

习题 1

1. 求下三角阵的逆矩阵的详细算法。

[解] 设下三角矩阵 L 的逆矩阵为 T

$$L = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}, \quad T = \begin{bmatrix} t_{11} & & & \\ t_{21} & t_{22} & & \\ \vdots & \vdots & \ddots & \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{bmatrix}$$

我们可以使用待定法，求出矩阵 T 的各列向量。为此我们将 T 按列分块如下：

$$T = [T_1, T_2, \dots, T_n]$$

注意到

$$LT = [LT_1, LT_2, \dots, LT_n] = [e_1, e_2, \dots, e_n] = I$$

我们只需运用算法 1 · 1 · 1，逐一求解方程

$$LT_i = e_i, \quad i = 1, 2, \dots, n$$

便可求得 $T = L^{-1}$ 。

[注意] 考虑到内存空间的节省，我们可以置结果矩阵 T 的初始状态为单位矩阵。

这样，我们便得到如下具体的算法：

算法（求解下三角矩阵 L 的逆矩阵 T ，前代法）

预置步 置 $T = I$

for $j = 1:n$

for $k = 1:n-1$

$$T(k, j) = T(k, j) / L(k, k)$$

$$T(k+1:n, j) = T(k+1:n, j) - T(k, j)L(j+1:n, k)$$

end

$$T(n, j) = T(n, j) / L(n, n)$$

end

2. 设 $S, T \in R^{n \times n}$ 为两个上三角矩阵，而且线性方程组 $(ST - \lambda I)x = b$ 是非奇异的，试给出一种运算量为 $O(n^2)$ 的算法，求解该方程组。

[解] 因 $(ST - \lambda I) = (S - \lambda T^{-1})T$ ，故为求解线性方程组 $(ST - \lambda I)x = b$ ，可先求得上三角矩阵 T 的逆矩阵 T^{-1} ，依照上题的思想我们很容易得到计算 T^{-1} 的算法。于是对该问题我们有如下解题的步骤：

(1) 计算上三角矩阵 T 的逆矩阵 T^{-1} ，算法如下：

算法 1（求解上三角矩阵的逆矩阵，回代法。该算法的的运算量为 n^3 ）

预置步 置 $\tilde{T} = I$

for $j = 1:n$

for $k = n:-1:2$

$$\tilde{T}(k, j) = \tilde{T}(k, j) / T(k, k)$$

$$\tilde{T}(1:k-1, j) = \tilde{T}(1:k-1, j) - \tilde{T}(k, j)T(1:k-1, k)$$

end

$$\tilde{T}(1, j) = \tilde{T}(1, j) / T(1, 1)$$

end

(2) 计算上三角矩阵 $\tilde{S} = S - \lambda T^{-1}$ 。运算量大约为 $n^2/2$ 。

(3) 用回代法求解方程组: $\tilde{S}y = b$ 。运算量为 n^2 ;

(4) 用回代法求解方程组: $Tx = y$ 。运算量为 n^2 。

算法总运算量大约为: $n^3 + O(n^2)$ 。

3. 证明: 如果 $L_k = I - l_k e_k^T$ 是一个 Gauss 变换, 则 $L_k^{-1} = I + l_k e_k^T$ 也是一个 Gauss 变换。

【解】按 Gauss 变换矩阵的定义, 易知矩阵 $I + l_k e_k^T$ 是 Gauss 变换。下面我们只需证明它是 Gauss 变换 $L_k = I - l_k e_k^T$ 的逆矩阵。事实上

$$(I - l_k e_k^T)(I + l_k e_k^T) = I + l_k e_k^T - l_k e_k^T - l_k e_k^T l_k e_k^T = I - l_k e_k^T l_k e_k^T$$

注意到 $l_k = [0, \dots, 0, l_{k+1,k}, \dots, l_{n,k}]^T$, 则显然有 $e_k^T l_k = 0$ 。从而有

$$L_k^{-1} = I + l_k e_k^T$$

4. 确定一个 Gauss 变换 L, 使

$$L \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 8 \end{bmatrix}$$

【解】比较比较向量 $(2, 3, 4)^T$ 和 $(2, 7, 8)^T$ 可以发现 Gauss 变换 L 应具有功能: 使向量 $(2, 3, 4)^T$ 的第二行加上第一行的 2 倍; 使向量 $(2, 3, 4)^T$ 的第三行加上第一行的 2 倍。于是 Gauss 变换如下

$$L = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ 2 & 0 & 1 \end{bmatrix}$$

5. 证明: 如果 $A \in R^{n \times n}$ 有三角分解, 并且是非奇异的, 那么定理 1.1.2 中的 L 和 U 都是唯一的。

【证明】设 $A = L_1 U_1 = L_2 U_2$, 其中 L_1, L_2 都是单位下三角阵, U_1, U_2 都是上三角阵。因为 A 非奇异的, 于是

$$L_2^{-1} L_1 = U_2 U_1^{-1}$$

注意到, 单位下三角阵的逆仍是单位下三角阵, 两个单位下三角阵的乘积仍是单位下三角阵; 上三角阵的逆仍是上三角阵, 两个上三角阵的乘积仍是上三角阵。

因此, 上述等将是一个单位下三角阵与一个上三角阵相等, 故此, 它们都必是单位矩阵。即 $L_2^{-1}L_1 = U_2U_1^{-1} = I$,

从而

$$L_1 = L_2, U_1 = U_2.$$

即 A 的 LU 分解是唯一的。

6. 设 $A = [a_{ij}] \in R^{n \times n}$ 的定义如下

$$a_{ij} = \begin{cases} 1, & \text{如果 } i = j \text{ 或 } j = n, \\ -1, & \text{如果 } i > j, \\ 0, & \text{其他} \end{cases}$$

证明 A 有满足 $|l_{ij}| \leq 1$ 和 $u_{nn} = 2^{n-1}$ 的三角分解。

[证明] 令 $L = [l_{ij}] \in R^{n \times n}$ 是单位下三角阵, $U = [u_{ij}] \in R^{n \times n}$ 是上三角阵。定义如下

$$l_{ij} = \begin{cases} 0, & i < j \\ 1, & i = j \\ -1, & i > j \end{cases}$$

$$u_{ij} = \begin{cases} 2^{i-1} & i = 1, 2, \dots, n; j = n. \\ 1 & i = j \\ 0 & \text{其余} \end{cases}$$

容易验证: $A = LU$.

