

计算方法

Computational Methods

Lecture 7: 特征值、谱分解与幂迭代

Eigenvalues, Spectral Decomposition, and Power Iteration

刘景铎

南京大学 · 计算机学院

课程信息

计算方法 Fall 2026

授课教师 刘景铎 liu@nju.edu.cn

上课 周四 5-6 节 · 1-18 周 · 仙 I-319

答疑 周四下午 4 点

助教 周宇恒 · 侯哲

作业 nm_nju_2026@163.com

课程主页 TCS Wiki · Fall 2026

QQ 群 1054896504

入群备注: 姓名、专业、学号

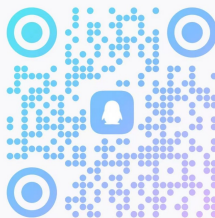
教材: Timothy Sauer, 《数值分析》(原书第 2 版)。

最新通知以课程主页为准。



计算方法 Fall 2026

群号: 1054896504



扫一扫二维码, 加入群聊



回顾：线性迭代的收敛性与特征值

回顾：解 $Ax = b$ 时，可以用 Jacobi 或 Gauss-Seidel 方法构造迭代方程

$$x^{(k+1)} = Bx^{(k)} + c, \quad x^* = Bx^* + c.$$

设 $e^{(k)} = x^{(k)} - x^*$ ，则

$$e^{(k)} = x^{(k)} - x^* = B^k e^{(0)}.$$

对任意初值，迭代均收敛到唯一不动点，当且仅当

$$B^k \rightarrow 0 \iff \rho(B) < 1.$$

特定初值也可能收敛：

即使 $\rho(B) \geq 1$ ，只要初始误差在合适的子空间中，迭代仍可能收敛。

这节课的问题：

能否通过迭代求出 $Av = \lambda v$ 的特征值和特征向量？

观察：沿特征向量方向，矩阵乘法只会改变向量的长度（或符号）。这正是研究矩阵迭代的出发点。

正定矩阵与半正定矩阵

对实对称矩阵 $A = A^\top$ ，考虑二次型 $x^\top Ax$ ：

$$A \succ 0 \iff x^\top Ax > 0 \quad (x \neq 0),$$

$$A \succeq 0 \iff x^\top Ax \geq 0 \quad (\text{任意 } x).$$

分别称为**正定矩阵** (positive definite) 和**半正定矩阵** (positive semidefinite, PSD)。

例： $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$

$$x^\top Ax = 2x_1^2 + x_2^2 > 0 \quad (x \neq 0)$$

例： $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

$$x^\top Ax = x_1^2 \geq 0, \text{ 但并非正定}$$

问题：是否必须检查所有向量，才能判断一个矩阵是否正定？

正规方程与条件数

若 A 可逆，求解 $Ax = b$ 可以转化为

$$\underbrace{A^\top A}_{\text{对称正定}} x = A^\top b.$$

对于任意 A ， $A^\top A$ 对称、半正定；若 A 列满秩，则正定。

需要注意条件数：当 A 可逆时，

$$\kappa_2(A^\top A) = \kappa_2(A)^2.$$

例：若 $A = \text{diag}(1, 10^{-3})$ ，则 $\kappa_2(A) = 10^3$ ， $\kappa_2(A^\top A) = 10^6$ 。

联系第四讲：正规方程会使条件数变成原来的平方。为避免放大数值误差，最小二乘问题通常采用 QR 分解。稍后证明上面的等式。

回顾：特征值与特征向量

非零向量 v 满足 $Av = \lambda v$ ，则 λ 为**特征值**， v 为**特征向量**。

$$\chi_A(t) = \det(tI - A), \quad \chi_A(\lambda) = 0.$$

- ▶ **代数重数**： λ 是特征多项式的几重根。
- ▶ **几何重数**： $\dim \ker(A - \lambda I)$ ，即对应特征空间维数。
- ▶ 如果 A 可对角化，那么每个特征值的代数重数与几何重数相等。

