Numerical Algebra Chapter 6

@rosefantasie

2022 年 12 月 1 日

1. 设 A 是 $n\times m$ 矩阵，B 是 $m\times n$ 矩阵，且 $m\geq n$. 证明：

$$\lambda(BA)=\lambda(AB)\cup\{0,\cdots,0\}.(m-n\text{个}0)$$

解.

$$\begin{aligned}\det(\lambda I_m-BA)&=\det\begin{pmatrix}I_n & A\\ & \lambda I_m-BA\end{pmatrix}=\det\begin{pmatrix}I_n & B\\ & \lambda I_m\end{pmatrix}\\&=\det\begin{pmatrix}\lambda^{-1}(\lambda I_n-AB) & A\\ & \lambda I_m\end{pmatrix}=\lambda^{m-n}\det(\lambda I_n-AB)\end{aligned}$$

□

2. 设 $\lim\limits_{k\to\infty}A_k=A, A_k=Q_kT_kQ_k^*$ 是 A_k 的 Schur 分解。证明：$\{Q_k\}$ 有收敛的子序列 $\{Q_{k_i}\}$; 若记 $Q=\lim\limits_{i\to\infty}Q_{k_i}$，则 Q^*AQ 是上三角阵。

解. Q_k 是酉矩阵，故 Q_k 有界，由 B-W 定理知 $\{Q_k\}$ 有收敛子列，记作 $\{Q_{k_i}\}$. 于是 $Q^*AQ=\lim\limits_{i\to\infty}Q_{k_i}^*AQ_{k_i}=\lim\limits_{i\to\infty}T_{k_i}$ 是上三角阵。 □

3. 设 $A\in\mathbb{C}^{n\times n}$ 没有重特征值，$B\in\mathbb{C}^{n\times n}$ 满足 $AB=BA$. 证明：若 $A=QTQ^*$ 是 A 的 Schur 分解，则 Q^*BQ 是上三角阵。

解. 由 $AB=BA$ 得 $QTQ^*B=BQTQ^*, TQ^*BQ=Q^*BQT.(1)$
归纳证明：若 $TX=XT, T_{ii}\neq T_{jj}(i\neq j)$ 上三角, 则 X 是上三角阵。
$n=2$ 时可直接验证得到。
设对于 $n-1\geqslant 2$ 结论成立. 将 T,X 分块

$$T=\begin{pmatrix}T_{n-1} & \vec{\gamma}\\ & \lambda_n\end{pmatrix}, X=\begin{pmatrix}X_{n-1} & \vec{\alpha}\\ \vec{\beta}^T & x_n\end{pmatrix}$$

$$TX=XT\Leftrightarrow\begin{pmatrix}T_{n-1}X_{n-1}+\vec{\gamma}\vec{\beta}^T & T_{n-1}\vec{\alpha}+\vec{\gamma}x_n\\ \lambda_n\vec{\beta}^T & \lambda_nx_n\end{pmatrix}=\begin{pmatrix}X_{n-1}T_{n-1} & X_{n-1}\vec{\gamma}+\vec{\alpha}\lambda_n\\ \vec{\beta}^TT_{n-1} & \vec{\beta}^T\vec{\gamma}+\lambda_nx_n\end{pmatrix}$$

由 $\lambda_n\vec{\beta}^T=\vec{\beta}^TT_{n-1}$ 得 $(\lambda_nI-T_{n-1})^T\vec{\beta}=0$, $T_{ii}\neq T_{jj}(i\neq j)$ 故 $(\lambda_nI-T_{n-1})^T$ 可逆，因而 $\vec{\beta}$ 此时分块矩阵第二行得等号均自然满足。再由 $T_{n-1}X_{n-1}+\vec{\gamma}\vec{\beta}^T=T_{n-1}X_{n-1}=X_{n-1}T_{n-1}$ 及归纳假设知 X_{n-1} 是上三角阵。因而 X 是上三角阵。再由式 (1) 知 Q^*BQ 是上三角阵。 □

4. 设 $A \in \mathbb{C}^{n\times n}$. 对于给定的非零向量 $x \in \mathbb{C}^n$, 定义

$$R(x) = \frac{x^*Ax}{x^*x},$$

称之为 x 对 A 的 Rayleigh 商。证明：对任意的 $x \in \mathbb{C}^n (x \neq 0)$,

$$\|Ax - R(x)x\|_2 = \inf_{\mu\in\mathbb{C}} ||Ax - \mu x||_2,$$

即 Rayleigh 商有极小剩余性。

解.

