　给定

$$A=\begin{bmatrix}1&0\\3&0\\0&1\\0&3\end{bmatrix},\qquad b=(1,2,4,5)^{\mathsf T}.$$

求最小二乘系数、拟合向量、残差和投影矩阵。

解　两列独立，且

$$A^{\mathsf T}A=\begin{bmatrix}10&0\\0&10\end{bmatrix},\qquad A^{\mathsf T}b=(7,19)^{\mathsf T}.$$

法方程给 $\widehat{x}=(7/10,19/10)^{\mathsf{T}}$。因此

$$p=A\widehat{x}=(7/10,21/10,19/10,57/10)^{\mathsf{T}},\quad e=b-p=(3/10,-1/10,21/10,-7/10)^{\mathsf{T}}.$$

直接核对 $A^{\mathsf{T}}e=(0,0)^{\mathsf{T}}$，所以 $e\in N(A^{\mathsf{T}})$。投影矩阵为

$$P=A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}=\frac{1}{10}\begin{bmatrix}1&3&0&0\\3&9&0&0\\0&0&1&3\\0&0&3&9\end{bmatrix},$$

并满足 $P^2=P=P^{\mathsf{T}}$。 ♦

点评　先检查列独立，才可使用 $(A^{\mathsf{T}}A)^{-1}$。法方程的几何含义是「残差垂直于列空间」；只算出系数而不核对 $A^{\mathsf{T}}e=0$，容易在投影题中漏项。

■ 例 R3.5 **Final Exam 第 1 题改编：由特殊解恢复四个基本空间**　设 A 是 3×4 矩阵，且 $Ax=0$ 的两个特殊解为

$$x_1=(1,1,1,0)^{\mathsf{T}},\qquad x_2=(-2,-1,0,1)^{\mathsf{T}}.$$

给出 $N(A)$、$C(A^{\mathsf{T}})$ 的基和四个基本空间的维数；说明如何从实际矩阵 A 取 $C(A)$ 的基。

解　两向量独立且题设说它们是特殊解，因此

$$N(A)=\operatorname{span}\{x_1,x_2\},\qquad \dim N(A)=2.$$

把它们写成自由变量参数，得到一个可取的行最简形

$$R=\begin{bmatrix}1&0&-1&2\\0&1&-1&1\\0&0&0&0\end{bmatrix},$$

其零空间正好由 x_1,x_2 张成。因此 $C(A^{\mathsf{T}})=\mathrm{Row}(A)$ 可取 R 的前两行作为基，维数为 2。秩–零度关系给出

$$\dim C(A)=2,\qquad \dim N(A^{\mathsf{T}})=3-2=1.$$

实际的 $C(A)$ 不能只凭两个特殊解写出具体向量；先对给定的 A 消元，取原矩阵中与 R 的主元列编号相同的前两列作为列空间基。 ♦

点评　特殊解先回答输入端的零空间，再用秩–零度关系补齐其余三维数。列空间基必须回到原矩阵取，不能把 R 的列直接当作 A 的列。

■ 例 R3.6 **第 31 讲例题改编：秩一矩阵的伪逆**　求

$$A=\begin{bmatrix}4&3\\8&6\end{bmatrix}$$

的伪逆 A^{+}，并写出 AA^{+} 与 $A^{+}A$ 的像空间。

解　矩阵秩为 1。取

$$u_1=\frac{1}{\sqrt{5}}(1,2)^{\mathsf T},\quad v_1=\frac{1}{5}(4,3)^{\mathsf T},\quad \sigma_1=5\sqrt{5}.$$

于是 $A=\sigma_1 u_1 v_1^{\mathsf T}$，并且

$$A^{+}=\sigma_1^{-1}v_1u_1^{\mathsf T}=\frac{1}{125}\begin{bmatrix}4&8\\3&6\end{bmatrix}.$$