实对称矩阵的谱定理：所有特征值均为实数，且存在一组**标准正交的特征向量**构成 $\mathbb{R}^n$ 的基。

正定矩阵的 Cholesky 分解

回顾：不需要交换行时，高斯消元给出 $A = LU$ ；一般情况下需要写成 $PA = LU$ 。对于**对称正定矩阵**，不需要交换行，而且有

$$A = R^{\top}R, \quad R \text{ 为上三角矩阵, 且 } r_{ii} > 0.$$

这称为 **Cholesky 分解**。

先进行行消元：

由此得到

$$L^{-1}A = U.$$

$$A = LDL^{\top}, \quad D = \text{diag}(d_i), \quad d_i > 0,$$

利用 A 的对称性，再作相应的列消元：

将 D 的平方根合并到三角矩阵中：

$$L^{-1}AL^{-\top} = D.$$

$$A = (D^{1/2}L^{\top})^{\top}(D^{1/2}L^{\top}).$$

与第四讲的联系： $x^{\top}Ax = \|Rx\|_2^2$ 。正定二次型可以写成欧氏范数的平方。

为什么正定矩阵可以逐步消元？

将 A 写成分块形式, $a > 0$:

$$A = \begin{pmatrix} a & r^\top \\ r & B \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ r/a & I \end{pmatrix}}_L \underbrace{\begin{pmatrix} a & 0 \\ 0 & B - rr^\top/a \end{pmatrix}}_{D \text{ 的分块}} \underbrace{\begin{pmatrix} 1 & r^\top/a \\ 0 & I \end{pmatrix}}_{L^\top}.$$

关键是 $S = B - rr^\top/a$ 仍然正定。事实上, 对任意 $y \neq 0$,

$$y^\top S y = \begin{pmatrix} -r^\top y/a \\ y \end{pmatrix}^\top A \begin{pmatrix} -r^\top y/a \\ y \end{pmatrix} > 0.$$

因此, 可以对剩余部分继续消元, 且每一步的主元都为正, 不必交换行。

正定性 $\iff$ 所有特征值为正

定理： 对实对称矩阵 A ,

$$A \succ 0 \iff \lambda_i(A) > 0 \quad (i = 1, \dots, n).$$

($\Rightarrow$) 若 $Av = \lambda v$ 且 $v \neq 0$, 则

$$0 < v^\top Av = \lambda \|v\|_2^2 \quad \Rightarrow \quad \lambda > 0.$$

($\Leftarrow$) 用标准正交特征向量展开 $x = \sum_i \alpha_i v_i$:

$$x^\top Ax = \sum_{i=1}^n \lambda_i \alpha_i^2 > 0 \quad (x \neq 0).$$

同理： $A \succeq 0 \iff \lambda_i(A) \geq 0$ 对所有 i 成立。

Loewner 偏序

对同阶实对称矩阵 A, B , 定义

$$A \succeq B \iff A - B \succeq 0 \iff x^\top A x \geq x^\top B x \ (\forall x).$$

- ▶ 若 $\lambda_{\min}(A) \geq \lambda_{\max}(B)$, 则 $A \succeq B$ (充分条件)。
- ▶ 若 A, B 有一组共同的标准正交特征向量, 则 $A \succeq B$ **当且仅当** 每个共同方向上 A 的特征值不小于 B 的对应特征值。
- ▶ **注意:** 这是偏序而不是全序。仅比较各自排序后的特征值, 并不能判断 $A \succeq B$ 。例如, $\text{diag}(2, 1)$ 与 $\text{diag}(1, 2)$ 的特征值相同, 但两个矩阵在 Loewner 偏序下不可比较。