7. 设 A 对称且 $a_{11} \neq 0$, 并假定经过一步 Gauss 消去之后, A 具有如下形式

$$\begin{bmatrix} a_{11} & a_1^T \\ 0 & A_2 \end{bmatrix}$$

证明 A_2 仍是对称阵。

[证明] 根据 Gauss 变换的属性, 显然做矩阵 A 的 LU 分解的第一步中的 Gauss 变换为

$$L_1 = \begin{bmatrix} 1 & & & \\ -\frac{a_{21}}{a_{11}} & 1 & & \\ \vdots & \vdots & \ddots & \\ -\frac{a_{n1}}{a_{11}} & 0 & \dots & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{a_1}{a_{11}} & I \end{bmatrix}$$

其中 $a_1^T = (a_{21}, \dots, a_{n1})$, 将 A 分块为

$$A = \begin{bmatrix} a_{11} & a_1^T \\ a_1 & A_{22}^{(0)} \end{bmatrix}$$

那么

$$L_1 A = \begin{bmatrix} 1 & 0 \\ -\frac{a_1}{a_{11}} & I \end{bmatrix} \begin{bmatrix} a_{11} & a_1^T \\ a_1 & A_{22}^{(0)} \end{bmatrix} = \begin{bmatrix} a_{11} & a_1^T \\ 0 & A_{22}^{(0)} - \frac{a_1 a_1^T}{a_{11}} \end{bmatrix}$$

即

$$A_2 = A_{22}^{(0)} - a_1 a_1^T / a_{11}$$

由 A 的对称性, A_2 对称性则是显而易见的。

8. 设 $A = [a_{ij}] \in R^{n \times n}$ 是严格对角占优阵, 即 A 满足

$$|a_{kk}| > \sum_{j=1, j \neq k}^n |a_{kj}|, \quad k = 1, 2, \dots, n.$$

又设经过一步 Gauss 消去后, A 具有如下形式

$$\begin{bmatrix} a_{11} & a_1^T \\ 0 & A_2 \end{bmatrix}$$

试证: 矩阵 A_2 仍是严格对角占优阵。由此推断: 对于对称的严格对角占优矩阵来说, 用 Gauss 消去法和列主元 Gauss 消去法可得同样的结果。

[证明] 依上题的分析过程易知, 题中的

$$A_2 = [\tilde{a}_{ij}], \text{ 其中: } \tilde{a}_{ij} = a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}}, i, j = 2, \dots, n.$$

于是 A_2 主对角线上的元素满足

$$|\tilde{a}_{ii}| = \left| a_{ii} - \frac{a_{i1}a_{1i}}{a_{11}} \right| > |a_{ii}| - \frac{|a_{i1}||a_{1i}|}{|a_{11}|}, \quad i = 2, \dots, n. \quad (1)$$

A_2 非主对角线上的元素满足

$$\sum_{k=2, k \neq i}^n |\tilde{a}_{ik}| = \sum_{k=2, k \neq i}^n \left| a_{ik} - \frac{a_{i1}a_{1k}}{a_{11}} \right| \leq \sum_{k=2, k \neq i}^n |a_{ik}| + \frac{|a_{i1}|}{|a_{11}|} \sum_{k=2, k \neq i}^n |a_{1k}|,$$

$$|a_{kk}| > \sum_{j=1, j \neq k}^n |a_{kj}|, \quad k = 1, 2, \dots, n.$$

由于 A 是严格对角占优的, 即

故

$$\sum_{k=2, k \neq i}^n |a_{ik}| < |a_{ii}| - |a_{i1}|,$$

$$\sum_{k=2, k \neq i}^n |a_{1k}| < |a_{11}| - |a_{1i}|.$$

从而

$$\sum_{k=2, k \neq i}^n |\tilde{a}_{ik}| < |a_{ii}| - \frac{|a_{i1}||a_{1i}|}{|a_{11}|}, \quad i = 2, \dots, n. \quad (2)$$

综合 (1) 和 (2) 得

$$|\tilde{a}_{ii}| > \sum_{k=2, k \neq i}^n |\tilde{a}_{ik}|, \quad i = 2, \dots, n.$$

即, 矩阵 A_2 仍是严格对角占优阵。

9. 设 $A \in R^{n \times n}$ 有三角分解。指出当把 Gauss 消去法应用于 $n \times (n+1)$ 矩阵 $[A, b]$ 时, 怎样才能不必存储 L 而解出 $Ax=b$? 需要多少次乘法运算?

[解] 用 Gauss 消去法作 A 的 LU 分解, 实际上就是对系数矩阵 A 作了一组初等行变换, 将其化为上三角矩阵 U 。而这一组的初等行变换对应的变换矩阵就是 L^{-1} , 即

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

如果把这一组初等行变换施加于方程右端向量 b 上, 即有

$$L^{-1}[A, b] = [U, L^{-1}b]$$

这就是说, 方程组 $Ax = b$ 和 $Ux = L^{-1}b$ 是同解方程。而后者是上三角形方程组, 可运用本章算法 1.1.2 求解。这样我们就不必存储 L , 通求解方程组 $Ux = L^{-1}b$, 来求解原方程组 $Ax = b$ 。算法如下:

(1) 用初等变换化 $[A, b]$ 为 $[U, L^{-1}b]$;

for $k=1:n-1$

$$A(k+1:n, k) = A(k+1:n, k) / A(k, k)$$

$$A(k+1:n, k+1:n) = A(k+1:n, k+1:n) - A(k+1:n, k)A(k, k+1:n)$$

$$b(k+1:n) = b(k+1:n) - A(k+1:n, k)b(k)$$

end

(2) 利用回代法求解方程组 $Ux = L^{-1}b$ 。

该算法所需要的加、减、乘、除运算次数为

$$\sum_{k=1}^{n-1} (3k + 2k^2) + n^2 = \frac{5}{3}n^3 + 2n^2 - \frac{8}{3}n$$

10. A 是正定阵, 如果对 A 执行 Gauss 消去一步产生一个形式为

$$\begin{bmatrix} a_{11} & a_1^T \\ 0 & A_2 \end{bmatrix}$$

的矩阵, 证明 A_2 仍是正定阵。

[证明] 不妨设

$$L_1 A = \begin{bmatrix} a_{11} & a_1^T \\ 0 & A_2 \end{bmatrix}, \quad \text{则有} \quad L_1 = \begin{bmatrix} 1 & 0 \\ -\frac{a_1}{a_{11}} & I \end{bmatrix}$$

从而有

$$L_1 A L_1^T = \begin{bmatrix} a_{11} & a_1^T \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{a_1}{a_{11}} & I \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ 0 & A_2 \end{bmatrix}$$

由于 L_1 非奇异, 故对 $\forall x \in R^{n-1}$ 且 $x \neq 0$, 构造 $\tilde{x} = (0, x^T)^T$, 及 $y = L_1^T \tilde{x}$, 则由 A 的正定性有

$$0 < y^T A y = \tilde{x}^T L_1 A L_1^T \tilde{x} = (0, x^T) \begin{bmatrix} a_{11} & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} 0 \\ x \end{bmatrix} = (0, x^T) \begin{bmatrix} 0 \\ A_2 x \end{bmatrix} = x^T A_2 x$$

