计算方法 作业 4

刘彦铭 ID: 122033910081

Last Edited: 2022 年 11 月 14 日

李庆杨等, 数值分析, 第 5 版, 华中科大, P.217, 1,2,3,4,5(1),7,8,9,11,12,13,14,15,16,18,19

1. 习题 1

$A = \begin{bmatrix} 5 & 2 & 1 \\ -1 & 4 & 2 \\ 2 & -3 & 10 \end{bmatrix}, b = [-12, 20, 3]^\top$, $Ax = b$. Jacobi 迭代: $x \leftarrow Bx + f$, 其中

$B = D^{-1}(L+U) = \begin{bmatrix} 0 & -0.4 & -0.2 \\ 0.25 & 0 & -0.5 \\ -0.2 & 0.3 & 0 \end{bmatrix}, f = D^{-1}b = \begin{bmatrix} -2.4 \\ 5 \\ 0.3 \end{bmatrix}$.

Gauss-Seidel 迭代: $x \leftarrow Gx + g$, 其中

$G = (D-L)^{-1}U = \begin{bmatrix} 0 & -0.4 & -0.2 \\ 0 & -0.1 & -0.55 \\ 0 & 0.05 & -0.125 \end{bmatrix}, g = (D-L)^{-1}b = \begin{bmatrix} -2.4 \\ 4.4 \\ 2.1 \end{bmatrix}$.

(1) 计算得到 $\rho(B) = 0.506 < 1, \rho(G) = 0.200 < 1$, 所以 Jacobi 迭代和 Gauss-Seidel 迭代在本例中都是收敛的。

(2) 两种迭代方法都能得到 $x = \begin{bmatrix} -4.0000 \\ 3.0000 \\ 2.0000 \end{bmatrix}$ 的解

2. 习题 2

$A = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$ 是幂零矩阵，$A^k = 0, \forall k \geq 2$. 所以 $\forall k \geq 2, I + A + A^2 + \cdots + A^k = I + A$, 该级数收敛

3. 习题 3

不妨设 A 为 n 阶方阵, $m = \max_{ij} |A_{ij}|$, 下用数学归纳法证明: $\max_{ij} |(A^k)_{ij}| \leq n^{k-1}m^k$.

(1) $k = 1$ 时显然成立;

(2) $\forall 1 \leq i, j \leq n$, $|(A^{k+1})_{ij}| = |\sum_k (A^k)_{ik} A_{kj}| \leq n \cdot (n^{k-1}m^k \cdot m) = n^k m^{k+1}$

所以存在常数 $N = \lceil 2nm \rceil$, 当 $k > N$ 时, $\max_{ij} |(A^k/k!)_{ij}| \leq \dfrac{n^{N-1}m^N}{N!} \cdot \left(\dfrac{1}{2}\right)^{k-N}$, 从而有 $\lim_{k\to\infty} \max_{ij} |(A^k/k!)_{ij}| = 0$. 由于 $|P|_\infty \leq n \cdot \max_{ij} |P_{ij}|, \forall P \in \mathbb{R}^{n\times n}$, 所以 $\lim_{k\to\infty} |A^k/k!|_\infty = 0$, 即该序列收敛于零。

4. 习题 4

写作矩阵形式 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leftarrow \begin{bmatrix} & -a_{12}/a_{11} \\ -a_{21}/a_{22} & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \end{bmatrix} = Bx + f$

收敛当且仅当 $\rho(B) < 1$. 考虑到 $\det(xI - B) = x^2 - (a_{12}a_{21})/(a_{11}a_{22})$, 故 $\rho(B) = \sqrt{|a_{12}a_{21}|/|a_{11}a_{22}|}$. 由此得收敛的充要条件是 $\left|\frac{a_{12}a_{21}}{a_{11}a_{22}}\right| < 1$.

5. 习题 5(1)

Jacobi 迭代: $D^{-1}(L+U) = \begin{bmatrix} 0 & -0.4 & -0.4 \\ -0.4 & 0 & -0.8 \\ -0.4 & -0.8 & 0 \end{bmatrix}$, $\rho(D^{-1}(L+U)) = 1.093 > 1$, 故不收敛;

Gauss-Seidel 迭代: $(D-L)^{-1}U = \begin{bmatrix} 0 & -0.4 & -0.4 \\ 0 & 0.16 & -0.64 \\ 0 & 0.032 & 0.672 \end{bmatrix}$, $\rho((D-L)^{-1}U) = 0.628 < 1$, 故收敛.