常见写法: $\alpha I \preceq A \preceq \beta I$ 表示 A 的所有特征值都在 $[\alpha, \beta]$ 内。

实对称矩阵的谱分解

令 $V = (v_1, \dots, v_n)$ 为标准正交特征向量构成的矩阵,

$$V^\top V = VV^\top = I, \quad D = \text{diag}(\lambda_1, \dots, \lambda_n).$$

从 $Av_i = \lambda_i v_i$ 得到

$$AV = VD \implies A = VDV^\top.$$

在原坐标系下：计算 Ax 要进行一般的矩阵乘法。

在特征向量坐标系下：设 $y = V^\top x$, 只需计算 Dy 。

因此，谱分解把矩阵乘法化成了各坐标上独立的乘法。

用正交投影表示谱分解

由于 V 正交,

$$I = \sum_{i=1}^n v_i v_i^\top,$$

$$A = \sum_{i=1}^n \lambda_i v_i v_i^\top.$$

注意: $v_i v_i^\top$ 是到 v_i 张成的直线上的正交投影。 A 在这些方向上分别乘以 λ_i 。

对任意向量 x ,

$$Ax = \sum_i \lambda_i \underbrace{(v_i^\top x) v_i}_{x \text{ 在第 } i \text{ 个方向的投影}}.$$

回顾第四讲: 我们已经用正交投影解释过最小二乘法。

由谱分解得到逆矩阵与伪逆

若所有 $\lambda_i \neq 0$,

$$A^{-1} = VD^{-1}V^{\top} = \sum_{i=1}^n \frac{1}{\lambda_i} v_i v_i^{\top}.$$

若部分特征值为零, 则不能求逆; 但可定义 **Moore-Penrose 伪逆**:

$$A^+ = \sum_{i: \lambda_i \neq 0} \frac{1}{\lambda_i} v_i v_i^{\top}.$$

- ▶ 对非零特征值取倒数; 对应零特征值的部分仍为零。
- ▶ 若 $A \succeq 0$ 且 b 属于 A 的列空间, $x = A^+b$ 是 $Ax = b$ 的**最小范数解**。

证明: 所有解都可以写成 $A^+b + z$ ($z \in \ker A$); 两部分正交, 故 $\|A^+b + z\|_2^2 = \|A^+b\|_2^2 + \|z\|_2^2$ 。

后续: 一般矩阵的伪逆将由第八讲的 SVD 统一解释。

迹、行列式与 Frobenius 范数

令 A 为实对称矩阵，特征值为 $\lambda_1, \dots, \lambda_n$ 。由 $A = VDV^\top$ 可立即得到：

$$\mathrm{tr}(A) = \sum_{i=1}^n A_{ii} = \sum_{i=1}^n \lambda_i,$$

$$\det(A) = \prod_{i=1}^n \lambda_i,$$

$$\|A\|_F^2 = \mathrm{tr}(A^\top A) = \sum_{i,j} A_{ij}^2 = \sum_i \lambda_i^2.$$

注意：迹等于特征值之和，行列式等于特征值之积；对于实对称矩阵，Frobenius 范数的平方等于特征值的平方和。

Gershgorin 圆盘定理

不直接求特征多项式的根，能否先确定特征值可能所在的区域？**Gershgorin circle theorem:**
对任意 $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ ，定义

$$R_i = \sum_{j \neq i} |a_{ij}|, \quad D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i\}.$$

记 $\text{spec}(A)$ 为特征值的集合。定理指出：**所有特征值**都落在这些圆盘的并集中，即

$$\text{spec}(A) \subseteq \bigcup_{i=1}^n D_i.$$

特别地，当各行的非对角元绝对值之和较小时，特征值只能出现在相应对角元附近。

Gershgorin 圆盘定理的证明

取任意特征值 λ 及非零特征向量 v : $Av = \lambda v$ 。选 i 使 $|v_i| = \max_j |v_j| > 0$ 。第 i 行满足

$$\sum_{j \neq i} a_{ij} v_j = (\lambda - a_{ii}) v_i.$$

因此

$$\begin{aligned} |\lambda - a_{ii}| &= \left| \sum_{j \neq i} a_{ij} \frac{v_j}{v_i} \right| \\ &\leq \sum_{j \neq i} |a_{ij}| \frac{|v_j|}{|v_i|} \leq \sum_{j \neq i} |a_{ij}| = R_i. \end{aligned}$$