由 x 的任意性知, A_2 正定。

11. 设

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{matrix} k \\ n-k \end{matrix}$$

并且 A_{11} 是非奇异的。矩阵

$$S = A_{22} - A_{21}A_{11}^{-1}A_{12}$$

称为是 A_{11} 在 A 中的 Schur 余阵。证明：如果 A_{11} 有三角分解，那么经过 k 步 Gauss 消去以后， S 正好等于 $(1 \cdot 1 \cdot 4)$ 的矩阵 $A_{22}^{(k)}$ 。

[证明] 因为 A_{11} 有三角分解，所以矩阵 A 可保证前 k 步 Gauss 消去法可以顺利完成。即有如下单位下三角矩阵

$$L = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{matrix} k \\ n-k \end{matrix}$$

使

$$LA = \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ 0 & A_{22}^{(k)} \end{bmatrix} \begin{matrix} k \\ n-k \end{matrix}$$

注意到

$$LA = \begin{bmatrix} L_{11}A_{11} & L_{11}A_{12} \\ L_{21}A_{11} + A_{21} & L_{21}A_{12} + A_{22} \end{bmatrix}$$

比较两式便知， $L_{21} = -A_{21}A_{11}^{-1}$ ，故有

$$A_{22}^{(k)} = A_{22} - A_{21}A_{11}^{-1}A_{12}$$

12. 证明：如果用全主元 Gauss 消去法得到 $PAQ=LU$ ，则对任意 i 有

$$|u_{ii}| \geq |u_{ij}|, j = i+1, \dots, n.$$

[证明] 略。

13. 利用列主元 Gauss 消去法给出一种求逆矩阵的实用算法。

[解] 设 A 是非奇异的，则应用列主元 Gauss 消去法可得到

$$PA = LU$$

这里： P 是置换阵， L 是单位下三角阵， U 是上三角阵。于是，通过求解下列 n 个方程组

$$LUx_j = e_j, \quad j = 1, 2, \dots, n.$$

便可求得

$$(PA)^{-1} = A^{-1}P = [x_1, x_2, \dots, x_n]$$

于是

$$A^{-1} = [x_1, x_2, \dots, x_n]P$$

也就是说，求 A 的逆矩阵，可按下列方案进行：

(1) 用列主元 Gauss 消去法得到： $PA = LU$ ；

(2) 经求解： $LUx_j = e_j, \quad j = 1, 2, \dots, n.$ 得 $(PA)^{-1} = X, \quad X \in R^{n \times n}$ ；

(3) 对 X 进行列置换得： $A^{-1} = XP$ 。

14. 假定已知 $A \in R^{n \times n}$ 的三角分解： $A=LU$ 。试设计一个算法来计算 A^{-1} 的 (i, j) 元素。

[解] 求解方程组

$$LUx = e_j$$

则 x 的第 i 个分量 x_i 就是 A^{-1} 的 (i, j) 元素。

15. 证明: 如果 $A^T \in R^{n \times n}$ 是严格对角占优阵 (参见第 8 题), 那么 A 有三角分解 $A=LU$ 并且 $\|v\| < 1$.

[证明] 仿照第 8 题的证明, 容易证明: 对于 $A^T \in R^{n \times n}$ 是严格对角占优阵, 经过一步 Gauss 消去后, 得到

$$A^{(1)} = \begin{bmatrix} a_{11} & a_1^T \\ 0 & A_2 \end{bmatrix} = L_1 A$$

其中 $A_2^T \in R^{(n-1) \times (n-1)}$ 仍是严格对角占优阵。A 的三角分解 $A=LU$ 中 $\|v\| < 1$.

这样, 我们在对 A 进行列主元三角分解时, 不需要选择主元, 因为每次消元时, 主元位置上的元素恰好是列主元。因此,

16. 形如 $N(y, k) = I - ye_k^T$ 的矩阵称作 Gauss-Jordan 变换, 其中 $y \in R^n$.

(1) 假定 $N(y, k)$ 非奇异, 试给出计算其逆矩阵的公式。

(2) 向量 $x \in R^n$ 满足何种条件才能保证存在 $y \in R^n$ 使得 $N(y, k)x = e_k$?

(3) 给出一种利用 Gauss-Jordan 变换求 $A \in R^{n \times n}$ 的逆矩阵 A^{-1} 的算法。并且说明 A 满足何种条件才能保证你的算法能够进行到底。

[解] 为解决本问题, 我们引入 Gauss-Jordan 变换的两个性质:

性质 1: $\det(N_k) = 1 - y_k$.

事实上,

$$\det(N_k) = \begin{vmatrix} 1 & & & -y_1 & & \\ & \ddots & & \vdots & & \\ & & 1 & -y_{k-1} & & \\ & & & 1 - y_k & & \\ & & & -y_{k+1} & 1 & \\ & & & \vdots & & \ddots \\ & & & -y_n & & & 1 \end{vmatrix} = \begin{vmatrix} 1 - y_k & & & & & \\ -y_{k+1} & 1 & & & & \\ \vdots & & \ddots & & & \\ -y_n & & & & 1 \end{vmatrix} = 1 - y_k.$$

性质 2: Gauss-Jordan 变换 $N_k = N(y, k)$ 非奇异的充分必要条件是 $y_k \neq 1$.

(1) 运用待定法, 首先设 $N(y, k) = I - ye_k^T$ 的逆矩阵为 $N(x, k) = I - xe_k^T$, 则有

$$\begin{aligned} I &= N(x, k)N(y, k) = (I - xe_k^T)(I - ye_k^T) = I - ye_k^T - xe_k^T + xe_k^T ye_k^T \\ &= I - (y + x - y_k x) e_k^T \end{aligned}$$

故应有

$$x = y / (y_k - 1), \quad y_k \neq 1$$

(2) 欲使 $N(y, k)x = e_k$, 则应有

$$x - x_k y = e_k$$

即

$$y = (x - e_k) / x_k = \left(\frac{x_1}{x_k}, \dots, \frac{x_{k-1}}{x_k}, 1 - \frac{1}{x_k}, \frac{x_{k+1}}{x_k}, \dots, \frac{x_n}{x_k} \right)$$

因此, $x \in R^n$ 应满足 $x_k \neq 0$, 便可按上述方法得到 $y \in R^n$ 使得 $N(y, k)x = e_k$.

(3) 设 A 的逆矩阵 $A^{-1} = T = [T_1, T_2, \dots, T_n]$, 则应有

$$AT_j = e_j, \quad j = 1, \dots, n.$$

下面我们给出利用 Gauss-Jordan 变换求解方程组的计算方法。算法如下: 假定 A 的各阶主子阵非零, 记 $A^{(0)} = A$, $T^{(0)} = I$.