$$\begin{aligned}\|Ax-\mu x\|_2^2 = (Ax-\mu x)^*(Ax-\mu x) &= x^*A^*Ax - \bar{\mu}x^*Ax - \mu x^*Ax + |\mu|^2x^*x \\ &= x^*A^*Ax - 2\mathrm{Re}(\mu)x^*Ax + |\mu|^2x^*x \\ &\geqslant x^*A^*Ax - |\mu|x^*Ax + |\mu|^2x^*x =: f(|\mu|)\end{aligned}$$

取等条件为 μ 是实数。考虑 $f(u)$,

$$\frac{\mathrm{d}f}{\mathrm{d}u} = 2(x^*xu - x^*Ax),$$

故 $f(u)$ 在 $u = \dfrac{x^*Ax}{x^*x} = R(x)$ 处取最小值。即 $\mu = |\mu| = R(x)$ 时取到 $\|Ax - R(x)x\|_2$ 的最小值。 □

5. 设 $A = \begin{pmatrix}\alpha & \gamma \\ 0 & \beta\end{pmatrix}, \alpha \neq \beta$. 求 A 的特征值 α 和 β 的条件数。

解. 参考 P164 计算过程。

$$\mathrm{cond}(\alpha) = \mathrm{cond}(\beta) = \frac{\sqrt{(\alpha-\beta)^2+\gamma^2}}{|\alpha-\beta|}$$

□

6. 证明：特征值和特征向量的条件数在酉相似变换下保持不变。

解. 设 λ 是 A 的特征值，x 是对应的单位特征向量，y 是 λ 的左特征向量，则有

$$y^TA = \lambda y^T, y^Tx = 1$$

设 Q 是酉方阵，下面计算 $\widetilde{A} = Q^*AQ$ 的属于特征值 $\tilde{\lambda}$ 的左特征向量 $\tilde{y}$.

由于 $\widetilde{A}Q^*x = Q^*Ax = \lambda Q^*x$, 故 $\tilde{x} = Q^*x$ 是 $\widetilde{A}$ 的特征向量。令 $\tilde{y} = Q^Ty$, 下证 $\tilde{y}$ 是 $\widetilde{A}$ 的左特征向量。

$\tilde{y}^T\widetilde{A} = y^TQQ^*AQ = y^TAQ = \lambda y^TQ = \lambda\tilde{y}^T$, 且 $\tilde{y}^T\tilde{x} = y^TQQ^*x = y^Tx = 1$.

故 $\mathrm{cond}(\tilde{\lambda}) = \|\tilde{y}\|_2 = \|y\|_2 = \mathrm{cond}(\lambda)$.

再看特征向量的条件数。

设

$$U = (x, U_2), U^*AU = \begin{pmatrix}\lambda & x^*AU_2 \\ & A_2\end{pmatrix}, \Sigma^\perp = U_2(\lambda I - A_2)^{-1}U_2^*.$$

$$\widetilde{U} = Q^*U, \widetilde{U}^*\widetilde{A}\widetilde{U} = U^*AU = \begin{pmatrix}\lambda & x^*AU_2 \\ & A_2\end{pmatrix}, \widetilde{\Sigma}^\perp = Q^*U_2(\lambda I - A_2)^{-1}U_2^*Q$$

从而 $\|\widetilde{\Sigma}^\perp\|_2 = \|Q^*\Sigma^\perp Q\|_2 = \|\Sigma^\perp\|_2$. □

7. 分别应用幂法于矩阵

$$A=\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix} \quad 和 \quad B=\begin{pmatrix}\lambda & 1\\ 0 & -\lambda\end{pmatrix}(\lambda\neq 0)$$

并考察所得序列的特性。

解. 取 $u_0=(1,1)$.

$$\begin{aligned}A:&u_k=\left(1,\frac{\lambda}{\lambda+k}\right)\to(1,0)\\ B:&u_1=\left(1,-\frac{\lambda}{\lambda+1}\right),u_2=(1,1)\text{ 不收敛}\end{aligned}$$

可取 $u_0=(1,0)$ 一步得到结果。 □

8. 在幂法中，取 $A=\begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}, u_0=(0,0,1)^T$. 得到一个精确到 5 位数字的特征向量需要多少次迭代?