证明的关键：选择绝对值最大的坐标 v_i ，使得 $|v_j/v_i| \leq 1$ 。

对称矩阵

$$A = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

的 Gershgorin 圆盘与**实轴**的交集分别为 $[3, 5]$ 、 $[1, 5]$ 、 $[1, 3]$ 。

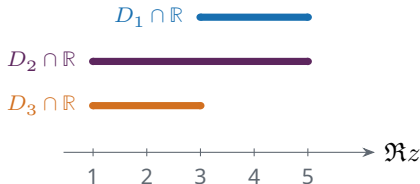
因为 A 实对称，所有特征值都在实轴上；因此

$$\text{spec}(A) \subseteq [1, 5].$$

故 $\lambda_{\min}(A) \geq 1$ ，从而

$$A \succeq I \succ 0.$$

推论：如果实对称矩阵满足 $a_{ii} > \sum_{j \neq i} |a_{ij}|$ （所有 i ），那么它是正定矩阵。



Rayleigh 商

对于实对称矩阵 A ，定义**Rayleigh 商** (Rayleigh quotient)：

$$R_A(x) = \frac{x^\top Ax}{x^\top x}, \quad x \neq 0.$$

在标准正交特征基中写 $x = \sum_i \alpha_i v_i$ ，则

$$R_A(x) = \frac{\sum_i \lambda_i \alpha_i^2}{\sum_i \alpha_i^2}.$$

- ▶ Rayleigh 商是特征值的**加权平均**，其中权重 $\alpha_i^2 / \sum_j \alpha_j^2$ 非负且和为 1。
- ▶ 因此 $\lambda_{\min}(A) \leq R_A(x) \leq \lambda_{\max}(A)$ 。
- ▶ 取相应的特征向量时，可以达到等号。

另外，令 $r = Ax - R_A(x)x$ ，则 $x^\top r = 0$ 。因此 $R_A(x)x$ 是 Ax 在 x 方向上的正交投影。

Courant-Fischer 定理：最大与最小特征值

设 $A = A^\top$ ，并按 $\lambda_1 \leq \dots \leq \lambda_n$ 排序。

$$\lambda_n = \max_{x \neq 0} R_A(x), \quad \lambda_1 = \min_{x \neq 0} R_A(x).$$

证明：用标准正交特征向量展开 $x = \sum_i \alpha_i v_i$ ：

$$\lambda_1 \sum_i \alpha_i^2 \leq \sum_i \lambda_i \alpha_i^2 \leq \lambda_n \sum_i \alpha_i^2.$$

除以 $\sum_i \alpha_i^2 > 0$ 即得不等式；取 $x = v_1$ 或 $x = v_n$ 可达等号。

等价地，也可以写成：

$$\lambda_1 = \min_{x^\top x = 1} x^\top A x.$$

即在 $x^\top x = 1$ 的约束下，求二次型 $x^\top A x$ 的最小值。

Courant-Fischer 定理：一般形式

把 $\lambda_1 \leq \dots \leq \lambda_n$ 对应的单位特征向量记为 v_i 。

$$\lambda_k = \min_{\substack{x \neq 0 \\ x \perp v_1, \dots, v_{k-1}}} R_A(x) = \max_{\substack{x \neq 0 \\ x \perp v_{k+1}, \dots, v_n}} R_A(x).$$

还可以写成不依赖于特征向量选取的形式：

$$\lambda_k = \min_{\substack{U \leq \mathbb{R}^n \\ \dim U = k}} \max_{x \in U \setminus \{0\}} R_A(x).$$

$$\lambda_k = \max_{\substack{U \leq \mathbb{R}^n \\ \dim U = n-k+1}} \min_{x \in U \setminus \{0\}} R_A(x).$$