第 1 步: 假若 $a_{11}^{(0)} \neq 0$, 令 $y_1 = (1 - 1/a_{11}^{(0)}, a_{21}^{(0)}/a_{11}^{(0)}, \dots, a_{n1}^{(0)}/a_{11}^{(0)})^T$, 构造 $N_1 = N(y_1, 1)$, 用 N_1 左乘 $A^{(0)}$ 和 $T^{(0)}$, 得到

$$A^{(1)} = N_1 A^{(0)}, \quad T^{(1)} = N_1 T^{(0)};$$

$$A^{(1)} = \begin{bmatrix} 1 & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{bmatrix},$$

其中

$$a_{1j}^{(1)} = \frac{a_{1j}^{(0)}}{a_{11}^{(0)}}, \quad j = 2, \dots, n.$$

$$a_{ij}^{(1)} = a_{ij}^{(0)} - \frac{a_{i1}^{(0)}}{a_{11}^{(0)}} a_{1j}^{(0)}, \quad i = 2, \dots, n; j = 2, \dots, n.$$

第 2 步: 假定 $a_{22}^{(1)} \neq 0$, 令 $y_2 = (a_{12}^{(1)}/a_{22}^{(1)}, 1 - 1/a_{22}^{(1)}, a_{32}^{(1)}/a_{22}^{(1)}, \dots, a_{n2}^{(1)}/a_{22}^{(1)})^T$, 构造 $N_2 = N(y_2, 2)$, 用 N_2 左乘 $A^{(1)}$ 和 $T^{(1)}$, 得到

$$A^{(2)} = N_2 A^{(1)}, \quad T^{(2)} = N_2 T^{(1)};$$

$$A^{(2)} = \begin{bmatrix} 1 & 0 & a_{13}^{(2)} & \dots & a_{1n}^{(2)} \\ 0 & 1 & a_{23}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{bmatrix}.$$

其中

$$a_{2j}^{(2)} = \frac{a_{2j}^{(1)}}{a_{22}^{(1)}}, \quad j = 3, \dots, n.$$

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{i1}^{(1)}}{a_{22}^{(1)}} a_{1j}^{(1)}, \quad i = 1, 3, 4, \dots, n; j = 3, \dots, n.$$

照此下去, 直到第 n 步: 假定 $a_{nn}^{(n-1)} \neq 0$,

$y_n = (a_{1n}^{(n-1)}/a_{nn}^{(n-1)}, \dots, a_{n-1,n}^{(n-1)}/a_{nn}^{(n-1)}, 1 - 1/a_{nn}^{(n-1)})^T$, 构造 $N_n = N(y_n, n)$, 用 N_n 左乘 $A^{(n-1)}$ 和 $T^{(n-1)}$, 得到

$$A^{(n)} = N_n A^{(n-1)}, \quad T^{(n)} = N_n T^{(n-1)};$$

$$A^{(n)} = I.$$

经上述 n 步, 我们得知:

$$I = A^{(n)} = N_n N_{n-1} \cdots N_1 A, \quad T^{(n)} = N_n N_{n-1} \cdots N_1$$

故

$$A^{-1} = T^{(n)}$$

从上面的约化过程可知, 要保证算法进行到底, 必须保证: $\alpha_{kk}^{(k-1)} \neq 0, \quad k=1, \dots, n$. 我们可以仿照定理 1.1.2 给出下列定理。

定理: $\alpha_{kk}^{(k-1)} \neq 0, \quad (k=1, \dots, n)$ 的充分必要条件是矩阵 A 的各阶顺序主子阵非奇异。

[证明] 对于 k 用归纳法。当 $k=1$ 时, $A_1 = \alpha_{11}^{(0)}$, 定理显然成立。假定定理直到 $k-1$ 成立, 下面只需证明: 若 A_1, \dots, A_{k-1} 非奇异, 则 A_k 非奇异的充要条件是 $\alpha_{kk}^{(k-1)} \neq 0$ 即可。由归纳假定知 $\alpha_{ii}^{(i-1)} \neq 0, \quad i=1, \dots, k-1$. 因此, Gauss-Jordan 约化过程至少可以进行 $k-1$ 步, 即可得到 $k-1$ 个 Gauss-Jordan 变换 N_1, \dots, N_{k-1} 使

$$A^{(k-1)} = N_{k-1} N_{k-2} \cdots N_1 A = \begin{bmatrix} I_{k-1} & A_{12}^{(k-1)} \\ 0 & A_{22}^{(k-1)} \end{bmatrix} \quad (16-1)$$

由此可知 $A^{(k-1)}$ 的 k 阶顺序主子阵有如下形式

$$\begin{bmatrix} I_{k-1} & * \\ 0 & \alpha_{kk}^{(k-1)} \end{bmatrix}$$

若将 N_1, \dots, N_{k-1} 的 k 阶顺序主子阵分别记为 $(N_1)_k, \dots, (N_{k-1})_k$, 则由 (16-1) 知

$$(N_{k-1})_k \cdots (N_1)_k A_k = \begin{bmatrix} I_{k-1} & * \\ 0 & \alpha_{kk}^{(k-1)} \end{bmatrix}$$

注意到 $\det(N_i)_k = \frac{1}{\alpha_{ii}^{(i-1)}}, \quad i=1, \dots, k-1$. 所以

$$\det A_k = \prod_{i=1}^k \alpha_{ii}^{(i-1)}.$$

即 A_k 非奇异的充要条件是 $\alpha_{kk}^{(k-1)} \neq 0$.

17. 证明定理 1·3·1 中的下三角阵 L 是唯一的。

[证明] 因 A 是正定对称矩阵, 故其各阶主子式均非零, 因此 A 非奇异。为证明 L 的唯一性, 不妨设有 L_1 和 L_2 使

$$A = L_1 L_1^T = L_2 L_2^T$$

那么

$$L_1^{-1} L_2 = L_1^T (L_2^T)^{-1}$$

注意到: L_1 和 L_2 是下三角阵, L_1^T 和 L_2^T 为上三角阵, 故它们的逆矩阵也分别是下三角阵和上三角阵。因此, $L_1^{-1} L_2$ 和 $L_1^T (L_2^T)^{-1}$ 只能是对角阵, 即

$$L_1^{-1} L_2 = L_1^T (L_2^T)^{-1} = D,$$

从而

$$L_2 = L_1 D, \quad (L_2^T)^{-1} = (D L_1^T)^{-1}, \quad D = L_1^T (D L_1^T)^{-1} = D^{-1}, \quad D = D^{-1} = I$$