6. 习题 7

$Ax = b$, 其中 A 对称正定。Jacobi 迭代有 $x^{(k+1)} = Bx^{(k)} + f$ 其中 $B = D^{-1}(L+U) = I - D^{-1}A$. 若 λ 是 $D^{-1}A$ 的特征值，那么 $1-\lambda$ 是 $B = I - D^{-1}A$ 的特征值。就 5(1) 中的例子而言, A 是对称正定矩阵, $D^{-1}A$ 的一个特征值为 2.093, 故 -1.093 是 B 的一个特征值，从而 $\rho(B) \geq 1.093 > 1$, 迭代不收敛。

7. 习题 8

$Ax = b$, $A = \begin{bmatrix} 1 & 0 & -0.25 & -0.25 \\ 0 & 1 & -0.25 & -0.25 \\ -0.25 & -0.25 & 1 & 0 \\ -0.25 & -0.25 & 0 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}$,

(1) Jacobi 迭代: $x \leftarrow B_0 x + f$, $B_0 = \begin{bmatrix} 0 & 0 & 0.25 & 0.25 \\ 0 & 0 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0 & 0 \\ 0.25 & 0.25 & 0 & 0 \end{bmatrix}$, $f = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}$. 计算得到 $\det(xI - B_0) = x^2(x^2 - 1/4)$, 故 $\rho(B_0) = 1/2 < 1$.

(2) Gauss-Seidel 迭代: $B_0 = (D-L)^{-1}U$, $B_0 = \begin{bmatrix} 0 & 0 & 0.25 & 0.25 \\ 0 & 0 & 0.25 & 0.25 \\ 0 & 0 & 0.125 & 0.125 \\ 0 & 0 & 0.125 & 0.125 \end{bmatrix}$. 计算得到 $\det(xI - B_0) = x^3(x - 1/4)$, 故 $\rho(B_0) = 1/4 < 1$.

(3) 由 (1) (2) 知，改方程组的 Jacobi 迭代与 Gauss-Seidel 迭代都收敛.

8. 习题 9

矩阵形式的迭代公式 $x \leftarrow (D - \omega L)^{-1}((1-\omega)D + \omega U)x + \omega(D - \omega L)^{-1}b$.

取 $x_0 = [0,0,0]^\top$, 编程计算可得：解为 $x = [0.50000, 1.00000, -0.50000]$. 当 ω 取 $1.03, 1, 1.1$ 时，达到题设精度要求的迭代次数分别为 $5, 6, 6$ 次。

9. 习题 11

   整理得到迭代公式 $x \leftarrow (I-\omega A)x+\omega b$, 迭代收敛当且仅当 $\rho(I-\omega A)<1$.

   由于 A 对称正定，故存在正交矩阵 Q 使得 $A=Q\Lambda Q^\top$，其中 $\Lambda=\mathtt{diag}(\lambda_1,\cdots,\lambda_n)$，$\lambda_i$ 是矩阵 A 的特征值，且均为正实数。故 $I-\omega A=QQ^\top-\omega Q\Lambda Q^\top=Q(I-\omega\Lambda)Q^\top$. 所以对于 $I-\omega A$ 的任意特征值 γ, 存在 A 的特征值 λ, 使得 $\gamma=1-\omega\lambda$. 当 $0<\omega<2/\beta$，$0<\alpha\le\lambda(A)\le\beta$ 时有 $-1<1-\omega\beta\le\gamma=1-\omega\lambda\le 1-\alpha\omega<1$, 此时 $\rho(I-\omega A)<1$.

10. 习题 12

   (1) 根据 Gauss-Seidel 迭代的公式有: $x_i^{(k+1)}=\dfrac{1}{a_{ii}}\cdot\left(-\sum_{j<i}a_{ij}x_j^{(k+1)}-\sum_{j>i}a_{ij}x_j^{(k)}+b_i\right)$，所以 $x_i^{(k+1)}=\dfrac{1}{a_{ii}}\cdot\left(-\sum_{j<i}a_{ij}x_j^{(k+1)}-\sum_{j\ge i}a_{ij}x_j^{(k)}+a_{ii}x_i^{(k)}+b_i\right)=x_i^{(k)}+r_i^{(k+1)}/a_{ii}$.