相乘得到

$$AA^{+}=\frac{1}{5}\begin{bmatrix}1&2\\2&4\end{bmatrix},\qquad A^{+}A=\frac{1}{25}\begin{bmatrix}16&12\\12&9\end{bmatrix}.$$

前者的像空间是 $C(A)=\operatorname{span}\{(1,2)^{\mathsf T}\}$，后者的像空间是 $C(A^{\mathsf T})=\operatorname{span}\{(4,3)^{\mathsf T}\}$。两者都是对称幂等的正交投影。 ♦

点评　秩亏时只在正奇异方向上取倒数。若误把 $A^{\mathsf T}A$ 的普通逆写出来，会在零特征值处除以零；伪逆正是把这个方向保留为零。

■ **例 R3.7　第 30 讲例题改编：最佳秩一近似**　把 $M=\begin{bmatrix}3&1\\1&3\end{bmatrix}$ 看成一个 2×2 数据块，求其最佳秩一近似和 Frobenius 误差。

解　M 对称正定，特征向量 $u_1=v_1=(1,1)^{\mathsf T}/\sqrt{2}$、$u_2=v_2=(1,-1)^{\mathsf T}/\sqrt{2}$ 对应特征值（也即奇异值）$4,2$。故

$$M_1=4u_1v_1^{\mathsf T}=\begin{bmatrix}2&2\\2&2\end{bmatrix},\qquad \|M-M_1\|_F=2.$$

原矩阵平方范数为 20，保留第一方向的能量为 16，相对误差为 $2/\sqrt{20}=1/\sqrt{5}$。 ♦

点评　奇异值按降序时，保留最大的 4 而不是最先写出的任意方向。尾部公式直接给误差 $\sqrt{2^2}=2$，无需逐项比较所有秩一矩阵。

R3.4　常见错误

把正交误当作正定　错误说法：方阵的奇异值全为 1 就一定对称、正定。反例是旋转矩阵 $\begin{bmatrix}0&-1\\1&0\end{bmatrix}$，奇异值全为 1，但不对称且二次型不恒正。正确说法：奇异值全为 1 只说明矩阵正交。

把 $A^{\mathsf T}A$ 的特征向量当成 A 的　错误说法：$A^{\mathsf T}A$ 与 A 总有同一组特征向量。反例是例 R3.2：A 非对称，$A^{\mathsf T}A$ 是对称半正定矩阵，维数和方向都可能不同。正确说法：$A^{\mathsf T}A$ 的特征向量是右奇异向量，需通过 Av_i 才得到左方向。

重复特征值只看代数重数　错误说法：特征值有 n 次重数就自动有 n 个特征向量。反例 $J_2(4)$ 只有一个特征向量，不能对角化。正确说法：要比较几何重数 $\dim N(A-\lambda I)$；Jordan 块大小记录缺少的方向。

换基公式因端点而反序　错误说法：输入、输出基不同也写 $P^{-1}AP$。若输入翻译矩阵为 P、输出为 Q，正确公式是 $Q^{-1}AP$；只有 $P=Q$ 时才是相似变换。

伪逆在零方向取倒数　错误说法：把所有对角元都替换成倒数。零奇异值没有倒数，必须在 Σ^{+} 中保留零；否则 AA^{+} 不再是列空间投影，且会制造不存在的输入方向。

R3.5　全课程回顾

全课程可以沿着一条问题链复习：先问 $Ax=b$ 是否可解，再问解的自由度和最近点；随后用行列式、特征值和矩阵函数理解变换，最后用 SVD、换基和伪逆把这些结构放到统一坐标中。图 R3.2 将每个阶段和本复习课的收束点连起来。