于是得知

$$L_2 = L_1$$

18. 证明：如果 A 是一个带宽为 $2m+1$ 的对称正定带状矩阵，则其 Cholesky 因子 L 也是带状矩阵。 L 的带宽为多少？

[证明] 带宽为 $2m+1$ 的矩阵的认识：当 $m=1$ 时， $2m+1=3$ ，该带宽矩阵形为：

$$\begin{bmatrix} \times & \times & & & & \\ & \times & \times & \times & & \\ & & \times & \times & \times & \\ & & & \ddots & \ddots & \ddots \\ & & & & \times & \times & \times \\ & & & & & \times & \times \end{bmatrix}$$

对 m 为任意一个合适的正整数来说，带宽为 $2m+1$ 的矩阵元素有如下特征：

$$a_{ij} = 0, \quad |i - j| > m$$

结合这一特征，对于带宽为 $2m+1$ 的对称正定带状矩阵 A 的 Colicky 分解算法，可改写成下列形式：

```
for k = 1:n
    A(k,k) = sqrt(A(k,k))
    n1 = min(n, k+m)
    A(k+1:n1, k) = A(k+1:n1, k) / A(k, k)
    for j = k+1:n1
        A(j:n1, j) = A(j:n1, j) - A(j:n1, k)A(j, k)
    end
end
```

从算法不难看出：Colicky 因子 L 是下三角带状矩阵， L 的带宽为 $m+1$ 。

19. 若 $A = LL^T$ 是 A 的 Cholesky 分解，试证 L 的 i 阶顺序主子阵 L_i 正好是 A 的 i 阶顺序主子阵 A_i 的 Cholesky 因子。

[证明] 将 A 和 L 作如下分块

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad L = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}$$

其中： A_{11} ， L_{11} 为矩阵 A 和 L 的 i 阶顺序主子阵。 $A_{21}^T = A_{12}$ 。显然

$$A = \begin{bmatrix} A_{11} & A_{12}^T \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} L_{11}^T & L_{21}^T \\ 0 & L_{22}^T \end{bmatrix} = \begin{bmatrix} L_{11}L_{11}^T & L_{11}L_{21}^T \\ L_{21}L_{11}^T & L_{21}L_{21}^T + L_{22}L_{22}^T \end{bmatrix}$$

故有 $A_{11} = L_{11}L_{11}^T$ 。即 L_{11} 是 A_{11} 的 Colicky 分解。

20. 证明：若 $A \in \mathbb{R}^{n \times n}$ 是对称的，而且其前 $n-1$ 个顺序主子阵均非奇异，则 A 有唯一的分解式

$$A = LDL^T$$

其中 L 是单位下三角矩阵， D 是对角矩阵。

[证明] 先证明存在性。根据定理 1.1.2 知，存在单位下三角阵 L 和上三角阵

U ，使 $A=LU$ ，且 U 的主对角线上元素除 u_{nn} 外，其余都不为零。令

$D = \text{diag}(u_{11}, \dots, u_{n-1, n-1}, u_{nn})$ ，则有单位上三角阵 \tilde{U} 使 $U = D\tilde{U}$ ，即有

$$A = LD\tilde{U}$$

又因为 $A^T = A$ ，则

$$LD\tilde{U} = \tilde{U}^T D L^T$$

从而根据 L 和 \tilde{U} 的可逆性知：

$$D\tilde{U}(L^T)^{-1} = L^{-1}\tilde{U}^T D$$

该等式左端是一个上三角阵，右端是下三角阵。因此它们等于对角阵。再注意到单位上三角阵的乘积仍是单位上三角阵，单位下三角阵的乘积仍是单位下三角阵。因此两端都等于 D 。于是

$$\tilde{U}(L^T)^{-1} = L^{-1}\tilde{U}^T = I, \text{ 即 } L^T = \tilde{U}$$

从而有

$$A = LDL^T$$

再证唯一性。令 $A = L_1 D_1 L_1^T = L_2 D_2 L_2^T$ ，故有 $L_2^{-1} L_1 D_1 = D_2 L_2^T (L_1^T)^{-1}$ 。左边为下三角阵，右边为上三角阵，故等于对角阵。又因 $L_2^{-1} L_1 = L_2^T (L_1^T)^{-1} = I$ ，故 $L_1 = L_2$ ， $D_1 = D_2$ 。

21. 给出按行计算 Cholesky 因子 L 的详细算法。

[解] 略。

22. 利用改进的平方根法设计一种计算正定对称矩阵的逆的算法。

[解] 算法可分为以下几个步骤：

(1) 首先利用算法 1·3·2 计算出正定矩阵的如下分解

$$A = LDL^T$$

其中， L 是单位下三角阵， D 是对角阵。

(2) 求解矩阵方程

$$LX = I, \quad X \in R^{n \times n}$$

其解矩阵 $X = L^{-1}$ 。

(3) 求解矩阵方程

$$DY = X, \quad Y \in R^{n \times n}$$

其解矩阵 $Y = (LD)^{-1}$ 。

(4) 求解矩阵方程

$$L^T Z = Y, \quad Z \in R^{n \times n}$$

其解矩阵 $Z = (LDL^T)^{-1} = A^{-1}$ 。

[注意] 以上 (2)、(3)、(4) 步都是求解非常简单的方程组，算法实现起来很容易。

23. 设

$$A = \begin{bmatrix} 16 & 4 & 8 & 4 \\ 4 & 10 & 8 & 4 \\ 8 & 8 & 12 & 10 \\ 4 & 4 & 10 & 12 \end{bmatrix}, \quad b = \begin{bmatrix} 32 \\ 26 \\ 38 \\ 30 \end{bmatrix}$$

用平方根法证明 A 是正定的，并给出方程组 $Ax = b$ 的解。

[解] 由 Colicky 分解可得

$$A = LL^T$$

其中

$$L = \begin{bmatrix} 4 & & & \\ 1 & 3 & & \\ 2 & 2 & 2 & \\ 1 & 1 & 3 & 1 \end{bmatrix}$$

显然, L 是非奇异矩阵。因此, 对 $x \neq 0 \in \mathbb{R}^n, Lx \neq 0$. 于是

$$x^T Ax = x^T L L^T x = (L^T x)^T L^T x > 0$$

所以 A 是正定的。

由方程组 $L y = b$, 解得 $y = (8, 6, 5, 1)^T$, 再由方程组 $L^T x = y$, 解得 $x = (1, 1, 1, 1)^T$.

24. 设 $H = A + iB$ 是一个正定 Hermite 矩阵, 其中 $A, B \in \mathbb{R}^{n \times n}$.