解. 经过计算得

$$\begin{gathered}u_k=\left(1,\frac{2}{k-1},\frac{2}{k(k-1)}\right)^T,\hat{u}=(1,0,0)^T,\\ \|e_k\|_\infty=\frac{2}{k-1}\leqslant 10^{-5},k\geqslant 200001\end{gathered}$$

□

9. 设 $A\in\mathbb{C}^{n\times n}$ 有实特征值并满足 $\lambda_1>\lambda_2\geq\cdots\geq\lambda_{n-1}>\lambda_n$. 现应用幂法于矩阵 $A-\mu I$. 试证：选择 $\mu=(\lambda_2+\lambda_n)/2$ 时，所产生的向量序列收敛到属于 λ_1 的特征向量的速度最快。

解. 设 $\{\tilde{\lambda}_{(k)}\}$ 是 $A-\mu I$ 按绝对值从大到小排列得特征值。

只有取 $\mu\in[\lambda_n,\lambda_2]$ 时有 $|\lambda_1-\mu|=\max|\lambda_i-\mu|$, 即 $\tilde{\lambda}_{(1)}=\lambda_1-\mu, |\tilde{\lambda}_{(2)}|=\max(|\lambda_2-\mu|,|\lambda_n-\mu|)$. 由定理 6.2.1 的证明过程知当 $\dfrac{\tilde{\lambda}_{(2)}}{\tilde{\lambda}_{(1)}}$ 取到最小时收敛速度最快。而 $\min\max(|\lambda_2-\mu|,|\lambda_n-\mu|)=\dfrac{|\lambda_2-\lambda_n|}{2}$, 在 $\mu=(\lambda_2+\lambda_n)/2$ 时取到，故此时收敛速度最快。

□

10. 应用幂法给出求多项式

$$p(z)=z^n+\alpha_1 z^{n-1}+\cdots+\alpha_n$$

的模最大根的一种算法。

解. 对该多项式的友方阵应用幂法即可。 □

11. 利用反幂法计算矩阵

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

对应于近似特征值 $\tilde{\lambda} = 1.2679$(精确特征值是 $\lambda = 3 - \sqrt{3}$) 的近似特征向量。

解.

$$A = \begin{pmatrix} 2 & 1 & \\ 1 & 3 & 1 \\ 0 & 1 & 4 \end{pmatrix}, A^{-1} = \frac{1}{18}\begin{pmatrix} 11 & -4 & 1 \\ -4 & 8 & -2 \\ 1 & -2 & 5 \end{pmatrix}$$

直接应用反幂法计算，可以自己计算准确值来验算。 □

12. 设 $A \in \mathbb{C}^{n\times n}$，并假定 $\lambda \in \mathbb{C}$ 和 $u \in \mathbb{C}^n$ 已给定 $(u \neq 0)$, 且 λ 不是 A 的特征值。证明：可选择 $E \in \mathbb{C}^{n\times n}$ 满足

$$||E||_F = \frac{||u||_2}{||v||_2},$$

使得向量 $v = (\lambda I - A)^{-1}u$ 是 $A + E$ 的一个特征向量。

解. 取 $E = \dfrac{uv^T}{\|v\|_2^2}$, 则 $\|E\|_F = \dfrac{\|u\|_2}{\|v\|_2}$. 下证 v 是 $A + E$ 的特征向量。

$$(A+E)v - \lambda v = (A - \mu I)(\lambda I - A)^{-1}u + \frac{uv^Tv}{\|v\|_2^2} = (A - \lambda I)(\lambda I - A)^{-1}u + u = 0.$$

故 v 是 $A + E$ 的特征向量。 □

13. 设 $A, E \in \mathbb{C}^{n\times n}$，并假定 λ 是 $A + E$ 的特征值但不是 A 的特征值。证明：存在向量 $u, v \in \mathbb{C}^n$ 使得

$$v = (\lambda I - A)^{-1}u,\ 且 \frac{\|u\|_2}{\|v\|_2} \leq \|E\|_2.$$

解. 设 v 满足 $(A + E)v = \lambda v$, $u = (\lambda I - A)v = Ev$.

$$\frac{\|u\|_2}{\|v\|_2} = \frac{\|Ev\|_2}{\|v\|_2} \leqslant \|E\|_2.$$

□

14. 应用 QR 算法的基本迭代格式 (6.4.1) 于矩阵

$$\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

并考察所得矩阵序列的特点。它是否收敛？

解.

$$A_0 = A = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}, A_1 = \frac{1}{2}\begin{pmatrix} 1 & -3 \\ -1 & -1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

故该序列不收敛。 □

15. 设 $A \in \mathbb{R}^{n\times n}$. 证明: 存在初等置换矩阵 P 和初等下三角阵 M, 使得 $(MP)A(MP)^{-1}$ 具有如下形状:

$$(MP)A(MP)^{-1} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ 0 & \alpha_{32} & \cdots & \alpha_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}.$$

解. 若第一列从第二行开始全是 0 则自然成立。

若 $a_{21} \neq 0$, 取 $P = I$,

$$\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & -\frac{a_{31}}{a_{21}} & 1 & & \\ & \vdots & & \ddots & \\ & -\frac{a_{n1}}{a_{21}} & & & 1 \end{pmatrix}$$

否则存在 $i_0 \geqslant 3$, $a_{i_0 1} \neq 0$, 左乘排列方阵将第 i_0 行与第 2 行交换，即 P，然后重复上述操作即可。 □

16. 利用第 15 题的结果, 设计一个利用非正交变换将 A 上 Hessenberg 化的实用算法.