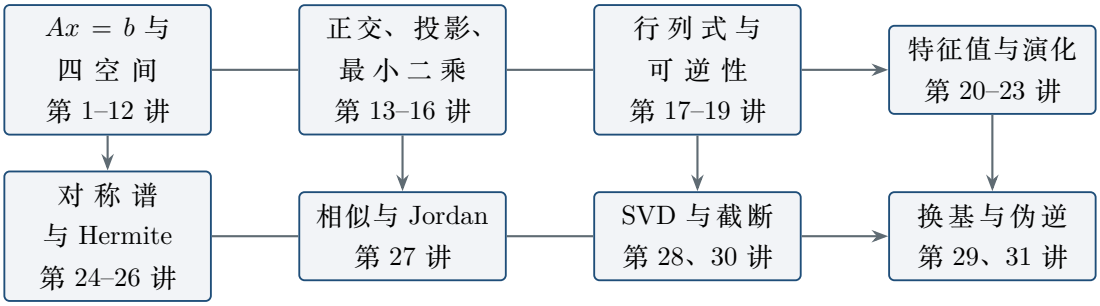


图 R3.2　先用消元和正交性回答可解与逼近，再用谱、SVD 和伪逆处理可逆方向与不可辨识方向。

第一部分：方程与空间（第 1–12 讲）。　消元把 $Ax=b$ 变成主元和自由变量；列空间回答右端是否可达，零空间回答解是否唯一。四个基本子空间把输入端、输出端和正交约束分开。后面的投影、最小二乘和伪逆都在重复使用这套语言。

第二部分：正交与逼近（第 13–16 讲）。　正交补把误差写成最近点条件，法方程把它变成 $A^{\mathsf{T}}e=0$。Gram–Schmidt 与 QR 只是把同一列空间换成正交坐标，计算更稳定，也为 Fourier 和 SVD 的正交方向作准备。

第三部分：行列式与系统（第 17–19 讲）。　行列式记录可逆性、面积和体积缩放；Cramer 法则适合小型、非奇异方阵。遇到奇异矩阵时，仍应回到相容性、秩和零空间，不能把行列式比值当成通用公式。

第四部分：特征值与演化（第 20–23 讲）。　特征方向把矩阵幂、矩阵指数和迭代分解成标量问题。对称矩阵有正交特征基，Markov 和 Fourier 则分别把长期状态与周期信号写成特殊的谱坐标。

第五部分：三种分解的关系（第 24–31 讲）。 实对称谱分解要求同一组正交输入输出方向；SVD 对任意长方形矩阵分别选择输入、输出方向。相似变换保留同一线性变换的谱不变量，换基公式处理不同端点的坐标；伪逆只反转非零奇异方向，因而同时给出投影、最小二乘和最小范数解。

本讲小结

- 实对称 A 才能写成 $Q\Lambda Q^{\mathsf{T}}$；Hermite 情形把转置换为共轭转置，Fourier 满足 $F_n^H F_n = nI$。
- 正定性可由正特征值、$x^{\mathsf{T}}Ax > 0$ 或顺序主子式全正判定；$A^{\mathsf{T}}A$ 永远半正定，列独立时才正定。
- 相似保留特征多项式、迹、行列式和秩；重复特征值还要看几何重数与 Jordan 块。
- SVD 的正奇异方向给出 $C(A^{\mathsf{T}})$ 与 $C(A)$，零方向给出 $N(A)$ 与 $N(A^{\mathsf{T}})$。
- SVD 截断的平方误差是尾部奇异值平方和；应按奇异值降序保留方向。
- 输入、输出基不同用 $Q^{-1}AP$；伪逆 $A^+ = V\Sigma^+U^{\mathsf{T}}$ 在可辨识方向上反转，在零方向上取零。

