

线性变换一章：计算、证明与应用策略

矩阵表示 · 特征值与特征子空间 · 对角化 · 值域与核 · 哈密顿-凯莱定理

本章统一主线：把“抽象、未知”的线性变换翻译为“具体、已知”的矩阵问题

- 选定线性空间的一组基，求基的像的坐标，求得线性变换在一组基下的矩阵 A 。
- 线性变换的运算、特征值、特征向量、核、值域、对角化等问题，分别转化为矩阵运算、矩阵的特征值、特征向量、相似对角化、矩阵的秩等问题。
- 在矩阵层面完成计算或证明后，再把坐标向量翻译回原来的抽象向量。
- 核心思想不是“多记一套线性变换公式”，而是建立稳定的翻译链： $T \rightarrow A \rightarrow$ 已学矩阵工具 $\rightarrow T$ 。

一、全章知识结构与“翻译规则”

线性变换 T 的问题	矩阵化翻译	主要工具	回译后的结论
求 T 的矩阵 A	求 T (基向量) 的坐标，作为矩阵 A 的列	坐标、矩阵乘法	$[T(v)] = A[v]$
求特征值、特征向量	$T(v) = \lambda v \Leftrightarrow Ax = \lambda x$	特征多项式、齐次方程组	$V_\lambda = \ker(T - \lambda I)$
判断 T 可对角化	找足够多线性无关特征向量	特征子空间维数、相似	$V = \bigoplus V_\lambda$
求 $\ker T$ 与 $\operatorname{Im} T$	$Ax = 0$ 与 A 的列空间	基础解系、秩、主元列	$\dim \ker T + \dim \operatorname{Im} T = \dim V$
处理 T 的高次幂/逆/多项式	处理 A 的高次幂/逆/多项式	哈密顿-凯莱定理	利用特征多项式降次计算

二、线性变换的矩阵及运算的矩阵：计算策略

设 $T: V \rightarrow V$, V 的基为 $\alpha = (\alpha_1, \dots, \alpha_n)$ 。矩阵 A 的第 j 列，恰好是 $T(\alpha_j)$ 在基 α 下的坐标。

$$[T(\alpha_1)]_\alpha [T(\alpha_2)]_\alpha \cdots [T(\alpha_n)]_\alpha = A, \text{ 且 } [T(v)]_\alpha = A[v]_\alpha$$

策略 1：求线性变换矩阵——“逐个作用基向量，坐标按列排”

第 1 步：先取定一组基 $\alpha_1, \dots, \alpha_n$ 。

第 2 步：依次计算 $T(\alpha_1), \dots, T(\alpha_n)$ 。

第 3 步：把每个 $T(\alpha_j)$ 用基 $\alpha_1, \dots, \alpha_n$ 线性表示，得到坐标列向量。

第 4 步：将这些坐标列向量按 $j=1, \dots, n$ 的顺序拼成矩阵 A 。

检查：任取 v ，若 $[v]_\alpha = x$ ，应满足 $[T(v)]_\alpha = Ax$ 。

策略 2：线性变换运算的矩阵——“先矩阵化，再按矩阵规则运算”