固定任意 k 维子空间 U 。由于

$$\dim U + \dim \operatorname{span}(v_k, \dots, v_n) = k + (n - k + 1) > n,$$

两个子空间必有非零交向量 x 。它只能由 $v_k, \dots, v_n$ 线性组合得到，因此 $R_A(x) \geq \lambda_k$ 。于是

$$\max_{y \in U \setminus \{0\}} R_A(y) \geq \lambda_k.$$

另一方面，取 $U_* = \operatorname{span}(v_1, \dots, v_k)$ ，则

$$\max_{y \in U_* \setminus \{0\}} R_A(y) = \lambda_k.$$

因此，对任意 k 维子空间都有上述下界；取 U_* 时恰好等号成立。

谱范数、条件数与正规方程

对任意实矩阵 A ，由 Rayleigh 商的极值刻画：

$$\|A\|_2^2 = \max_{\|x\|_2=1} x^\top A^\top A x = \lambda_{\max}(A^\top A).$$

若 A 可逆，同理有

$$\min_{\|x\|_2=1} \|Ax\|_2^2 = \lambda_{\min}(A^\top A) = \|A^{-1}\|_2^{-2}.$$

因此，

$$\kappa_2(A^\top A) = \frac{\lambda_{\max}(A^\top A)}{\lambda_{\min}(A^\top A)} = \kappa_2(A)^2.$$

特别地，若 $A \succ 0$ ，因 $A^\top A = A^2$ ，所以

$$\|A\|_2 = \lambda_{\max}(A), \quad \kappa_2(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}.$$

矩阵多项式

若 $Av = \lambda v$, 则 $A^k v = \lambda^k v$ 。给定多项式 $p(t) = c_0 + c_1 t + \cdots + c_m t^m$, 定义

$$p(A) = c_0 I + c_1 A + \cdots + c_m A^m.$$

立刻得到

$$p(A)v = p(\lambda)v.$$

对实对称矩阵, $A = VDV^\top$, 因此

$$p(A) = Vp(D)V^\top = \sum_i p(\lambda_i) v_i v_i^\top.$$

与第三讲的联系：我们已经用多项式近似函数。这里则是把多项式作用于矩阵：每个特征值 λ_i 变成 $p(\lambda_i)$ 。

矩阵多项式对向量的作用

将 x 沿标准正交特征向量展开： $x = \sum_i \alpha_i v_i$ 。

$$Ax = \sum_i \alpha_i \lambda_i v_i,$$

$$A^2x = \sum_i \alpha_i \lambda_i^2 v_i,$$

$$A^kx = \sum_i \alpha_i \lambda_i^k v_i,$$

$$p(A)x = \sum_i \alpha_i p(\lambda_i) v_i.$$

观察：不同特征向量方向分别乘上 $p(\lambda_i)$ ，而不必直接计算 $p(A)$ 的所有元素。

后面还会用到：计算 A^kx 相当于取 $p(t) = t^k$ 。第九讲的共轭梯度法也可以从矩阵多项式的角度理解。

为什么不直接求特征多项式的根？

理论上可以求 $\det(tI - A)$ 的零点。为什么数值计算通常不这样做？

- ▶ **系数扰动的影响可能很大：**如 Wilkinson 多项式 $\prod_{j=1}^{20} (t - j)$ ，系数的小扰动可能使根显著偏移。
- ▶ **计算量：**用牛顿法需要反复计算特征多项式及其导数；即使用割线法，也需要反复计算行列式。
- ▶ **注意：** $|\det(\mu I - A)|$ 很小，**不一定**说明 μ 靠近特征值。

例： $A = 0_{n \times n}$ ，取 $\mu = \frac{1}{2}$ 。则 $|\det(\mu I - A)| = 2^{-n} \rightarrow 0$ ，但 μ 与任何特征值的距离始终为 $1/2$ 。

下面介绍另一种办法： 只做矩阵与向量的乘法，逐步逼近特征向量。

幂迭代 (Power Method)

假设 $A \in \mathbb{R}^{n \times n}$ 可在实数域对角化，且有一个特征值的绝对值严格大于其余特征值的绝对值：

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|, \quad x^{(0)} = \sum_{i=1}^n \alpha_i v_i, \quad \alpha_1 \neq 0.$$

连续进行矩阵与向量的乘法，得到

$$A^k x^{(0)} = \lambda_1^k \left(\alpha_1 v_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right).$$