解. 对 A 做一次 15 题中操作后，对右下角的 $(n-1)\times(n-1)$ 的矩阵重复，依次往下即可得到上 Hessenberg 阵。 □

17. 设 $A \in \mathbb{C}^{n\times n}, x \in \mathbb{C}^n, X = (x, Ax, \cdots, A^{n-1}x)$. 证明: 如果 X 是非奇异的，则 $X^{-1}AX$ 是上 Hessenberg 矩阵。

解. X 非奇异，故 $x, Ax, \cdots, A^{n-1}x$ 构成向量空间的 1 组基，设 $A_n x = \sum\limits_{i=0}^{n-1} a_i A^i x$, 则

$$AX = X\begin{pmatrix} 0 & & & & a_0 \\ 1 & 0 & & & a_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & a_{n-2} \\ & & & 1 & a_{n-1} \end{pmatrix}, X^{-1}AX = \begin{pmatrix} 0 & & & & a_0 \\ 1 & 0 & & & a_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & a_{n-2} \\ & & & 1 & a_{n-1} \end{pmatrix}$$

是上 Hessenberg 阵。 □

18. 证明: 若 H 是一个非亏损的不可约上 Hessenberg 矩阵, 则 H 没有重特征值.

解. 假设 H 的特征值 λ 的代数重数 $\geqslant 2$. 由于 H 不可约，次对角元非零。

$$H-\lambda I=\begin{pmatrix} h_{11}-\lambda & h_{12} & \cdots & \cdots & h_{1n} \\ h_{21} & h_{22}-\lambda & \cdots & \cdots & h_{2n} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_{n,n-1} & h_{nn}-\lambda \end{pmatrix}$$

故左下角的 $(n-1)\times(n-1)$ 的矩阵是可逆的，即 $\operatorname{rank}(\lambda I-H)=n-1$. 即 λ 的几何重数 $=1\neq$ 代数重数，H 不是非亏损的，矛盾。 $\square$

19. 设 H 是一个不可约的上 Hessenberg 矩阵. 证明: 存在一个对角阵 D, 使得 $D^{-1}HD$ 的次对角元均为 1. $\kappa_2(D)=\|D\|_2\left\|D^{-1}\right\|_2$ 是多少?

解. 设 $D=\operatorname{diag}(d_1,\cdots,d_n)$, 则

$$\frac{d_i}{d_{i+1}}h_{i+1,i}=1$$

取 $d_1=1, d_k=h_{k,k-1}\cdots h_{21}$ 即可。 $\square$

20. 设 $H\in\mathbb{R}^{n\times n}$ 是一个上 Hessenberg 矩阵, 并假定 $x\in\mathbb{R}^n$ 是 H 的对应于实特征值 λ 的一个特征向量. 试给出一个算法计算正交矩阵 Q, 使得

$$Q^{\mathrm{T}}HQ=\begin{pmatrix} \lambda & w^{\mathrm{T}} \\ 0 & H_1 \end{pmatrix}$$

其中 H_1 是 $n-1$ 阶上 Hessenberg 矩阵.

解. 存在正交阵 $P=(x,P_1)$ 使

$$P^THP=\begin{pmatrix} \lambda & * \\ & H_1 \end{pmatrix}, H_1=P_1^THP_1$$

同时存在 $n-1$ 行的正交阵 Q_0 使 $Q_0^TH_1Q_0$ 是上 Hessenberg 阵。

取 $Q=P\begin{pmatrix} 1 & \\ & Q_0 \end{pmatrix}$, 则

$$Q^THQ=\begin{pmatrix} \lambda & * \\ & Q_0^TH_1Q_0 \end{pmatrix}$$

$\square$

21. 设 H 是一个奇异的不可约上 Hessenberg 矩阵. 证明: 进行一次基本的 QR 迭代后, H 的零特征值将出现.

解.

$$\operatorname{rank}(H)=n-1,\text{其 QR 分解为} H=Q_1R_1,$$

R_1 为最后一行为 0 的上三角阵，故 $A_1=R_1Q_1=\begin{pmatrix}\widetilde{A_1}\\ 0\end{pmatrix}$. □

22. 证明：若给定 $H=H_0$, 并由

$$H_k-\mu_kI=U_kR_k\ \text{和} H_{k+1}=R_kU_k+\mu_kI$$

产生矩阵 H_k, 则

$$\left(U_0\cdots U_j\right)\left(R_j\cdots R_0\right)=\left(H-\mu_0I\right)\cdots\left(H-\mu_jI\right).$$

解. 归纳，$j=0$ 时显然有 $H_0-\mu_0I=U_0R_0$.