证明: 矩阵

$$C = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$$

是正定对称的。

试给出一种仅用实数运算的算法来求解线性方程组

$$(A + iB)(x + iy) = (b + ic), \quad x, y, b, c \in \mathbb{R}^n.$$

[解] 既然 $H = A + iB$ 是正定的, 又对 $\forall z \in \mathbb{C}^n$, 有 $z = x + iy$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$, 且 $z^* = x^T - iy^T$. 且 $z \neq 0 \Leftrightarrow$ 或 $x \neq 0$, 或 $y \neq 0$. 注意到

$$z^* H z = (x^T - iy^T)(A + iB)(x + iy) = (x^T A x + y^T A y) + i(x^T B x + y^T B y)$$

显然 H 正等价于 A, B 正定。

对 $\forall z = (x^T, y^T)^T \in \mathbb{R}^{2n}$, 则有

$$z^T C z = (x^T, y^T) \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^T A x + y^T A y$$

由前面的讨论, 知道若 H 是正定的, 则 A 是正定的, 故矩阵 C 是正定的。

由于

$$(A + iB)(x + iy) = (Ax - By) + i(Ay + Bx)$$

于是求解原复数方程组, 等价于求解下列实方程组

$$\begin{cases} Ax - By = b \\ Ay + Bx = c \end{cases}$$

其矩阵形式为:

$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix}$$

由 (1) 得知系数矩阵正定, 故该方程可采用平方根算法求解。

习题 2 线性方程组的敏度分析与消去法的舍入误差分析

2.1 设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 n 个正数。证明：由

$$v(x) = \left(\sum_{i=1}^n \alpha_i x_i^2 \right)^{\frac{1}{2}}$$

定义的函数 $v: \mathbf{R}^n \rightarrow \mathbf{R}$ 是一个范数。

证明 只需验证 $v(x)$ 满足定义 2.1.1 的三个条件。其中 (1) 和 (2)，即正定性和齐次性显然成立，下面给出 (3) 三角不等式的证明。像 2 范数的证明一样，要证明三角不等式，需要用到 Cauchy-Schwartz 不等式

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n y_i^2 \right)^{\frac{1}{2}}, \quad x, y \in \mathbf{R}^n.$$

欲证明这个不等式，只需证明：对任意的 $x, y \in \mathbf{R}^n$ ，有下列等式成立

$$\left(\sum_{i=1}^n x_i y_i \right)^2 = \left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{i=1}^n y_i^2 \right) - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (x_i y_j - y_i x_j)^2,$$

用数学归纳法证明。当 $n=1$ 时，等式显然成立。不妨归纳假设当 $n \leq k$ 时，等式仍然成立，即有

$$\left(\sum_{i=1}^k x_i y_i \right)^2 = \left(\sum_{i=1}^k x_i^2 \right) \left(\sum_{i=1}^k y_i^2 \right) - \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k (x_i y_j - y_i x_j)^2, \quad (\text{E2.1})$$

现在来考虑 $n = k+1$ 时的情形，注意到

$$\begin{aligned} & \left(\sum_{i=1}^{k+1} x_i^2 \right) \left(\sum_{i=1}^{k+1} y_i^2 \right) - \frac{1}{2} \sum_{i=1}^{k+1} \sum_{j=1}^{k+1} (x_i y_j - y_i x_j)^2 \\ &= \left(\sum_{i=1}^k x_i^2 + x_{k+1}^2 \right) \left(\sum_{i=1}^k y_i^2 + y_{k+1}^2 \right) - \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^{k+1} (x_i y_j - y_i x_j)^2 - \frac{1}{2} \sum_{j=1}^{k+1} (x_{k+1} y_j - y_{k+1} x_j)^2 \\ &= \left(\sum_{i=1}^k x_i^2 \right) \left(\sum_{i=1}^k y_i^2 \right) + x_{k+1}^2 \left(\sum_{i=1}^k y_i^2 \right) + y_{k+1}^2 \left(\sum_{i=1}^k x_i^2 \right) + x_{k+1}^2 y_{k+1}^2 \\ & \quad - \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k (x_i y_j - y_i x_j)^2 - \frac{1}{2} \sum_{i=1}^k (x_i y_{k+1} - y_i x_{k+1})^2 - \frac{1}{2} \sum_{j=1}^k (x_{k+1} y_j - y_{k+1} x_j)^2 \\ &= \left(\sum_{i=1}^k x_i^2 \right) \left(\sum_{i=1}^k y_i^2 \right) - \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k (x_i y_j - y_i x_j)^2 \\ & \quad + x_{k+1}^2 \left(\sum_{i=1}^k y_i^2 \right) + y_{k+1}^2 \left(\sum_{i=1}^k x_i^2 \right) - \sum_{i=1}^k (x_{k+1}^2 y_i^2 - 2x_i y_i x_{k+1} y_{k+1} + y_{k+1}^2 x_i^2) + x_{k+1}^2 y_{k+1}^2 \\ &= \left(\sum_{i=1}^k x_i y_i \right)^2 + 2x_{k+1} y_{k+1} \left(\sum_{i=1}^k x_i y_i \right) + x_{k+1}^2 y_{k+1}^2 = \left(\sum_{i=1}^{k+1} x_i y_i \right)^2 \end{aligned}$$

至此，我们便证明了前述等式。亦即证明了 Cauchy-Schwartz 不等式。

又因为 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 n 个正数，因此有

$$\left| \sum_{i=1}^n \alpha_i x_i y_i \right| \leq \left(\sum_{i=1}^n \alpha_i x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n \alpha_i y_i^2 \right)^{\frac{1}{2}} = v(x) \cdot v(y)$$

从而对 $x, y \in \mathbf{R}^n$, 我们有

$$\begin{aligned} v(x+y) &= \left(\sum_{i=1}^n \alpha_i (x_i + y_i)^2 \right)^{\frac{1}{2}} \\ &= \left(\sum_{i=1}^n \alpha_i x_i^2 + 2 \sum_{i=1}^n \alpha_i x_i y_i + \sum_{i=1}^n \alpha_i y_i^2 \right)^{\frac{1}{2}} \\ &\leq \left(v^2(x) + 2v(x) \cdot v(y) + v^2(y) \right)^{\frac{1}{2}} \\ &= v(x) + v(y). \end{aligned}$$

2.2 证明: 当且仅当 x 和 y 线性相关且 $x^T y \geq 0$ 时, 才有 $\|x+y\|_2 = \|x\|_2 + \|y\|_2$.

