

计算方法

Computational Methods

Lecture 9: 逆迭代、同时迭代与 QR 算法

Inverse Iteration, Orthogonal Iteration, and the QR Algorithm

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

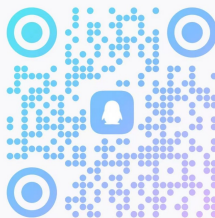
教材: Timothy Sauer, 《数值分析》(原书第 2 版)。

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：谱分解、SVD 与幂迭代

第七讲：对实对称矩阵 B ，存在正交特征向量基 v_i ，满足

$$B = V\Lambda V^\top, \quad Bv_i = \lambda_i v_i.$$

第八讲：对任意实矩阵 $A \in \mathbb{R}^{m \times n}$ ，利用 $A^\top A$ 的谱分解得到

$$A = U\Sigma V^\top, \quad A^\top A v_i = \sigma_i^2 v_i \quad (1 \leq i \leq r),$$

其中 $r = \text{rank}(A)$ ；宽矩阵的 $A^\top A$ 还会有额外的零特征值。

存在性不等于计算方法：上面的分解解释了奇异值和最优低秩逼近，却没有告诉我们如何高效求出特征值及相应特征向量。

第六讲：Jacobi 和 Gauss-Seidel 将 $Ax = b$ 改写为不动点迭代。**第七讲：**虽然 $Ax = \lambda x$ 中 λ, x 都未知，但幂迭代仍通过反复乘矩阵得到占优特征方向。

今天的问题：如何寻找其它特征值，甚至同时求出多个特征值？

本讲的三个问题

已知幂迭代能够找到模最大的特征值及相应特征向量。接下来要解决：

1. 如何找到**指定位置附近**的特征值？
2. 如何得到**多个**特征向量，而不只是一个？
3. 能否利用第四讲的 **QR 分解**，计算整个矩阵的特征值？

计算方法：逆迭代与位移 $\rightarrow$ 投影与同时迭代 $\rightarrow$ QR 迭代。

先讨论实对称矩阵；逆迭代的部分结论也适用于更一般的可对角化矩阵。

逆矩阵会怎样改变特征值？

设 A 可逆，且 $Av_i = \lambda_i v_i$ 。则

$$A^{-1}v_i = \frac{1}{\lambda_i}v_i.$$

对 A 做幂迭代

$$|\lambda_1| > |\lambda_2| > \dots$$

得到 λ_1 对应的特征向量。

对 A^{-1} 做幂迭代

$$\frac{1}{|\lambda_n|} > \frac{1}{|\lambda_{n-1}|} > \dots$$

得到模最小的非零特征值所对应的特征向量。

问题：如果想找靠近某个给定数 q 的特征值呢？

位移后的逆迭代 (Shifted Inverse Iteration)

设 $q \notin \{\lambda_1, \dots, \lambda_n\}$ 。由 $Av_i = \lambda_i v_i$,

$$(A - qI)^{-1}v_i = \frac{1}{\lambda_i - q}v_i.$$

因此, **离 q 最近的特征值**成为 $(A - qI)^{-1}$ 的模最大特征值。

算法: 给定 q 和单位初值 $x^{(0)}$, 迭代

$$(A - qI)y^{(k+1)} = x^{(k)}, \quad x^{(k+1)} = \frac{y^{(k+1)}}{\|y^{(k+1)}\|_2}.$$

注意: 每一步是解一个线性方程组, **不必显式计算** $(A - qI)^{-1}$ 。

例子：位移改变了哪个方向占优

$$A = \text{diag}(1, 3, 8), \quad x^{(0)} = (1, 1, 1)^\top / \sqrt{3}.$$



直接幂迭代:

$|\lambda_*| = 8$, 趋向于 e_3 。

误差的渐近衰减比为 $3/8$ 。

位移逆迭代:

$$\frac{1}{1-q} = -\frac{5}{9}, \quad \frac{1}{3-q} = 5, \quad \frac{1}{8-q} = \frac{5}{26}.$$

初值含 e_2 分量, 迭代趋向于 e_2 。

误差的渐近衰减比:

$$\max \left\{ \frac{|3-q|}{|1-q|}, \frac{|3-q|}{|8-q|} \right\} = \frac{1}{9}.$$