练习

R3.1 ★ 对例 R3.1 的 H，验证 $H^{\mathsf{T}}H = I$，并判断 H 是否必然对称。

R3.2 ★ 设 $F_4 = (i^{jk})_{0\le j,k<4}$。写出 $F_4^H F_4$，并说明 F_4^{-1} 的归一化因子。

R3.3 ★★ 判断 $\begin{bmatrix}1 & 2\\ 2 & 5\end{bmatrix}$ 是否正定；分别用顺序主子式和二次型给出理由。

R3.4 ★★ $J = \begin{bmatrix}3 & 1\\ 0 & 3\end{bmatrix}$ 的代数重数和几何重数各是多少？它是否可对角化？

R3.5 ★★ 对 $A = \begin{bmatrix}3 & 1\\ 1 & 3\end{bmatrix}$ 求奇异值，并写出最佳秩一近似的 Frobenius 误差。

R3.6 ★★ 若 A 为 5×3、秩为 2，写出四个基本空间的维数。

R3.7 ★★ 设输入基变换为 $P = \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$，输出基变换为 $Q = \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}$，且旧坐标矩阵为 $A = I_2$。求新坐标矩阵。

R3.8 ★★★ 对 $A = \begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}$ 写出 A^+、AA^+ 和 A^+A，并说明两个投影的像空间。

R3.9 ★★★ 对例 R3.4 的 A, b，验证 $A^{\mathsf{T}}(b - A\hat{x}) = 0$，并计算最小残差平方范数。

R3.10 ★★★ 综合题：若 $A = U\Sigma V^{\mathsf{T}}$ 的正奇异值是 $5, 2$，且 $b = 3u_1 - 4u_2 + u_3$（$u_3 \in N(A^{\mathsf{T}})$），求 A^+b、AA^+b 和残差。

练习答案与提示

R3.1 逐列内积得到 $H^{\mathsf{T}}H = I_4$；但 $H_{12} = 1/2$ 而 $H_{21} = 1/2$、$H_{13} = -1/2$ 而 $H_{31} = 1/2$，所以 $H \neq H^{\mathsf{T}}$，正交不推出对称。

R3.2 不同频率列的内积是 $\sum_{j=0}^{3} i^{j(k-l)}$，故 $F_4^H F_4 = 4I_4$，从而 $F_4^{-1} = \frac{1}{4}F_4^H$。

R3.3 $D_1 = 1 > 0$、$D_2 = 5 - 4 = 1 > 0$，所以正定。直接配方 $x^{\mathsf{T}}Ax = (x+2y)^2 + y^2 > 0$（$x, y$ 不全为零）。

R3.4 唯一特征值 3 的代数重数为 2；$J - 3I$ 的零空间是一维，所以几何重数为 1，不可对角化，只有一个大小为 2 的 Jordan 块。

R3.5 特征值为 $4, 2$，故奇异值为 $4, 2$。最佳秩一近似保留 4 方向，尾部误差为 $\sqrt{2^2} = 2$。

R3.6 $\dim C(A) = \dim C(A^{\mathsf{T}}) = 2$，$\dim N(A) = 3 - 2 = 1$，$\dim N(A^{\mathsf{T}}) = 5 - 2 = 3$。

R3.7 $Q^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$，所以 $Q^{-1}AP = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$。

R3.8 $A = uv^{\mathsf{T}}$（$u = (1,2)^{\mathsf{T}}$、$v = (1,1)^{\mathsf{T}}$），$A^{+} = \frac{1}{10}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$；$AA^{+} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$，像为 $\operatorname{span}\{(1,2)^{\mathsf{T}}\}$；$A^{+}A = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$，像为 $\operatorname{span}\{(1,1)^{\mathsf{T}}\}$。

R3.9 例题中的残差为 $(3/10, -1/10, 21/10, -7/10)^{\mathsf{T}}$，所以 $A^{\mathsf{T}}e = (0,0)^{\mathsf{T}}$；平方范数为 $9/100 + 1/100 + 441/100 + 49/100 = 5$。

R3.10 $A^{+}b = (3/5)v_1 - (4/2)v_2 = \frac{3}{5}v_1 - 2v_2$；$AA^{+}b = 3u_1 - 4u_2$；残差为 $b - AA^{+}b = u_3$。左零空间分量被保留在残差中，正奇异方向被反转。