固定一组基，线性变换 T, S 在该基下的矩阵为 $[T], [S]$ ：则 $[T+S]=[T]+[S]$ ， $[cT]=c[T]$ 。

复合最易出错：若先做 S 、再做 T ，则 $[T \circ S]=[T][S]$ ；矩阵乘法顺序与作用顺序相反。

若 T 可逆，则 $[T^{-1}]=[T]^{-1}$ ； $[T^k]=[T]^k$ ；对多项式 p ，有 $[p(T)]=p([T])$ 。

同一线性变换在不同基下的矩阵彼此相似：若坐标变换矩阵为 P ，则 $B=P^{-1}AP$ 。

例 1：由定义求线性变换的矩阵

在 $P[x]_3$ 中取基 $B=(1, x, x^2)$ ，定义 $T(f)=f'+f(0)(1+x)$ 。求 $[T]B$ 。

1. 计算基向量的像： $T(1)=1+x$ ， $T(x)=1$ ， $T(x^2)=2x$ 。
2. 写成 B 坐标： $[T(1)]B=(1, 1, 0)^T$ ， $[T(x)]B=(1, 0, 0)^T$ ， $[T(x^2)]B=(0, 2, 0)^T$ 。
3. 按列排列： $A=[T]B=[[1, 1, 0], [1, 0, 2], [0, 0, 0]]$ 。
4. 验证：若 $f=a+bx+cx^2$ ，则 $A(a, b, c)^T=(a+b, a+2c, 0)^T$ ，对应 $T(f)=(a+b)+(a+2c)x$ 。

结论：线性变换的作用规律已经完全转化为坐标矩阵 A 的乘法规律。

例 2：复合线性变换的矩阵

在 R^2 标准基下， $S(x, y)=(x+y, y)$ ， $T(x, y)=(2x, x-y)$ 。求 $T \circ S$ 的矩阵。

1. $[S]=[[1, 1], [0, 1]]$ ， $[T]=[[2, 0], [1, -1]]$ 。
2. 因为先做 S 再做 T ，所以 $[T \circ S]=[T][S]=[[2, 2], [1, 0]]$ 。
3. 直接检验： $T(S(x, y))=T(x+y, y)=(2x+2y, x)$ ，与矩阵结果一致。

结论：复合问题不必重新从抽象定义展开，矩阵化后直接用矩阵乘法即可；但必须牢记乘法顺序。

三、线性变换的特征值、特征向量与特征子空间：计算策略

设 $T: V \rightarrow V$ ，选定基 B 后 $A=[T]B$ 。若 $v \neq 0$ ，令 $x=[v]B$ ，则

$$T(v)=\lambda v \Leftrightarrow Ax=\lambda x \Leftrightarrow (A-\lambda E)x=0$$

因此，抽象的特征向量问题被转化为矩阵的齐次线性方程组问题。特征子空间 $V_\lambda=\ker(T-\lambda E)$ ，在坐标下对应 $\ker(A-\lambda E)$ 。

标准计算流程：特征多项式 → 特征值 → 齐次方程组 → 特征子空间

第 1 步：选基得到 $A=[T]$ 。

第 2 步：解 $\det(\lambda E - A) = 0$ ，得到全部特征值。

第 3 步：对每个 λ ，解 $(A - \lambda E)x = 0$ ；所有非零解对应 λ 的特征向量。

第 4 步：基础解系给出特征子空间的坐标基；再按所选基把坐标向量翻译回 V 中的向量。

第 5 步：记录 $\dim V_\lambda = n - r(A - \lambda E)$ 。它就是 λ 的几何重数，是判断可对角化的关键数据。

重要区分：特征值的代数重数 = 它作为特征多项式根的重数；几何重数 = $\dim V_\lambda$ 。总有 $1 \leq \text{几何重数} \leq \text{代数重数}$ 。

例 3：重复特征值下的特征子空间维数

设 T 在某组基下的矩阵 $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ 。求特征值、各特征子空间的维数与基。

1. A 为上三角矩阵，所以其特征多项式 $= (\lambda - 2)^2(\lambda - 3)$ 。特征值为 2（代数重数 2）和 3（代数重数 1）。
2. $\lambda = 2$ ： $A - 2E = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ ，故 $x_2 = x_3 = 0$ ， x_1 自由。坐标特征子空间 $\ker(A - 2I) = \text{span}\{(1, 0, 0)^T\}$ 。
3. 因此 $\dim V_2 = 1$ ；若原基为 $B = (\alpha_1, \alpha_2, \alpha_3)$ ，则 $E_2 = \text{span}\{\alpha_1\}$ ，一组基为 $\{\alpha_1\}$ 。
4. $\lambda = 3$ ：解 $(A - 3I)x = 0$ 得 $x_1 = x_2 = 0$ ， x_3 自由，所以 $\dim V_3 = 1$ ， $V_3 = \text{span}\{\alpha_3\}$ ，一组基为 $\{\alpha_3\}$ 。