证明 因为对任意的 $x, y \in \mathbf{R}^n$

$$\begin{aligned} \|x+y\|_2^2 &= (x+y)^T (x+y) = x^T x + x^T y + y^T x + y^T y \\ &= \|x\|_2^2 + 2x^T y + \|y\|_2^2 \end{aligned}$$

于是,

$$\|x+y\|_2 = \|x\|_2 + \|y\|_2$$

当且仅当

$$x^T y = \|x\|_2 \|y\|_2 \geq 0.$$

由等式 (E2.1) 可知, $x^T y = \|x\|_2 \|y\|_2$ 当且仅当

$$\sum_{i=1}^n \sum_{j=1}^n (x_i y_j - y_i x_j)^2 = 0,$$

即, 对任意的 $0 \leq i, j \leq n, x_i y_j = y_i x_j$, 此式成立不外乎二种情形: 或 $x = 0$; 或 $y = 0$; 或 $\alpha x + \beta y = 0$ 且 $|\alpha| + |\beta| \neq 0$. 即 x 和 y 线性相关。

2.3 证明: 如果 $A = [a_1, a_2, \dots, a_n]$ 是按列分块的, 那么

$$\|A\|_F^2 = \|a_1\|_2^2 + \|a_2\|_2^2 + \dots + \|a_n\|_2^2.$$

证明 因为

$$\|A\|_F^2 = \sum_{i,j=1}^n |a_{ij}|^2 = \sum_{j=1}^n \sum_{i=1}^n |a_{ij}|^2 = \sum_{j=1}^n \|a_j\|_2^2.$$

2.4 证明: $\|AB\|_F \leq \|A\|_2 \|B\|_F$ 和 $\|AB\|_F \leq \|A\|_F \|B\|_2$.

证明 记 $B = [b_1, b_2, \dots, b_n]$, 那么, 根据第 3 题的结果我们有

$$\|AB\|_F = \left(\sum_{j=1}^n \|Ab_j\|_2^2 \right)^{\frac{1}{2}} \leq \left(\|A\|_2^2 \sum_{j=1}^n \|b_j\|_2^2 \right)^{\frac{1}{2}} = \|A\|_2 \|B\|_F.$$

根据 Frobenius 范数定义易知, 对 $A \in \mathbf{R}^{n \times n}, \|A^T\|_F = \|A\|_F, \|A^T\|_2 = \|A\|_2$. 于是

$$\|AB\|_F = \|B^T A^T\|_F \leq \|B^T\|_2 \|A^T\|_F = \|B\|_2 \|A\|_F = \|A\|_F \|B\|_2.$$

2.5 设 $\nu: \mathbf{R}^{n \times n} \rightarrow \mathbf{R}$ 是由

$$\nu(A) = \max_{1 \leq i, j \leq n} |a_{ij}|, \quad A \in \mathbf{R}^{n \times n}$$

定义的。证明 $\nu_1 = n\nu$ 是矩阵范数，并且举例说明 ν 不满足矩阵范数的相容性。

证明 (1) 证明 $\nu_1 = n\nu$ 是矩阵范数。因为

$$\nu_1(A) = \frac{1}{n} \max_{1 \leq i, j \leq n} |a_{ij}|, \quad A \in \mathbf{R}^{n \times n}$$

显然 ν_1 满足矩阵范数定义中的前三条：正定性、齐次性、三角不等式。下面我们证明 ν_1 还满足“相容性”。对任意 $A, B \in \mathbf{R}^{n \times n}$ ，记 $A = (a_{ij})$ ， $B = (b_{ij})$ ，且

$$a = \max_{1 \leq i, j \leq n} |a_{ij}|, \quad b = \max_{1 \leq i, j \leq n} |b_{ij}|,$$

则 $\nu_1(A) = \frac{1}{n}a$ ， $\nu_1(B) = \frac{1}{n}b$ ，且

$$\begin{aligned} \nu_1(AB) &= \frac{1}{n} \max_{1 \leq i, j \leq n} \left| \sum_{k=1}^n a_{ik} b_{kj} \right| \leq \frac{1}{n} \max_{1 \leq i, j \leq n} \sum_{k=1}^n |a_{ik} b_{kj}| \\ &\leq \frac{1}{n} \sum_{k=1}^n ab = \frac{1}{n}a \cdot \frac{1}{n}b = \nu_1(A) \cdot \nu_1(B). \end{aligned}$$

(2) 一个 ν 不满足矩阵范数的相容性的例子。取 $n = 2$ ，

$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$ ， $B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ ，则 $AB = \begin{bmatrix} 3 & 2 \\ 5 & 3 \end{bmatrix}$ 。于是 $\nu(A) = 2$ ， $\nu(B) = 2$ ， $\nu(AB) = 5$ ，从而 $\nu(AB) > \nu(A) \cdot \nu(B)$ 。

2.6 证明：在 \mathbf{R}^n 上，当且仅当 A 是正定矩阵时，函数 $f(x) = (x^T A x)^{\frac{1}{2}}$ 是一个向量范数。

证明 由于 A 是正定矩阵，不妨设 $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ 是 A 的特征值， $\xi_1, \xi_2, \dots, \xi_n$ 是其对应的标准正交特征向量，即

$$A\xi_i = \lambda_i \xi_i, \quad i = 1, 2, \dots, n.$$

$$\xi_i^T \xi_j = \delta_{ij}, \quad i, j = 1, 2, \dots, n.$$

显然， $\xi_1, \xi_2, \dots, \xi_n$ 是线性无关的。因此， $\mathbf{R}^n = \text{span}\{\xi_1, \xi_2, \dots, \xi_n\}$ 。记 $Q = [\xi_1, \xi_2, \dots, \xi_n]$ ， $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ ，那么 $Q^T A Q = \Lambda$ ， $A = Q \Lambda Q^T$ ，且对任意 $x \in \mathbf{R}^n$ ，总有 $\bar{x} \in \mathbf{R}^n$ 使 $x = Q\bar{x}$ ， $\bar{x} = Q^T x$ 。

命题的充分性是很显然的。因为 $f(x) = (x^T A x)^{\frac{1}{2}}$ 是 \mathbf{R}^n 上的向量范数，则由其正定性可知 A 必为正定矩阵。

现在我们来证明命题的必要性。即假设 A 是正定矩阵，则函数 $f(x) = (x^T A x)^{\frac{1}{2}}$ 满足向量范数定义三条性质：

正定性。由 A 的正定性，正定性显然成立。

齐次性。对任意的 $x \in \mathbf{R}^n$, $\alpha \in \mathbf{R}$, 因为

$$(\alpha^2 x^T A x)^{\frac{1}{2}} = |\alpha| (x^T A x)^{\frac{1}{2}}, \text{ 故有 } f(\alpha x) = |\alpha| f(x).$$

三角不等式。对于任意给定的 $x, y \in \mathbf{R}^n$, 有 $\bar{x}, \bar{y} \in \mathbf{R}^n$, 使

$$\begin{aligned} f(x+y) &= \sqrt{(x+y)^T A (x+y)} = \sqrt{(\bar{x}+\bar{y})^T Q^T A Q (\bar{x}+\bar{y})} \\ &= \sqrt{(\bar{x}+\bar{y})^T \Lambda (\bar{x}+\bar{y})} = \sqrt{\sum_{i=1}^n \lambda_i (\bar{x}_i + \bar{y}_i)^2} \end{aligned}$$