逆迭代的收敛速度

设 A 为实对称矩阵, λ_j 是**唯一**最接近 q 的特征值, 且初值在 v_j 方向的分量非零。

$$x^{(0)} = \alpha_j v_j + \sum_{i \neq j} \alpha_i v_i.$$

对 $B = (A - qI)^{-1}$ 做 k 次幂迭代:

$$B^k x^{(0)} = (\lambda_j - q)^{-k} \left(\alpha_j v_j + \sum_{i \neq j} \alpha_i \left(\frac{\lambda_j - q}{\lambda_i - q} \right)^k v_i \right).$$

非目标方向的统一衰减界:

$$\gamma(q) = \max_{i \neq j} \frac{|\lambda_j - q|}{|\lambda_i - q|} < 1.$$

注意: 实际衰减可更快 (某些 $\alpha_i = 0$); 归一化向量也可能交替变号, 收敛的是**特征方向**。

结论: 位移接近目标特征值通常加快收敛, 但线性系统也更难解。

如何高效实现迭代？

每一步需要解 $(A - qI)y = x$ 。选择求解器时要看矩阵大小与结构。

稠密矩阵、固定 q

1. 先对 $A - qI$ 做一次带选主元的 LU 分解；
2. 每一步只做前代和回代；
3. 分解代价约 $O(n^3)$ ，每次迭代约 $O(n^2)$ 。

大规模稀疏矩阵

- ▶ 可利用稀疏直接分解；
- ▶ 也可用合适的迭代线性求解器；
- ▶ 不能仅用稠密矩阵的 $O(n^3)$ 估计实际代价。

与第五、六讲的联系：矩阵求逆通常不是计算的第一选择；问题实际要求的是解线性方程组。

如何选择位移 q ?

已知近似单位特征向量 x ，第七讲证明过：使 $\|Ax - \mu x\|_2$ 最小的 μ 是

$$\mu = R_A(x) = x^\top Ax.$$

- ▶ 如果已知目标特征值的大致位置，可选择一个**固定的** q 。
- ▶ 如果没有足够准确的估计，可以用当前 x 的 Rayleigh 商**更新位移**。

由此得到 Rayleigh 商迭代： 每一步先根据当前向量估计特征值，再用这个估计做位移迭代。

Rayleigh 商迭代 (Rayleigh Quotient Iteration)

设 $A = A^\top$, 从单位向量 $x^{(0)}$ 出发:

1. 计算当前近似特征值 $\mu_k = (x^{(k)})^\top A x^{(k)}$;
2. 解线性方程 $(A - \mu_k I)y^{(k+1)} = x^{(k)}$;
3. 归一化 $x^{(k+1)} = y^{(k+1)} / \|y^{(k+1)}\|_2$ 。

固定移位迭代:

系数矩阵不变, 可复用其分解。

停机与例外: 先检查残差 $\|Ax^{(k)} - \mu_k x^{(k)}\|_2$ 。

若残差为零则停止; 若位移矩阵奇异但残差不为零, 则需调整位移或另行处理。

Rayleigh 商迭代:

位移不断变化, 每一步通常都要重新分解。

位移越准确，线性方程越难解？

对实对称且 q 不是特征值的矩阵，第八讲 SVD 的条件数公式给出

$$\kappa_2(A - qI) = \frac{\max_i |\lambda_i - q|}{\min_i |\lambda_i - q|}.$$

q 越靠近目标特征值，逆迭代通常越快，但线性系统也越接近奇异。

一个可以算清楚的例子：取 $A = \text{diag}(1, 3)$ 、 $q = 3 - \varepsilon$ ，并令 $x = (1, 1)^\top / \sqrt{2}$ 。

$$(A - qI)y = x \quad \implies \quad y = \frac{1}{\sqrt{2}} \begin{pmatrix} (-2 + \varepsilon)^{-1} \\ \varepsilon^{-1} \end{pmatrix}.$$

当 $\varepsilon \rightarrow 0^+$ 时， $\|y\|_2 \rightarrow \infty$ ，但 $\frac{y}{\|y\|_2} \rightarrow e_2$ 。

结论：求解线性系统时的病态性，并不意味着归一化后的方向必然不准确；但仍须控制求解误差，并检查特征对的残差。若位移恰好等于特征值，还须处理奇异情形。

Rayleigh 商迭代为什么快?