因此，当 $k \rightarrow \infty$ 时，括号中其余各项趋于零。 $A^k x^{(0)}$ 的**方向**趋于 v_1 所在的直线。

特别地，若 $\lambda_1 = 1$ ，则 $A^k x^{(0)} \rightarrow \alpha_1 v_1$ 。如果 $\lambda_1 \neq 1$ ，则未经归一化的向量一般不收敛到 $\alpha_1 v_1$ 。

注意：必须按特征值的绝对值比较大小；负特征值也可能对应最大的绝对值。

归一化幂迭代

直接计算 $A^k x^{(0)}$ 可能出现以下问题：

- ▶ $|\lambda_1| > 1$ 时，向量长度可能指数增长，导致上溢。
- ▶ $|\lambda_1| < 1$ 时，向量长度可能指数缩小，导致下溢。

归一化幂迭代：

$$w^{(k+1)} = Ax^{(k)}, \quad x^{(k+1)} = \frac{w^{(k+1)}}{\|w^{(k+1)}\|_2}.$$

每一步把向量归一化，避免其长度过大或过小。若 $\lambda_1 < 0$ ，向量可能交替变号，但方向仍趋向于占优特征向量所在的直线。

注意：如果 $w^{(k+1)} = 0$ ，就不能继续归一化，应停止迭代。

思考：能否用其他范数归一化？为什么不直接除以 λ_1^k ？

幂迭代的收敛速度

当 A 实对称且 $|\lambda_1| > |\lambda_2|$, $\alpha_1 \neq 0$ 时, 其余特征向量分量相对主导分量的衰减速度由

$$\left| \frac{\lambda_2}{\lambda_1} \right|^k$$

决定 (常数与初值有关)。 $|\lambda_2/\lambda_1|$ **越接近 1, 收敛越慢。**

例 1:

$A = \text{diag}(4, 2, 1)$, 且 $x^{(0)} = (1, 1, 1)^\top$ 。
次要成分的相对大小为 $2^{-k}, 4^{-k}$ 。

例 2:

$A = \text{diag}(2, -2)$: 绝对值并列, 可能振荡。
若 $\alpha_1 = 0$, 迭代也永远得不到 v_1 的成分。

由近似特征向量估计特征值

假设已经得到近似特征向量 x 。如何选择 μ ，使 Ax 尽可能接近 μx ？

$$\min_{\mu \in \mathbb{R}} \|Ax - \mu x\|_2^2.$$

展开目标函数：

$$\|Ax - \mu x\|_2^2 = \|Ax\|_2^2 - 2\mu x^\top Ax + \mu^2 \|x\|_2^2.$$

令导数为零，或者写出一维正规方程 $x^\top (Ax - \mu x) = 0$ ，得

$$\mu_* = \frac{x^\top Ax}{x^\top x} = R_A(x).$$

这仍然是第四讲的最小二乘：在所有 μx 中，寻找与 Ax 距离最近的一个。

课堂练习（原讲义）：不求导，改用扰动法证明上述最优性。

$$\|Ax - (\mu_* + \delta)x\|_2^2 \geq \|Ax - \mu_*x\|_2^2 \quad (\forall \delta \in \mathbb{R}).$$

利用残差估计特征值误差

对实对称矩阵 A ，取 $\|x\|_2 = 1$ 、 $\mu \in \mathbb{R}$ ，定义残差 $r = Ax - \mu x$ 。

定理：存在 A 的一个特征值 λ_j ，使得

$$|\lambda_j - \mu| \leq \|r\|_2.$$

证明：展开 $x = \sum_i \alpha_i v_i$ ，其中 $\sum_i \alpha_i^2 = 1$ ，

$$\begin{aligned}\|r\|_2^2 &= \left\| \sum_i \alpha_i (\lambda_i - \mu) v_i \right\|_2^2 \\ &= \sum_i \alpha_i^2 (\lambda_i - \mu)^2 \\ &\geq \min_j |\lambda_j - \mu|^2 \sum_i \alpha_i^2 = \min_j |\lambda_j - \mu|^2.\end{aligned}$$