结论： $\lambda = 2$ 的几何重数 1 小于代数重数 2，这一信息将直接决定 T 不可对角化。

例 4：多项式空间中的特征子空间

在 P_2 上定义 $T(f(x)) = f(-x)$ ，取基 $B = (1, x, x^2)$ 。求特征值、特征子空间的维数与基。

1. $T(1) = 1$ ， $T(x) = -x$ ， $T(x^2) = x^2$ ，所以 $[T]B = \text{diag}(1, -1, 1)$ 。
2. 特征值为 1 和 -1。
3. $\lambda = 1$ 时， $V_1 = \{a + cx^2\} = \text{span}\{1, x^2\}$ ， $\dim V_1 = 2$ ，一组基为 $\{1, x^2\}$ 。
4. $\lambda = -1$ 时， $V_{-1} = \{bx\} = \text{span}\{x\}$ ， $\dim V_{-1} = 1$ ，一组基为 $\{x\}$ 。

结论：在抽象空间中，“偶多项式/奇多项式”结构通过对角矩阵一目了然。

四、线性变换可对角化：计算与证明策略

线性变换 T 可对角化，指存在 V 的一组基，使得 T 在这组基下的矩阵为对角矩阵。矩阵化后，问题等价于： A 是否相似于对角矩阵。

$$T \text{ 可对角化} \Leftrightarrow V \text{ 有一组由特征向量组成的基} \Leftrightarrow \sum \lambda \dim V_\lambda = n$$

计算判定策略 A：按“代数重数 vs. 几何重数”逐个检查

第 1 步：求特征多项式，列出不同特征值 $\lambda_1, \dots, \lambda_s$ 及其代数重数 $m_1, \dots, m_s$ 。

第 2 步：分别计算 $V\lambda_i$ ，得到几何重数 $g_i = \dim V\lambda_i$ 。

第 3 步：若对每个 i 都有 $g_i = m_i$ ，则可对角化；只要有一个 $g_i < m_i$ ，就不可对角化。

第 4 步：若可对角化，将各特征子空间的基合并成 n 个特征向量，作为 P 的列，则 $P^{-1}AP = D$ 。

证明策略 B：根据题目条件选“最短路径”

若已知有 n 个互不相同的特征值：直接用“不同特征值的特征向量线性无关”得到特征向量基。

若已知 $V = V\lambda_1 \oplus \dots \oplus V\lambda_s$ ：把各特征子空间的基合并，即得特征向量基，因此可对角化。

若要证明不可对角化：寻找某个重复特征值，使 $\dim V\lambda$ 小于其代数重数；或证明总共找不到 n 个线性无关特征向量。

若课程已讲最小多项式： T 可对角化 $\Leftrightarrow$ 最小多项式在数域上分裂且无重根。

例 5：构造对角化矩阵

设 $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ 。判断对应线性变换是否可对角化，并求一组对角化基。

- 特征多项式 $= (\lambda-1)(\lambda-2)(\lambda-3)$ ，有 3 个互不相同的特征值，所以必可对角化。
- $\lambda=1$ ：特征向量可取 $v_1 = (1, 0, 0)^T$ ； $\lambda=2$ ：解得可取 $v_2 = (1, 1, 0)^T$ ； $\lambda=3$ ：可取 $v_3 = (0, 0, 1)^T$ 。
- 令 $P = (v_1, v_2, v_3) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ ，则 $P^{-1}AP = \text{diag}(1, 2, 3)$ 。
- 若原线性空间的基为 B ，则把 v_1, v_2, v_3 作为 B 坐标翻译回 V ，即得到一组由 T 的特征向量组成的对角化基。