   (2) 根据题中定义有 $r_i^{(k+1)}=b_i-\sum_{j<i}a_{ij}x_j^{(k+1)}-\sum_{j\ge i}a_{ij}x_j^{(k)}=\sum_j a_{ij}x_j^*-\sum_{j<i}a_{ij}x_j^{(k+1)}-\sum_{j\ge i}a_{ij}x_j^{(k)}=\sum_{j<i}a_{ij}\epsilon_j^{(k+1)}+\sum_{j\ge i}a_{ij}\epsilon_j^{(k)}$，其中 $\epsilon^{(k)}:=x^*-x^{(k)}$. (题目中定义的符号应该反了)。由 (1) 有 $\epsilon_i^{(k+1)}=x_i^*-x_i^{(k+1)}=x_i^*-x_i^{(k)}-r_i^{(k+1)}/a_{ii}=\epsilon_i^{(k)}-r_i^{(k+1)}/a_{ii}$.

   (3) 为避免繁琐的上标，这里用 r 表示 $r^{(k+1)}$, 用 ϵ,ϵ' 分别表示 $\epsilon^{(k)},\epsilon^{(k+1)}$. 由于 A 对称，故而 $A=D-L-U=D-L-L^\top$. 将 (2) 中结论写作向量形式，有 $\begin{cases}\epsilon'=\epsilon-D^{-1}r\\ r=-L\epsilon'+D\epsilon-L^\top\epsilon\end{cases}$，消去 ϵ' 整理得到 $A\epsilon=(I-LD^{-1})r=(D-L)D^{-1}r$. 于是有

$$
\begin{aligned}
Q(\epsilon')-Q(\epsilon) &= (\epsilon-D^{-1}r)^\top A(\epsilon-D^{-1}r)-\epsilon^\top A\epsilon\\
&= -2r^\top D^{-1}A\epsilon+r^\top D^{-1}AD^{-1}r\\
&= -2r^\top D^{-1}(D-L)D^{-1}r+r^\top D^{-1}(D-L-L^\top)D^{-1}r\\
&= r^\top D^{-1}(-D+L-L^\top)D^{-1}r\\
&= -r^\top D^{-1}r+r^\top D^{-1}LD^{-1}r-r^\top D^{-1}L^\top D^{-1}r
\end{aligned}
$$

   注意到 $r^\top D^{-1}LD^{-1}r\in\mathbb{R}^{1\times1}$, 所以 $r^\top D^{-1}LD^{-1}r=(r^\top D^{-1}LD^{-1}r)^\top=r^\top D^{-1}L^\top D^{-1}r$. 所以有 $Q(\epsilon')-Q(\epsilon)=-r^\top D^{-1}r=-\sum_{j=1}^n\dfrac{\left(r_j^{(k+1)}\right)^2}{a_{jj}}$.

   (4) (或许应当限定 A 是对称矩阵?)

   根据 (3) 有 $\forall k, Q(\epsilon^{(k+1)})\le Q(\epsilon^{(k)})$, 所以 $\forall k, Q(\epsilon^{(0)})\ge Q(\epsilon^{(k)})$.

   对于任意的 $y\in\mathbb{R}^{n\times1}$, 取 $x_0=y$ 为初始值, 利用 Gauss-Seidel 迭代解方程 $Ax=0$. 由于 A 非奇异, 所以有理论上的唯一解 $x^*=0$. 由于对于任意初始值都收敛，所以当取定 $x_0=y$ 时, 迭代收敛于 $x^*=0$. 所以 $y^\top Ay=(x_0-x^*)^\top A(x_0-x^*)=Q(\epsilon^{(0)})\ge\lim_{k\to\infty}Q(\epsilon^{(k)})=0$. 可以验证当且仅当 $x_0=y=0$ 时取得等号，所以 A 正定。

11. 习题 13

   写作矩阵形式: $\begin{bmatrix}A & B\\ B & A\end{bmatrix}\begin{bmatrix}z_1\\ z_2\end{bmatrix}=\begin{bmatrix}b_1\\ b_2\end{bmatrix}$

(1) $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \leftarrow \begin{bmatrix} & -A^{-1}B \\ -A^{-1}B & \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} A^{-1}b_1 \\ A^{-1}b_2 \end{bmatrix}$. 注意到 $\det\left(\begin{bmatrix} xI & -A^{-1}B \\ -A^{-1}B & xI \end{bmatrix}\right) = \det(xI + A^{-1}B)\det(xI - A^{-1}B)$, 所以收敛当且仅当 $\rho\left(\begin{bmatrix} & -A^{-1}B \\ -A^{-1}B & \end{bmatrix}\right) = \rho(A^{-1}B) < 1$.