假设 $j-1$ 时结论成立。则

$$\begin{aligned}
U_1\cdots U_jR_j\cdots R_1&=(H_1-\mu_1I)\cdots(H_1-\mu_jI)\\
U_0\cdots U_jR_j\cdots R_0&=U_0\left(R_0U_0+\mu_0I-\mu_1I\right)\cdots\left(R_0U_0+\mu_0I-\mu_jI\right)R_0\\
&=U_0\left(R_0U_0+\mu_0I-\mu_1I\right)U_0^{-1}\cdots U_0\left(R_0U_0+\mu_0I-\mu_jI\right)U_0^{-1}U_0R_0\\
&=(H_0-\mu_1I)\cdots(H_0-\mu_jI)U_0R_0\\
&=(H-\mu_0I)\cdots(H-\mu_jI)
\end{aligned}$$

□

23. 设 $A\in\mathbb{R}^{n\times n}$ 是一个具有互不相同对角元的上三角阵. 给出计算 A 的全部特征向量的详细算法.

解. A=

$$\begin{pmatrix}\lambda_1 & & *\\ & \ddots & \\ & & \lambda_n\end{pmatrix}$$

对 $A-\lambda_iI$ 应用反幂法即可。 □

24. 证明：对任意的 $A\in\mathbb{C}^{n\times n}$, 有 QL 分解 $A=QL$, 其中 Q 是酉矩阵, L 是下三角阵. 利用这一分解给出求矩阵特征值的 QL 算法, 并给出 QL 算法类似于定理 6.4.1 的收敛性定理.

解. 和 QR 没什么区别。可以照着写。 □

25. 设 H 是上 Hessenberg 矩阵, 并且假定已经用列主元 Gauss 消去法求得分解 $PH=LU$, 其中 P 是排列方阵, L 是单位下三角阵, U 是上三角阵. 证明: $\widetilde{H}=U\left(P^{\mathrm{T}}L\right)$ 仍是上 Hessenberg 矩阵, 并且相似于 H.

解.

$$H=(P^TL)U=(P^TL)\widetilde{H}(P^TL)^{-1}$$

H 上 Hessenberg，故 P^TL 上 Hessenberg，故 $\widetilde{H}$ 也是上 Hessenberg 的。 □

26. 设

$$A=\begin{pmatrix}\alpha_{11} & v^T\\ 0 & T\end{pmatrix}\in\mathbb{R}^{3\times 3},$$

其中 T 是一个有一对复共轭特征值的 2×2 矩阵. 设计一种算法计算一个 3 阶正交矩阵 Q, 使得

$$Q^TAQ=\begin{pmatrix}\tilde{T} & n\\ 0 & \alpha_{11}\end{pmatrix}$$

其中 $\lambda(\widetilde{T})=\lambda(T)$.

解.

$$Q=(e_1,e_2,e_3),Q^TAQ=\begin{pmatrix}\tilde{T} & n\\ 0 & \alpha_{11}\end{pmatrix},$$

$$\Rightarrow e_3^TA(e_1,e_2)=0,e_3^TAe_3=\alpha_{11},$$

取 e_3 为 A 的左特征向量，则上述条件自然成立。 □

27. 设 $A\in\mathbb{R}^{n\times n}$ 是一个如下形状的拟上三角阵:

$$A=\begin{pmatrix}A_{11} & A_{12} & \cdots & A_{1k}\\ & A_{22} & \cdots & A_{2k}\\ & & \ddots & \vdots\\ & & & A_{kk}\end{pmatrix},$$

其中 $A_{ii}(i=1,\cdots,k)$ 是 1×1 的或是一个有一对复共轭特征值的 2×2 矩阵. 设计一种计算 A 的全部特征向量的数值方法.

解. 对于 1×1 的矩阵直接对 $A-\lambda I$ 应用反幂法，对于 2×2 的，对其共轭复根 $A-(a\pm bi)I$ 应用反幂法。 □

28. 借助幂法设计一种求一个给定矩阵的最大奇异值的算法, 并讨论你所设计算法的收敛性.

解. 对 A^TA 应用幂法。 □

29. 借助反幂法设计一种计算一个给定矩阵的左右奇异向量的数值方法.