假设 A 实对称, 目标是简单特征值 λ_j , 且单位向量 x 已非常接近相应特征向量 v_j 。写作

$$x = \cos \theta v_j + \sin \theta w, \quad w \perp v_j, \quad \|w\|_2 = 1.$$

由 Rayleigh 商的展开可得

$$R_A(x) - \lambda_j = \sin^2 \theta (w^\top A w - \lambda_j) = O(\theta^2).$$

关键：当特征向量的方向误差是 $O(\theta)$ 时, Rayleigh 商的误差是 $O(\theta^2)$ 。

进一步的结论：在实对称、目标特征值简单且初值充分接近的条件下, Rayleigh 商迭代具有局部**三次收敛**。课堂机动的 2×2 例子可直接算出三次关系；一般证明从略。

如果需要第二个特征向量呢？

设 $A = A^\top$ 的特征向量为标准正交基 $v_1, \dots, v_n$ ，且 $|\lambda_1| > |\lambda_2| > \dots$ 。对初值写成

$$x^{(0)} = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n.$$

如果能使 $\alpha_1 = 0$ ，幂迭代就不会沿 v_1 的方向放大。

投影：已知 v_1 时，把每一步的向量替换成

$$x \leftarrow (I - v_1 v_1^\top)x.$$

便能消去沿 v_1 的分量。

为什么可行？ 实对称矩阵的特征向量相互正交， $v_1^\perp$ 在 A 下保持不变；仍需在剩余特征值之间有模间隙。

用正交投影逐个求特征向量

若已求得标准正交的 $v_1, \dots, v_{j-1}$ ，设

$$P_{j-1} = I - \sum_{i=1}^{j-1} v_i v_i^\top.$$

从初值中去掉这些方向，再做幂迭代：

$$y^{(k+1)} = P_{j-1} A x^{(k)}, \quad x^{(k+1)} = \frac{y^{(k+1)}}{\|y^{(k+1)}\|_2}.$$

- ▶ 当目标特征值与其他特征值的模存在间隙时，可找到下一个方向。
- ▶ 对浮点计算，必须警惕向已消去的方向重新产生分量。

问题：与其逐个求出并排除特征向量，能否**同时**维护一组正交向量？

同时迭代：同时更新多个向量

把 r 个初始向量排成矩阵

$$Q_0 = [q_1^{(0)}, \dots, q_r^{(0)}], \quad Q_0^\top Q_0 = I_r.$$

每一步先乘 A ，再作正交化：

$$AQ_k = Q_{k+1}R_k, \quad Q_{k+1}^\top Q_{k+1} = I_r.$$

当 AQ_k 满列秩时，这是一次（薄）QR 分解。

与第 4 讲的联系：QR 分解就是把线性无关的列向量正交化。当 $r = 1$ 时，这就是归一化幂迭代； $r = n$ 时，在 AQ_k 满秩的条件下得到一个完整正交基。

为什么每步都要正交化？

假设 A 的特征值依模严格排序： $|\lambda_1| > \dots > |\lambda_n|$ 。

- ▶ 若直接对每一列反复乘 A ，这些列一般会**全部靠近**模最大特征值的方向。
- ▶ 通过 QR 分解保持两两正交，才能用不同列捕捉**不同的特征方向**。

几何解释：幂迭代放大某些方向；QR 正交化不改变 AQ_k 的列空间（假设 AQ_k 满列秩）。

$$\text{col}(AQ_k) = \text{col}(Q_{k+1}).$$

严格结论所需条件：目标特征子空间与其余特征值存在模间隙，初始子空间在相应方向上的投影非退化。

同时迭代怎样得到特征值？

同时迭代得到 r 个正交方向 $Q = [q_1, \dots, q_r]$ 。定义较小的矩阵

$$T := Q^\top A Q \in \mathbb{R}^{r \times r}.$$

对角元：

$$T_{ii} = q_i^\top A q_i$$

即 q_i 的 Rayleigh 商。

非对角元：

$$T_{ij} = q_i^\top A q_j$$

若各 q_i 是特征向量，则非对角元为 0。

为什么有用？ 若 $\text{col}(Q)$ 是 A 的不变子空间，则 $AQ = QT$ ； T 的特征值就是 A 在该子空间上的特征值。

当 $r = n$ 时： Q 是正交方阵， $T = Q^\top A Q$ 表示同一线性变换在另一组正交基下的矩阵。下面正式定义这种换基关系。

子空间还未收敛，特征值估计可靠吗？

同时迭代通常只得到**近似**不变子空间。设 $A = A^\top$, $Q^\top Q = I_r$, 并计算小矩阵 $T = Q^\top A Q$ 的特征对

$$Tz = \theta z, \quad \|z\|_2 = 1, \quad u := Qz.$$

则 $\|u\|_2 = 1$, 其在原矩阵上的**残差**为

$$Au - \theta u = (I - QQ^\top)AQz.$$

可验证的误差保证：记 A 的特征值为 $\lambda_1, \dots, \lambda_n$ 。由第七讲的谱分解，

$$\min_{1 \leq i \leq n} |\theta - \lambda_i| \leq \|Au - \theta u\|_2 \leq \|(I - QQ^\top)AQ\|_2.$$