结论：对角化的本质不是“把矩阵算成对角”，而是“找到一组特征向量基”。

例 6：用特征子空间维数证明不可对角化

在 P_2 中定义 $T(f) = f + f'$ 。判断 T 是否可对角化。

- 在 $B = (1, x, x^2)$ 下， $[T]B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ 。因此特征多项式 $= (\lambda-1)^3$ ，唯一特征值为 1，代数重数为 3。
- $V_1 = \ker(T-I) = \{f: f' = 0\} = \text{span}\{1\}$ ，所以 $\dim V_1 = 1$ 。
- 几何重数 $1 <$ 代数重数 3 ，因此不能找到 3 个线性无关的特征向量。

结论： T 不可对角化。证明非常短，因为抽象问题被转化为秩 $(A-E)$ 的问题。

五、线性变换的值域与核：计算、维数、基与证明策略

设 $T: V \rightarrow W$, $\dim V = n$, 取基 $\alpha = (\alpha_1, \dots, \alpha_n)$, $A = [T]\alpha$ 。

计算策略：核用“解方程”，值域用“找主元列”

求 $\text{Ker } T$: 解 $Ax=0$, 基础解系给出核的一组坐标基; 再按定义域基 α 回译为 V 中的向量。

求 $\text{Im } T$: $\text{Im } T = \text{span}\{T(\alpha_1), \dots, T(\alpha_n)\}$; 在矩阵 A 中求列秩, 选原矩阵 A 的主元列, 对应的 $T(\alpha_i)$ 构成值域的一组基。

维数: $\dim \text{Ker } T = n - r(A)$, $\dim \text{Im } T = r(A)$, 故 $\dim \text{Ker } T + \dim \text{Im } T = n$ 。

注意: 求值域基时必须取“原矩阵”的主元列, 不能把行变换后的阶梯矩阵列直接当作原值域的基。

证明策略：维数公式（秩–零度定理）的两条路径

矩阵路径: 坐标同构把 $\text{Ker } T$ 与 $\text{Ker } A$ 、 $\text{Im } T$ 与 A 的列空间对应起来, 再用 $n - r(A) + r(A) = n$ 。

空间路径: 先取 $\text{Ker } T$ 的基 $\alpha_1, \dots, \alpha_k$, 并扩充成 V 的基 $\alpha_1, \dots, \alpha_n$; 证明 $T(\alpha_{k+1}), \dots, T(\alpha_n)$ 正好是 $\text{Im } T$ 的一组基, 从而 $\dim V = \dim \text{Ker } T + \dim \text{Im } T$ 。

证明“值域等于某生成空间”时常用: 一边由 T (任意向量) 的线性展开得包含关系, 另一边注意每个 $T(\alpha_i)$ 本身就在 $\text{Im } T$ 中。

例 7: 导数变换的核与值域

在 P_2 上取 $B = (1, x, x^2)$, 定义 $D(f) = f'$ 。求 $\text{Ker } D$ 、 $\text{Im } D$ 的维数与基, 并验证维数公式。

- $[D]B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$, 秩为 2。
- 核: 解 $Ax=0$ 得 $b=c=0$, a 自由。因此 $\text{Ker } D = \text{span}\{1\}$, $\dim \text{Ker } D = 1$ 。
- 值域: A 的第 2、3 列为主元列, 对应 $D(x)=1$, $D(x^2)=2x$, 所以 $\text{Im } D = \text{span}\{1, x\}$, $\dim \text{Im } D = 2$ 。
- $\dim \text{Ker } D + \dim \text{Im } D = 1 + 2 = 3 = \dim P_2$ 。