解. 对 A^TA 应用反幂法。 □

30. 设 $A=X\Lambda X^{-1}$, 其中

$$X=(x_1,\cdots,x_n),\quad \Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_n),$$

并假定

$$\lambda_1=\mathrm{e}^{\mathrm{i}\theta}\lambda_2,\quad 0<\theta<2\pi,\quad |\lambda_1|=|\lambda_2|>|\lambda_3|\geqslant\cdots\geqslant|\lambda_n|.$$

证明: 若 $\theta=2s\pi/t$, 其中 s 和 t 是两个互素的正整数, 则由幂法产生的向量序列有 t 个收敛的子序列, 且分别收敛到向量

$$\mathrm{e}^{\mathrm{i}\frac{2ks}{t}\pi}(y_1^*u_0)\,x_1+(y_2^*u_0)\,x_2,\quad k=1,\cdots,t,$$

这里 y_i^* 表示 X^{-1} 的第 i 行向量.

解.

$$A^k u_0 = X\Lambda^k X^{-1} u_0 = X \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}^k \begin{pmatrix} y_1^* u_0 \\ \vdots \\ y_n^* u_0 \end{pmatrix} = X \begin{pmatrix} \lambda_1^k y_1^* u_0 \\ \vdots \\ \lambda_n^k y_n^* u_0 \end{pmatrix} = \sum_{i=1}^n \lambda_i^k x_i y_i^* u_0$$

$$= \lambda_2^k \left(\sum_{i=2}^n \left(\frac{\lambda_i}{\lambda_2} \right)^k x_i y_i^* u_0 + e^{ik\theta} x_1 y_1^* u_0 \right)$$

$$u_k = \frac{A^k u_0}{\lambda_2^k} = e^{ik\theta} x_1 y_1^* u_0 + x_2 y_2^* u_0 + \sum_{i=3}^n \left(\frac{\lambda_i}{\lambda_2} \right)^k x_i y_i^* u_0$$

$$= e^{i\frac{2ks}{t}\pi} x_1 y_1^* u_0 + x_2 y_2^* u_0 + \sum_{i=3}^n \left(\frac{\lambda_i}{\lambda_2} \right)^k x_i y_i^* u_0$$

故有 t 个收敛子列，收敛到 $e^{i\frac{2ks}{t}\pi} (y_1^* u_0) x_1 + (y_2^* u_0) x_2$. □

31. 设 $A \in \mathbb{C}^{n\times n}$ 是非亏损的, 并假定 A 的特征值满足 $|\lambda_1| > |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$. 定义

$$q_0 = \frac{u}{\|u\|_2}, \quad q_k = \frac{Aq_{k-1}}{\|Aq_{k-1}\|_2}, \quad k \geqslant 1,$$

其中 u 是一个在 λ_1 的特征子空间上投影不为零的向量. 试证:

$$|q_k^* A q_k - \lambda_1| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right);$$

若再假定 A 是 Hermite 矩阵, 则

$$|q_k^* A q_k - \lambda_1| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^{2k} \right).$$

解. 设 A 的特征向量为 $x_1, \cdots, x_n$, 并设 $u = \alpha_1 x_1 + \cdots + \alpha_n x_n$,

$$q_k = \frac{A^k u}{\|A^k u\|_2} = \frac{\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n}{\|\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n\|_2},$$

$$q_k^* A q_k = \frac{(A^k u)^* A^{k+1} u}{\|A^k u\|_2^2} = \frac{(\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n)^* (\alpha_1 \lambda_1^{k+1} x_1 + \cdots + \alpha_n \lambda_n^{k+1} x_n)}{(\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n)^* (\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n)}$$

$$= \frac{(\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n)^* \left[\lambda_1 \left(\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n \right) + \left(\alpha_2 \lambda_2^k (\lambda_2 - \lambda_1) x_2 + \cdots + \alpha_n \lambda_n^k (\lambda_n - \lambda_1) x_n \right) \right]}{(\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n)^* (\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n)}$$

$$= \lambda_1 + \frac{(\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n)^* \left(\alpha_2 (\frac{\lambda_2}{\lambda_1})^k (\lambda_2 - \lambda_1) x_2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k (\lambda_n - \lambda_1) x_n \right)}{(\alpha_1 \lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n)^* (\alpha_1 x_1 + \lambda_2 (\frac{\lambda_2}{\lambda_1})^k x_2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k x_n)}$$

由 $|\lambda_1| > |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$ 知

$$|q_k^* A q_k - \lambda_1| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right).$$

若 A 时 Hermite 阵，则

$$|q_k^* A q_k - \lambda| = \left| \frac{\sum_{i=2}^n \alpha_i^2 (\lambda_i/\lambda_1)^{2k} (\lambda_i - \lambda_1) x_i}{\sum_{i=1}^n \alpha_i^2 (\lambda_i/\lambda_1)^{2k}} \right| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^{2k} \right)$$

□

32. 反幂法与 Rayleigh 商相结合就得到著名的 Rayleigh 商迭代法:

$$q_0=\frac{u_0}{\|u_0\|_2},\quad \mu_0=q_0^*Aq_0,$$
$$(A-\mu_kI)\,z_k=q_{k-1},\quad q_k=\frac{z_k}{\|z_k\|_2},\quad \mu_k=q_k^*Aq_k,\quad k\geqslant 1,$$