(2) $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}^{(m+1)} = \begin{bmatrix} 0 & -A^{-1}B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}^{(m)} + \begin{bmatrix} 0 & 0 \\ -A^{-1}B & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}^{(m+1)} + \begin{bmatrix} A^{-1}b_1 \\ A^{-1}b \end{bmatrix}$,

即 $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \leftarrow \begin{bmatrix} I & 0 \\ -A^{-1}B & I \end{bmatrix} \begin{bmatrix} 0 & -A^{-1}B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} I & 0 \\ -A^{-1}B & I \end{bmatrix} \begin{bmatrix} A^{-1}b_1 \\ A^{-1}b \end{bmatrix}$.

记作 $z \leftarrow Gz + f$, $G = \begin{bmatrix} 0 & -A^{-1}B \\ 0 & A^{-1}BA^{-1}B \end{bmatrix}$, $\det(G) = x^n \cdot \det(xI - A^{-1}BA^{-1}B)$. 所以迭代收敛当且仅当 $\rho(G) = \rho(A^{-1}BA^{-1}B) < 1$.

考虑 $A^{-1}B$ 的 Schur 标准型 $A^{-1}B = URU^\star$, 有 $A^{-1}BA^{-1}B = UR^2U^\star$. 其中 R 是上三角阵，R^2 也是上三角阵且对角线元素是 R 中对应元素的平方。所以 $A^{-1}B$ 的非零特征值 λ 与 $A^{-1}BA^{-1}B$ 的非零特征值 λ^2 对应。所以迭代收敛当且仅当 $\rho(G) = \rho(A^{-1}BA^{-1}B) = \rho(A^{-1}B)^2 < 1$, 即 $\rho(A^{-1}B) < 1$.

根据上述推导可知，两种迭代方法同时收敛（或者不收敛）。当收敛时, 由于 $\rho(A^{-1}B)^2 < \rho(A^{-1}B) < 1$, 故第二种方法的收敛速度更快.

12. 习题 14

$\det(xI - A) = (x-1+a)^2(x-1-2a)$, 故 A 的特征值为 $1-a, 1-a, 1+2a$. 所以对于 $-1/2 < a < 1$, A 的全部特征值为正实数，A 正定. Jacobi 迭代收敛，当且仅当 $\rho(D^{-1}(L+U)) = \rho\left(\begin{bmatrix} 0 & -a & -a \\ -a & 0 & -a \\ -a & -a & 0 \end{bmatrix}\right)$.

其特征值为 $-a, -a, 2a$, 故要保证 $\rho < 1$ 需要有 $-1/2 < a < 1/2$.

13. 习题 15

取 $P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, 可以验证 $PAP^\top$ 是形如 $\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$ 的矩阵。

14. 习题 16

由于 C 的特征值全部为 0, 考虑 C 的 Schur 标准型, $C = URU^\star$, R 是对角线全 0 的上三角阵。容易验证: $\forall k \geq n, R^k = 0$, $C^k = UR^nU^\star = 0$. 因为 $x^{(k)} = C^k x^{(0)} + \sum_{0 \leq i < k} C^i g$, 所以 $\forall k \geq n$, $x^{(k)} = \sum_{0 \leq i < n} C^i g$, 即最多迭代 n 次就收敛。

补充对 $R^n = 0$ 的证明: 下归纳证明 $(R^k)_{ij} = 0, \forall i < j+k$: (1) 对于 $k=1$ 成立; (2) $(R^k)_{ij} = 0, \forall i < j+k$ $\Rightarrow$ $(R^{k+1})_{ij} = (R^kR)_{ij} = \sum_t (R^k)_{it} R_{tj}$. 当 $i < j+k+1$ 时, $i \geq t+k$ 与 $t \geq j+1$ 不能同时成立，故 $(R^k)_{it}$ 与 R_{tj} 中至少有一个为 0, 从而推出 $(R^{k+1})_{ij} = 0, \forall i < j+(k+1)$.

对于任意 $1 \leq i, j \leq n$, 有 $i < j+n$, 于是 $(R^n)_{ij} = 0$, 所以 $R^n = 0$.