意义：不必先得到精确不变子空间；残差小时， θ 就接近 A 的某个真特征值。

相似变换 (Similarity Transformation)

若 T 可逆, 且 $B = T^{-1}AT$, 称 A 与 B **相似**。

相似矩阵具有相同的特征值:

$$B(T^{-1}v) = T^{-1}Av = \lambda T^{-1}v \quad (Av = \lambda v).$$

特别地, 若 Q 是正交矩阵, 则

$$B = Q^{\top} A Q.$$

这是一次**正交相似变换**。当 $A = A^{\top}$ 时, B 也对称。**目标:** 在不改变特征值的前提下, 逐步把矩阵变为更容易读出特征值的形式。

QR 分解产生一次相似变换

设 A_k 是实方阵，对其做 QR 分解

$$A_k = Q_k R_k, \quad Q_k^\top Q_k = I.$$

交换两个因子的顺序：

$$A_{k+1} = R_k Q_k = Q_k^\top A_k Q_k.$$

两个重要性质：

1. A_{k+1} 与 A_k 相似，因此具有相同的特征值；
2. 若 A_k 对称，则 A_{k+1} 仍然对称。

与第四讲的联系： 当时 QR 分解用于解最小二乘；现在同一分解可用于特征值计算。

QR 迭代算法 (QR Algorithm)

输入： 实对称矩阵 A 。

1. 令 $A_0 = A$;

2. 对 $k = 0, 1, 2, \dots$:

$$A_k = Q_k R_k, \quad A_{k+1} = R_k Q_k.$$

3. 若迭代矩阵的非对角部分足够小，则从对角线上读出近似特征值。

每一步都保持：

$$A_k = Z_k^\top A Z_k, \quad Z_k = Q_0 Q_1 \cdots Q_{k-1}.$$

注意： 上述讨论实对称矩阵；即使如此，**不带位移的基本 QR 迭代**也不一定对角收敛（见补充反例）。

为什么非对角元小，就能读出特征值？

对实对称矩阵 A_k ，记

$$A_k = D_k + E_k, \quad D_k = \text{diag}((A_k)_{11}, \dots, (A_k)_{nn}), \quad E_k = \text{非对角部分}.$$

任取单位向量 x ，有

$$|x^\top A_k x - x^\top D_k x| = |x^\top E_k x| \leq \|E_k\|_2 \leq \|E_k\|_F.$$

因此，由第七讲的 Courant–Fischer 定理，若将 A_k 的特征值与 D_k 的对角元分别按**降序**排列，便得到

$$\max_i |\lambda_i(A) - d_{(i)}| \leq \|E_k\|_2 \leq \|E_k\|_F.$$

停机依据：基本 QR 迭代保持特征值不变；当非对角部分足够小时，对角元就是可控制误差的特征值近似。

一个 2×2 的 QR 迭代例子

$$A_0 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \underbrace{\frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}}_{Q_0} \underbrace{\frac{1}{\sqrt{5}} \begin{pmatrix} 5 & 4 \\ 0 & 3 \end{pmatrix}}_{R_0}.$$

$$A_1 = R_0 Q_0 = \begin{pmatrix} 14/5 & 3/5 \\ 3/5 & 6/5 \end{pmatrix}.$$

迭代前： 对角元 2, 2；非对角元 1。

迭代后： 对角元 2.8, 1.2；非对角元 0.6。

检查： $\text{tr}(A_k) = 4$ 、 $\det(A_k) = 3$ 始终不变，特征值仍为 3, 1。

QR 迭代与同时迭代的关系

同时迭代从 $Z_0 = I$ 开始，每一步对 AZ_k 做 QR 分解，得到下一组正交基 Z_{k+1} 。基本 QR 迭代产生 $A_k = Q_k R_k$ 及累积变换