因此，将 $\mu = R_A(x)$ 代入上式，便可用残差范数估计它到**某个**特征值的距离。但不能仅凭残差判断该特征值是否为绝对值最大的特征值。

本讲小结

1. **正定性与谱分解**: Cholesky 分解、Loewner 偏序及 $A = VDV^T$ 。
2. **特征值的估计**: Gershgorin 圆盘定理和 Courant–Fischer 定理。
3. **幂迭代**: 矩阵多项式的性质以及幂迭代的收敛条件和速度。
4. **近似特征值**: 用 Rayleigh 商估计特征值，并利用残差控制误差。

下一讲: 如果要求的不是绝对值最大的特征值，该怎样修改幂迭代？又怎样从矩阵中找出低维的近似表示？

逆幂迭代、位移、Rayleigh 商迭代、奇异值分解 (SVD)。

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix}, \quad R = \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}, \quad A = R^{\top} R.$$

比较矩阵元素可得

$$r_{11} = 2, \quad r_{12} = 1, \quad r_{22} = 2,$$

因此

$$\boxed{\begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.}$$

提问：将右下角 5 改成 1，分解过程中会发生什么？

对角矩阵 $A = \text{diag}(4, 3.9, 1)$ 和 $B = \text{diag}(4, 2, 1)$ ，初值取 $x^{(0)} = (1, 1, 1)^\top$ 。

A 的次大比例：

$$\left| \frac{3.9}{4} \right| = 0.975.$$

需要较多次迭代，才能使其他方向的分量充分衰减。

B 的次大比例：

$$\left| \frac{2}{4} \right| = 0.5.$$

每次迭代后，第二个分量与第一个分量的比值减半。

讨论：第六讲用谱半径判断迭代是否收敛；这里为什么还需要比较两个最大特征值的绝对值？

前面已经给出了以下公式（证明放在课堂机动部分）：

$$\lambda_k = \min_{\dim U=k} \max_{x \in U \setminus \{0\}} R_A(x).$$

类似地，对任意维数 $n - k + 1$ 的子空间 W ，维数计数表明

$$W \cap \operatorname{span}(v_1, \dots, v_k) \neq \{0\}.$$

于是 $\min_{x \in W \setminus \{0\}} R_A(x) \leq \lambda_k$ 。取 $W_* = \operatorname{span}(v_k, \dots, v_n)$ 时取等，故

$$\lambda_k = \max_{\dim W=n-k+1} \min_{x \in W \setminus \{0\}} R_A(x).$$

若 $Av_i = \lambda_i v_i$ ，则 $p(A)v_i = p(\lambda_i)v_i$ 。但不同的特征值可能对应同一个 $p(\lambda)$ ：例如

$$A = \text{diag}(1, -1), \quad p(t) = t^2, \quad p(A) = I.$$

矩阵 A 有两个不同的特征值；经过 p 映射，两者都变成 1。

因此，原来的特征向量仍是 $p(A)$ 的特征向量，但当 $p(\lambda_i) = p(\lambda_j)$ 时， $p(A)$ 的对应特征空间可能更大。

对实对称矩阵，小残差保证近似特征值靠近真实特征值。对一般非正规矩阵则未必。例如

$$A = \begin{pmatrix} 0 & M \\ 0 & 0 \end{pmatrix}, \quad M \gg 1.$$

唯一特征值为 0。取 $x = (1, \varepsilon)^\top / \sqrt{1 + \varepsilon^2}$ ，其 **Rayleigh 商** 为

$$\mu = R_A(x) = \frac{M\varepsilon}{1 + \varepsilon^2}, \quad \|Ax - \mu x\|_2 = \frac{M\varepsilon^2}{1 + \varepsilon^2}.$$

选择 $\varepsilon = M^{-1}$ ，则 $\mu = \frac{M^2}{M^2+1} \rightarrow 1$ ，而残差范数 $\frac{M}{M^2+1} \rightarrow 0$ ；可是谱始终只有 0。 **结论：** 对于非正规矩阵，即使 μ 取 Rayleigh 商，残差很小也不一定意味着 μ 接近某个特征值。