这里 $A\in\mathbb{C}^{n\times n}$ 和 $u_0\in\mathbb{C}^n$ 事先给定. 对于 Rayleigh 商迭代法, 我们可以用剩余

$$\rho_k=\|Aq_k-\mu_kq_k\|_2$$

来衡量 μ_k 和 q_k 作为 A 的近似特征值和特征向量的精度. 现假定 $\lim\limits_{k\to\infty}\rho_k=0$. 试证:

(1) 若取 $U_k\in\mathbb{C}^{n\times(n-1)}$, 使得 $Q_k=(q_k,U_k)$ 是酉矩阵, 则有

$$Q_k^*AQ_k=\begin{pmatrix}\mu_k & h_k^*\\ g_k & C_k\end{pmatrix},$$

其中 $g_k=U_k^*Aq_k, h_k^*=q_k^*AU_k, C_k=U_k^*AU_k$.

(2) $\rho_k=\|g_k\|_2$.

(3) 若定义

$$y_k=(\mu_{k-1}I-C_{k-1})^{-1}\,g_{k-1},\quad \delta_k=\left(1+\|y_k\|_2^2\right)^{\frac{1}{2}},$$

则有

$$q_k=\frac{1}{\delta_k}Q_{k-1}\begin{pmatrix}1\\ y_k\end{pmatrix}.$$

(4) 当 k 充分大时, 存在 $D\in\mathbb{C}^{(n-1)\times(n-1)}$, 使得

$$D^*\left(I+y_ky_k^*\right)D=I,$$

而且

$$D=I+O\left(\|g_{k-1}\|_2^2\right).$$

(5) $Q_k=Q_{k-1}\begin{pmatrix}1 & -y_k\\ y_k & I\end{pmatrix}\begin{pmatrix}\delta_k^{-1} & 0\\ 0 & D\end{pmatrix}$.

(6) 当 k 充分大时, 有

$$g_k=-\frac{h_{k-1}^*y_k}{\delta_k}y_k+O\left(\|g_{k-1}\|_2^3\right).$$

(7) 当 k 充分大时, 有 $\rho_k=O\left(\rho_{k-1}^2\right)$. 特别地, 当 A 是 Hermite 矩阵时, 有 $\rho_k=O\left(\rho_{k-1}^3\right)$. 这表明: 当 Rayleigh 商迭代法收敛时, 其收敛速度至少是二次的 (当 A 是 Hermite 矩阵时, 其收敛速度至少是三次的). 因此, 常用 Rayleigh 商迭代来加速收敛.

解. (1) 直接计算验证。

(2)

$$\begin{aligned}\rho_k&=\|Aq_k-\mu_kq_k\|_2=\|Q^*Aq_k-\mu_kQ^*q_k\|_2\\&=\left\|\begin{pmatrix}q_k^*Aq_k\\ U_k^*Aq_k\end{pmatrix}-\begin{pmatrix}\mu_kq_k^*q_k\\ \mu_kU_k^*q_k\end{pmatrix}\right\|_2=\left\|\begin{pmatrix}\mu_k\\ U_k^*Aq_k\end{pmatrix}-\begin{pmatrix}\mu_k\\ 0\end{pmatrix}\right\|_2\\&=\|U_k^*Aq_k\|_2=\|g_k\|_2\end{aligned}$$

(3)

$$q_k = \frac{z_k}{\|z_k\|_2} = \frac{(A-\mu_{k-1}I)^{-1}q_{k-1}}{\|z_k\|_2} = \frac{Q_{k-1}(Q_{k-1}^*AQ_{k-1}-\mu_{k-1}I)^{-1}Q_{k-1}^*q_{k-1}}{\|z_k\|_2}$$
$$= \frac{1}{\|z_k\|_2}Q_{k-1}\begin{pmatrix} 0 & h_{k-1}^* \\ g_{k-1} & C_{k-1}-\mu_{k-1}I \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

注意到

$$\begin{pmatrix} 0 & h_{k-1}^* \\ g_{k-1} & C_{k-1}-\mu_{k-1}I \end{pmatrix}\begin{pmatrix} 1 \\ (C_{k-1}-\mu_{k-1}I)^{-1}g_{k-1} \end{pmatrix} = \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}$$

ε 是一个常数。故

$$\begin{aligned} z_k &= Q_{k-1}\begin{pmatrix} 0 & h_{k-1}^* \\ g_{k-1} & C_{k-1}-\mu_{k-1}I \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \varepsilon\begin{pmatrix} 1 \\ y_k \end{pmatrix}, \\ q_k &= \frac{1}{\delta_k}Q_{k-1}\begin{pmatrix} 1 \\ y_k \end{pmatrix}. \end{aligned}$$