$$Z_k = Q_0 Q_1 \cdots Q_{k-1}, \quad A_k = Z_k^\top A Z_k.$$

联系：取相同的 QR 因子符号约定，累积矩阵 Z_k 就是同时迭代得到的正交基：

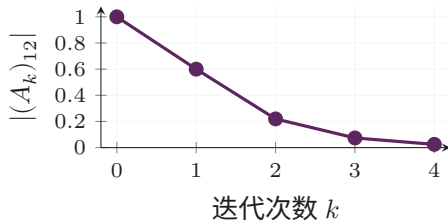
$$AZ_k = Z_k A_k = Z_k Q_k R_k = Z_{k+1} R_k.$$

同一种算法的两个视角：同时迭代更新一组向量；QR 迭代更新矩阵在这组基下的表示。

理解 QR 迭代的极限

每一步都有 $A_k = Z_k^\top A Z_k$ ：**特征值始终不变**，但矩阵的表示发生变化。

刚才的 2×2 例子： $A_0 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ 。



理想的极限：如果 $A_k \rightarrow D$ ，且 D 是对角矩阵，那么

$$D = \text{diag}(\lambda_{\pi(1)}, \dots, \lambda_{\pi(n)}).$$

此时对角元就是 A 的特征值（顺序可变）。

但不要倒推：正交相似只保证特征值不变，**不保证**基本 QR 迭代对任意矩阵都对角收敛。

实际算法：位移、三对角化及停机判断可以显著提高效率（参见补充材料）。

QR 迭代的定点与对易矩阵

如果 QR 迭代到达某个**定点** A_* ，即对 A_* 的分解满足

$$A_* = Q_* R_* = R_* Q_*,$$

那么

$$A_* Q_* = (Q_* R_*) Q_* = Q_* (R_* Q_*) = Q_* A_*.$$

观察：当 QR 迭代在上述意义下到达定点时， A_* 与其 QR 分解中的正交因子 Q_* **对易**。

接下来的问题：两个矩阵满足 $AB = BA$ ，能推出什么关于其特征向量的结论？

不能据此推出对角化：例如补充部分给出一个非对角矩阵，它本身就是基本 QR 迭代的定点。

对易矩阵与共同特征向量基

QR 迭代把矩阵转化到不同的正交基下。下面考虑一个相关的线性代数问题：

定理： 设 $A \in \mathbb{R}^{n \times n}$ 有 n 个**互不相同**的实特征值，且相应特征向量 $v_1, \dots, v_n$ 构成一组基。则对任意实方阵 B ，有

$$AB = BA \iff v_1, \dots, v_n \text{ 同时也是 } B \text{ 的特征向量。}$$

注意措辞： 结论是**存在一组共同特征向量基**，不是两矩阵的全部特征向量完全相同（例如 $B = I$ ）。

重特征值时： $AB = BA$ 仍保证 A 的每个特征空间在 B 下不变，但不保证其中**每个**向量都是 B 的特征向量。

证明：共同的特征向量 $\implies AB = BA$

假设 $v_1, \dots, v_n$ 是 A, B 共同的特征向量：

$$Av_i = \lambda_i v_i, \quad Bv_i = \mu_i v_i.$$

对每个 i ,

$$ABv_i = A(\mu_i v_i) = \mu_i \lambda_i v_i, \quad BAv_i = B(\lambda_i v_i) = \lambda_i \mu_i v_i.$$

因而 $(AB - BA)v_i = 0$ 。由于 $\{v_i\}_{i=1}^n$ 是 $\mathbb{R}^n$ 的一组基，

$$AB = BA.$$

这个方向不需要特征值互不相同，只需要共同的特征向量构成一组基。

证明: $AB = BA \implies$ 共同的特征向量

设 $Av_i = \lambda_i v_i$, 由 $AB = BA$,

$$A(Bv_i) = B(Av_i) = \lambda_i(Bv_i).$$

所以 Bv_i 仍属于 A 的 λ_i -特征空间。由于 A 的特征值互不相同, 该特征空间为一维:

$$Bv_i = \mu_i v_i \quad (\mu_i \in \mathbb{R}).$$

其中 $\mu_i = 0$ 的情况也包括在内。因此 v_i 同时是 B 的特征向量。

证明的关键: 与 A 对易的矩阵, 必须把 A 的每个特征空间映射到其自身; 当特征空间一维时, 这就意味着沿该方向只做数乘。

如何选择特征值算法？

根据**要找的特征值**、矩阵规模和可利用的计算操作选择方法。

目标	主要方法	计算上的关键差别
靠近给定 q 的特征值	位移迭代	反复求解 $(A - qI)y = x$ ；固定 q 时可复用分解
多个占优的特征方向	同时迭代	矩阵与多个向量相乘，每步 QR 正交化
实对称矩阵的全部特征值	QR 迭代	连续作正交相似变换；工程上常需位移和预处理