15. 习题 18

考查 $G=(D-\omega L)^{-1}\left((1-\omega)D+\omega U\right)$ 的特征值 λ, 有:

$\det(\lambda I-G)=\det\left((D-\omega L)^{-1}\right)\det\left(\lambda(D-\omega L)-(1-\omega)D-\omega U\right)=0,$

即 $\det(\lambda(D-\omega L)-(1-w)D-\omega U)=0$. 下面验证 $B=\lambda(D-\omega L)-(1-\omega)D-\omega U$ 在 $|\lambda|\geq 1$ 时是不可约弱对角占优矩阵:

(1) B 显然是不可约的，因为 $P^\top BP=(\lambda-1+\omega)P^\top DP-\lambda\omega P^\top LP-\omega P^\top UP$ 不可能具有分块上三角矩阵的形式，这可以由 $A=D-L-U$ 的不可约性质得到。

(2) $|B_{ii}|=|\lambda-1+\omega||a_{ii}|$. 当 $j<i$ 时, $|B_{ij}|=|-\lambda\omega a_{ij}|=|\lambda\omega||a_{ij}|$. 当 $j>i$ 时, $|B_{ij}|=|w||a_{ij}|$. 故 $|B_{ii}|-\sum_{j\neq i}|B_{ij}|\geq(|\lambda-1+\omega|-|\lambda\omega|)\sum_{j<i}|a_{ij}|+(|\lambda-1+\omega|-|\omega|)\sum_{j>i}|a_{ij}|\geq 0$. 由于 A 是弱对角占优的，且 $|\lambda-1+\omega|\neq 0$, 所以至少存在一个 i 使得上述不等式的第一个不等号不取等，所以 B 是弱对角占优矩阵。

所以对于 $|\lambda|\geq 1$, B 是不可约的弱对角占优矩阵，根据定理 8.6 可知 B 非奇异，即 $\det(B)\neq 0$. 故 G 不存在模长大于等于 1 的特征值，即 $\rho(G)<1$, SOR 迭代收敛。

这里补充对 $|\lambda|\geq 1\Rightarrow|\lambda-1+\omega|\geq|\lambda\omega|\geq|\omega|$ 的证明:

$$\begin{aligned}
|\lambda-1+\omega|^2-|\lambda\omega|^2 &= (\lambda-1+\omega)(\bar\lambda-1+\omega)-\omega^2\lambda\bar\lambda\\
&=(1-\omega^2)\lambda\bar\lambda-(1-\omega)(\lambda+\bar\lambda)+(1-\omega)^2\\
&=(1-\omega)\left(\lambda\bar\lambda-\lambda-\bar\lambda+1+\omega(\lambda\bar\lambda-1)\right)\\
&=(1-\omega)\left((\lambda-1)\overline{(\lambda-1)}+\omega(\lambda\bar\lambda-1)\right)\\
&\geq 0
\end{aligned}$$

所以 $|\lambda-1+\omega|\geq|\lambda\omega|=|\lambda||\omega|\geq|\omega|$.

16. 习题 19

(1) 对于任意 $x\in\mathbb{R}^{n\times 1}$, 有 $x^\top A^\top Ax=(Ax)^\top(Ax)\geq 0$. 当且仅当 $Ax=0$, $x=A^{-1}Ax=0$ 时取等，故 $A^\top A$ 是正定矩阵。其对称性显然。

(2) 考虑 A 的奇异值分解 $A=U\Lambda V$, 其中 U, V 是正交矩阵，Λ 是对角线矩阵，且由于 A 非奇异，Λ 对角线元素均非零。$A^\top A=V^\top\Lambda U^\top U\Lambda V=V^\top\Lambda^2V$. 记 Λ^2 中的最大值，最小值分别为 λ_M,λ_m. 易知 $A^\top AA^\top A=V^\top\Lambda^4V$. 考虑到 $\texttt{cond(A)}_2=\sqrt{\dfrac{\lambda_{\texttt{max}}(A^\top A)}{\lambda_{\texttt{min}}(A^\top A)}}=\sqrt{\dfrac{\lambda_{\texttt{max}}(\Lambda^2)}{\lambda_{\texttt{min}}(\Lambda^2)}}=\sqrt{\dfrac{\lambda_M}{\lambda_m}}$, $\texttt{cond(A}^\top\texttt{A)}_2=\sqrt{\dfrac{\lambda_{\texttt{max}}(A^\top AA^\top A)}{\lambda_{\texttt{min}}(A^\top AA^\top A)}}=\sqrt{\dfrac{\lambda_{\texttt{max}}(\Lambda^4)}{\lambda_{\texttt{min}}(\Lambda^4)}}=\dfrac{\lambda_M}{\lambda_m}$. 所以有 $\texttt{cond}(A^\top A)_2=\left(\texttt{cond}(A)_2\right)^2$