结论: 核描述“被 T 压成 0 的方向”, 值域描述“真正能被 T 输出的方向”; 两者的维数由矩阵秩统一控制。

例 8: 从矩阵同时求核与值域

设 $T: \mathbb{R}^4 \rightarrow \mathbb{R}^3$ 在标准基下的矩阵 $A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 3 & 2 & 1 \end{bmatrix}$ 。求 $\text{Ker } T$ 与 $\text{Im } T$ 的基和维数。

- 第 3 行 = 第 1 行 + 第 2 行, 且前两行线性无关, 所以 $r(A)=2$ 。原矩阵第 1、2 列可作为主元列。
- 因此 $\text{Im } T = \text{span}\{(1, 0, 1)^T, (2, 1, 3)^T\}$, $\dim \text{Im } T = 2$ 。
- 解 $Ax=0$: 由 $x_1 + 2x_2 + x_3 = 0$, $x_2 + x_3 + x_4 = 0$ 。令 $x_3 = s$, $x_4 = t$, 则 $x_2 = -s - t$, $x_1 = s + 2t$ 。

4. 故 $\text{Ker } T = \text{span}\{(1, -1, 1, 0)^T, (2, -1, 0, 1)^T\}$, $\dim \text{Ker } T = 2$ 。

5. 验证: $2+2=4=\dim \mathbb{R}^4$ 。

结论: 同一个矩阵 A 同时编码了“哪些方向被压到 0”和“输出空间中能到达哪些方向”。

六、哈密顿-凯莱定理：理解与应用策略

哈密顿-凯莱定理 (Cayley-Hamilton) : 若 A 是 n 阶矩阵, 特征多项式 $f = \det(\lambda E - A)$, 则

$$f(A) = 0. \quad \text{对线性变换 } T \text{ 同样有 } f(T) = 0.$$

理解关键: 特征多项式原本是关于标量 λ 的多项式; 定理说明把其中的 λ 按多项式运算规则替换为 A 后, 得到的是零矩阵。它不是“看到 $\det(\lambda I - A) = 0$ 就随意令 $\lambda = A$ ”, 而是一个需要定理保证的矩阵恒等式。

应用策略 1: 高次幂“降次”

若 $f(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \dots + c_0$, 则 $A^n = -c_{n-1}A^{n-1} - \dots - c_0E$ 。

从此以后所有 $A^k (k \geq n)$ 都可递推化为 $E, A, \dots, A^{n-1}$ 的线性组合。

对于 $p(T)$, 同样可先用特征多项式对多项式 p 做除法, 余式 r 次数 $< n$, 再用 $p(T) = r(T)$ 计算。

应用策略 2: 求逆矩阵 / 线性变换的逆

若 $\det A \neq 0$, 则特征多项式的常数项 $c_0 \neq 0$ 。由 $f(A) = 0$ 把含 E 的项移到一边, 并提取 A , 可写出 A^{-1} 为 $E, A, \dots, A^{n-1}$ 的线性组合。

对可逆线性变换 T , 可同样从 $f(T) = 0$ 写出 T^{-1} 的多项式表达。

这体现了“线性变换的逆仍可由 T 自身的有限次幂组合出来”。

例 9: 用哈密顿-凯莱定理计算高次幂

设 $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ 。不用反复相乘, 求 A^5 。

1. 特征多项式 $f(\lambda) = \lambda^2 - \lambda - 1$ 。由哈密顿-凯莱定理, $A^2 - A - E = 0$, 即 $A^2 = A + E$ 。

2. 递推: $A^3 = A(A + E) = 2A + E$; $A^4 = 3A + 2E$; $A^5 = 5A + 3E$ 。

3. 因此 $A^5 = 5 \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + 3I = \begin{bmatrix} 8 & 5 \\ 5 & 3 \end{bmatrix}$ 。