共同的工具：线性方程组求解、Rayleigh 商、正交投影、QR 分解及相似变换。

共同的问题：不变量是什么？在什么假设下收敛？每一步实际计算什么？

非对称矩阵也有 QR 算法，但其极限形式通常应理解为实 Schur 形（可能有 2×2 块），不是这里的对角矩阵。

本讲小结：从矩阵分解到迭代计算

1. **逆迭代与位移**：将离 q 最近的特征值变成占优特征值，迭代时只需解线性方程组。
2. **Rayleigh 商迭代**：根据当前向量调整位移，局部收敛速度更快。
3. **正交投影与同时迭代**：将幂迭代推广到多个特征方向，使用 QR 保持正交。
4. **QR 迭代**：通过正交相似变换改变矩阵的表示，逼近更容易读出特征值的形式。
5. **对易矩阵**：当 A 的特征值两两不同时， $AB = BA$ 与共同特征向量基之间的关系。

下一讲：不再尝试求完整特征分解，而是用矩阵-向量乘积高效求解 $Ax = b$ ：
Richardson 迭代、梯度下降与 Krylov 方法。

三次收敛：一个能算出精确关系的例子

考虑 $A = \text{diag}(1, 3)$ ，目标特征向量为 e_1 ，取

$$x = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \quad 0 < \theta < \pi/4.$$

Rayleigh 商为 $\mu = 1 + 2 \sin^2 \theta$ 。解 $(A - \mu I)y = x$ 得

$$y = \frac{1}{2} \begin{pmatrix} -\cos \theta / \sin^2 \theta \\ \sin \theta / \cos^2 \theta \end{pmatrix}, \quad \frac{y_2}{y_1} = -\tan^3 \theta.$$

精确的角度关系：归一化后，若 θ_{new} 表示与特征方向的锐角，则

$$\tan \theta_{\text{new}} = \tan^3 \theta.$$

这个例子说明**三次收敛从哪里来**。一般对称矩阵的局部结论仍需利用谱间隙。

若 $v_1, \dots, v_r$ 是已知的标准正交特征向量，记

$$V_r = [v_1, \dots, v_r], \quad P = I - V_r V_r^\top.$$

性质：

$$\begin{aligned} P^\top &= P, & P^2 &= P, \\ PV_r &= 0, & PA &= AP \quad (A = A^\top). \end{aligned}$$

于是 PAP 在已知特征方向上为零，在其正交补上与 A 具有同样的作用。**提醒：**若未知的下一个特征值是 0，投影后出现的零特征值需要区别对待。

为加快 QR 算法，常在第 k 步选择位移 μ_k ：

$$A_k - \mu_k I = Q_k R_k, \quad A_{k+1} = R_k Q_k + \mu_k I.$$

由 $A_k = Q_k R_k + \mu_k I$ 得

$$A_{k+1} = Q_k^\top A_k Q_k,$$

因此仍是**正交相似变换**，特征值不变。**实际求解**：对称矩阵通常先化成三对角形式，再进行带位移 QR 迭代，减少每次迭代的计算量。**本课程不要求**：Wilkinson 位移、隐式 QR 与完整工程实现的证明。

取

$$A = I_2, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

显然 $AB = BA$ 。但 A 的**每个非零向量**都是特征向量，

$$x = (1, 1)^\top$$

却**不是** B 的特征向量。

正确结论：若 $AB = BA$ ，则 B 保持 A 的各个特征空间不变；若这些特征空间是一维的，才可推出 A 的每个特征向量也都是 B 的特征向量。

看一个简单矩阵

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A^\top A = I.$$

在一个自然的 QR 选取中, $Q = A$ 、 $R = I$, 因此

$$A_{k+1} = RQ = A_k = A.$$

无论迭代多少次, 非对角元都没有变小。

结论: 相似变换保持特征值, 但不保证无位移 QR 迭代在所有输入上都收敛到对角矩阵。实际算法需要更精细的数值处理。