(4) $\lim\limits_{k\to\infty}\rho_k = 0$ 故当 $k\to\infty$ 有 $g_k\to 0$, 从而 $y_k\to 0$.
由于 $I+y_ky_k^*$ 时对称正定阵，则存在 D, $I+y_ky_k^* = (D^*)^{-1}D^{-1}$.

$$\begin{aligned} &DD^* = (I+y_ky_k^*)^{-1} = I - y_ky_k^* + \|y_k\|_2^2y_ky_k^* - \|y_k\|_2^4y_ky_k^* + \cdots =: (I+\widetilde{D})(I+\widetilde{D})^*, \\ &\widetilde{D}+\widetilde{D}^* = O(y_ky_k^*) = O(\|y_k\|_2^2) = O(\|g_{k-1}\|_2^2). \end{aligned}$$

因此 $D = I + O\left(\|g_{k-1}\|_2^2\right)$.

(5)

$$Q_{k-1}\begin{pmatrix} 1 & -y_k \\ y_k & I \end{pmatrix}\begin{pmatrix} \delta_k^{-1} & 0 \\ 0 & D \end{pmatrix} = Q_{k-1}\begin{pmatrix} \delta_k^{-1} & -y_k^*D \\ \delta_k^{-1}y_k & D \end{pmatrix} = \begin{pmatrix} q_k & * \end{pmatrix}.$$
$$\left(Q_{k-1}\begin{pmatrix} 1 & -y_k \\ y_k & I \end{pmatrix}\begin{pmatrix} \delta_k^{-1} & 0 \\ 0 & D \end{pmatrix}\right)^*\left(Q_{k-1}\begin{pmatrix} 1 & -y_k \\ y_k & I \end{pmatrix}\begin{pmatrix} \delta_k^{-1} & 0 \\ 0 & D \end{pmatrix}\right) = I,$$
$$\left(Q_{k-1}\begin{pmatrix} 1 & -y_k \\ y_k & I \end{pmatrix}\begin{pmatrix} \delta_k^{-1} & 0 \\ 0 & D \end{pmatrix}\right)\left(Q_{k-1}\begin{pmatrix} 1 & -y_k \\ y_k & I \end{pmatrix}\begin{pmatrix} \delta_k^{-1} & 0 \\ 0 & D \end{pmatrix}\right)^* = I.$$

故右边是第一列为 q_k 的酉方阵，故 $Q_k = Q_{k-1}\begin{pmatrix} 1 & -y_k \\ y_k & I \end{pmatrix}\begin{pmatrix} \delta_k^{-1} & 0 \\ 0 & D \end{pmatrix}$.

(6)

$$
\begin{aligned}
g_k = U_k^* A q_k &= (0, I_{k,k-1})\, Q_k^* A q_k \\
&= (0, I_{k,k-1}) \begin{pmatrix} \delta_k^{-1} & \delta_k^{-1} y_k^* \\ -D^* y_k & D^* \end{pmatrix} Q_{k-1}^* A Q_{k-1} Q_{k-1}^* q_k \\
&= \delta_k^{-1} \begin{pmatrix} -D^* y_k & D^* \end{pmatrix} \begin{pmatrix} \mu_{k-1} & h_{k-1}^* \\ g_{k-1} & C_{k-1} \end{pmatrix} \begin{pmatrix} 1 \\ y_k \end{pmatrix} \\
&= \delta_k^{-1} D^* \left[(C_{k-1} - \mu_{k-1} I)\, y_k + g_{k-1} - y_k h_{k-1}^* y_k \right] \\
&= -\delta_k^{-1} D^* y_k h_{k-1}^* y_k \\
&= -\delta_k^{-1} \left(I + O\left(\|g_{k-1}\|_2^2 \right) \right) y_k (h_{k-1}^* y_k) \\
&= -\frac{h_{k-1}^* y_k}{\delta_k} y_k + O\left(\|g_{k-1}\|_2^3 \right)
\end{aligned}
$$

(7)

$$
\rho_k = \|g_k\|_2 = O\left(\|y_k\|_2^2\right) = O\left(\|g_{k-1}\|_2^2\right) = O\left(\rho_{k-1}^2\right).
$$

$$
A \text{ Hermite}, h_{k-1}^* = g_{k-1}, \quad \|g_k\|_2^2 = O\left(\|y_k\|_2^2 \|g_{k-1}\|_2\right) = O\left(\|g_{k-1}\|_2^3\right) = O\left(\rho_{k-1}^3\right).
$$

$\square$