结论: 二阶矩阵的任意高次幂最终都可降为 $aA + bI$; 这里系数自然出现斐波那契递推。

例 10：用哈密顿–凯莱定理求逆

设 $A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$ 。用哈密顿–凯莱定理表示 A^{-1} 。

- 1. 特征多项式 $f(\lambda) = (\lambda - 2)(\lambda - 3) = \lambda^2 - 5\lambda + 6$ 。于是 $A^2 - 5A + 6E = 0$ 。
- 2. 因 $\det A = 6 \neq 0$ ， A 可逆。两边左乘 A^{-1} ： $A - 5E + 6A^{-1} = 0$ 。
- 3. 故 $A^{-1} = (5E - A)/6 = \begin{bmatrix} 1/2 & -1/6 \\ 0 & 1/3 \end{bmatrix}$ 。
- 4. 若 A 是 T 在某基下的矩阵，则相应地 $T^{-1} = (5I - T)/6$ 。

结论：哈密顿–凯莱定理把“求逆”转化为“低次多项式表示”，无需单独做伴随矩阵计算。

七、五类问题的一页式算法模板

问题	固定翻译动作	落地算法
求 T	对基向量逐个作用	$T(\alpha_j)$ 坐标 $\rightarrow$ 按列组成 A
求特征值/特征向量	$T(v) = \lambda v \rightarrow Ax = \lambda x$	$\det(\lambda E - A) = 0$ ；解 $(A - \lambda E)x = 0$ ；回译坐标
判断可对角化	找特征向量基	算各 V_λ 的维数； $\sum \dim V_\lambda = n$ ；构造 $P^{-1}AP = D$
求 $\text{Ker } T / \text{Im } T$	$\text{Ker} \rightarrow Ax = 0$ ； $\text{Im} \rightarrow$ 列空间	基础解系；原矩阵主元列；用秩读维数
用 C–H 定理	$f(T) = 0$	高次幂降次；求逆；化简 $p(T)$

八、学生最容易出现的错误与纠正提示

- 错误 1：矩阵的列排反了。** 纠正：第 j 列永远是“第 j 个定义域基向量的像”的陪域坐标。
- 错误 2：复合顺序写反。** 纠正：先 S 后 T ，对应 $[T \circ S] = [T][S]$ 。
- 错误 3：求出坐标特征向量后忘记回译。** 纠正：在 P_2 、矩阵空间等抽象空间中， $(1, 0, 1)^T$ 只是坐标；真正向量要按基还原。
- 错误 4：只看特征值个数就判断可对角化。** 纠正：重复特征值时必须算 $\dim E_\lambda$ ；核心是能否得到 n 个线性无关特征向量。
- 错误 5：把代数重数与几何重数混为一谈。** 纠正：代数重数看特征多项式，几何重数是 $n - \text{秩}(A - \lambda E)$ 。
- 错误 6：用阶梯矩阵的主元列作为 $\text{Im } T$ 的原空间基。** 纠正：主元位置由行变换确定，但值域基必须回到“原矩阵 A 的对应列”。
- 错误 7：把哈密顿–凯莱定理理解成“令 $\lambda = A$ ”。** 纠正： $f(A) = 0$ 是定理结论；行列式表达式本身不能直接进行这种替换。

最终记忆口令：一条主线 + 五个落点

主线：选基 $\rightarrow$ 坐标化 $\rightarrow$ 得矩阵 $A \rightarrow$ 使用已学矩阵知识 $\rightarrow$ 回到线性变换。

落点 1： T 的作用 = 矩阵乘法 Ax 。

落点 2：特征向量 = $(A - \lambda E)x = 0$ 的非零解；特征子空间维数 = $n - r(A - \lambda E)$ 。

落点 3：可对角化 = 有一组特征向量基 = 各特征子空间维数之和为 n 。

落点 4： $\text{Ker } T$ = 齐次方程组解空间； $\text{Im } T = A$ 的列空间；两者维数之和为 $\dim V$ 。

落点 5：特征多项式 f 使得 $f(T)=0$ ，把高次、求逆、多项式问题降到有限低次幂。