应用习题 2.1 的结果, 得

$$\begin{aligned} \sqrt{\sum_{i=1}^n \lambda_i (\bar{x}_i + \bar{y}_i)^2} &\leq \sqrt{\sum_{i=1}^n \lambda_i \bar{x}_i^2} + \sqrt{\sum_{i=1}^n \lambda_i \bar{y}_i^2} \\ &= \sqrt{\bar{x}^T \Lambda \bar{x}} + \sqrt{\bar{y}^T \Lambda \bar{y}} \\ &= \sqrt{x^T Q \Lambda Q^T x} + \sqrt{y^T Q \Lambda Q^T y} \\ &= \sqrt{x^T A x} + \sqrt{y^T A y} = f(x) + f(y) \end{aligned}$$

即有

$$f(x+y) \leq f(x) + f(y).$$

2.7 设 $\|\bullet\|$ 是 \mathbf{R}^m 上的一个向量范数, 并且设 $A \in \mathbf{R}^{m \times n}$. 证明: 若 $\text{rank}(A) = n$, 则 $\|x\|_A \stackrel{\text{def}}{=} \|Ax\|$ 是 \mathbf{R}^n 上的一个向量范数。

证明 当 $\text{rank}(A) = n$ 时, $Ax = 0$ 当且仅当 x 是 \mathbf{R}^n 上的零向量。再由假设 $\|\bullet\|$ 是 \mathbf{R}^m 上的一个向量范数, 于是可证得 $\|x\|_A \stackrel{\text{def}}{=} \|Ax\|$ 满足:

正定性。事实上, 对任意 $x \in \mathbf{R}^n$, $\|x\|_A = \|Ax\| \geq 0$, 而且 $\|x\|_A = \|Ax\| = 0$ 当且仅当 $x = 0$.

齐次性。事实上, 对所有的 $x \in \mathbf{R}^n$ 和 $\alpha \in \mathbf{R}$ 有 $\|A(\alpha x)\| = \|\alpha Ax\| = |\alpha| \|Ax\|$, 因此 $\|\alpha x\|_A = |\alpha| \|x\|_A$.

三角不等式。事实上, 对所有的 $x, y \in \mathbf{R}^n$ 有 $\|A(x+y)\| = \|Ax + Ay\| \leq \|Ax\| + \|Ay\|$, 因此有 $\|x+y\|_A \leq \|x\|_A + \|y\|_A$.

2.8 若 $\|A\| < 1$ 且 $\|I\| = 1$, 证明

$$\|(I-A)^{-1}\| \leq \frac{1}{1-\|A\|}.$$

证明 首先用反证法, 证明 $(I-A)^{-1}$ 的存在性。设 $(I-A)$ 奇异, 则 $(I-A)x = 0$

有非零解 x , 且 $x = Ax$, 于是 $0 < \|x\| = \|Ax\| \leq \|A\| \|x\|$, 从而 $\|A\| \geq 1$. 这与假设矛盾。

现在来证明命题中的不等式。注意到: $\|I\| = 1$, 且

$$I = (I - A)(I - A)^{-1} = (I - A)^{-1} - A(I - A)^{-1}$$

故有

$$1 = \|(I - A)^{-1} - A(I - A)^{-1}\| \geq \|(I - A)^{-1}\|(1 - \|A\|)$$

即

$$\|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$$

2.9 设 $\|\cdot\|$ 是由向量范数 $\|\cdot\|$ 诱导出的矩阵范数。证明：若 $A \in \mathbb{R}^{n \times n}$ 非奇异，则

$$\|A^{-1}\|^{-1} = \min_{x \neq 0} \|Ax\|$$

证明 因为 $\|\cdot\|$ 是向量范数诱导的矩阵范数，故 $\|I\| = 1$ ，且对 $\forall G \in \mathbb{R}^{n \times n}$ 和 $x \in \mathbb{R}^n$ ，有 $\|Gx\| \leq \|G\|\|x\|$ 。于是对 $x \in \mathbb{R}^n$ ，有 $\|x\| = \|A^{-1}Ax\| \leq \|A^{-1}\|\|Ax\|$ ，且当 $\|x\| = 1$ 时，有

$$\|A^{-1}\|^{-1} \leq \|Ax\|. \quad (\text{E2.2})$$

现在只需证明：存在 $x \in \mathbb{R}^n$ 且 $\|x\| = 1$ ，使 $\|Ax\| = \|A^{-1}\|^{-1}$ 即可。根据算子范数的定义，我们不妨假设 $y \in \mathbb{R}^n$ 且 $\|y\| = 1$ ，使 $\|A^{-1}\| = \|A^{-1}y\|$ 。再取 $x = A^{-1}y / \|A^{-1}\|$ ，显然 $\|x\| = 1$ ，且

$$\|Ax\| = \left\| \frac{AA^{-1}y}{\|A^{-1}\|} \right\| = \frac{\|y\|}{\|A^{-1}\|} = \frac{1}{\|A^{-1}\|}. \quad (\text{E2.3})$$

综合 (E2.2) 和 (E2.3) 得

$$\|A^{-1}\|^{-1} = \min_{x \neq 0} \|Ax\|.$$

2.10 设 $A = LU$ 是 $A \in \mathbb{R}^{n \times n}$ 的 LU 分解。这里 $|l_{ij}| \leq 1$ ，设 a_i^T 和 u_i^T 分别表示 A 和 U 的第 i 行，验证等式

$$u_i^T = a_i^T - \sum_{j=1}^{i-1} l_{ij} u_j^T,$$

并用它证明 $\|U\|_{\infty} \leq 2^{n-1} \|A\|_{\infty}$ 。

[解] 记

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix}, \quad L = \begin{bmatrix} l_1^T \\ l_2^T \\ \vdots \\ l_n^T \end{bmatrix}, \quad U = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix}$$

于是

$$a_i^T = l_i^T U = \sum_{j=1}^n l_{ij} u_j^T, \quad i = 1, 2, \dots, n.$$

注意到： $l_i^T = (l_{i1}, \dots, l_{i,i-1}, 1, 0, \dots, 0)$ 。则有

$$u_i^T = a_i^T - \sum_{j=1}^{i-1} l_{ij} u_j^T, \quad i = 1, 2, \dots, n.$$

现在来证明 $\|U\|_{\infty} \leq 2^{n-1} \|A\|_{\infty}$. 因为

$$\|u_1^T\|_{\infty} = \|a_1^T\|_{\infty} \leq \|A\|_{\infty},$$

$$\|u_2^T\|_{\infty} \leq \|a_2^T\|_{\infty} + |l_{21}| \|u_1^T\|_{\infty} \leq \|a_2^T\|_{\infty} + \|A\|_{\infty} \leq 2\|A\|_{\infty},$$

$$\|u_3^T\|_{\infty} \leq \|a_3^T\|_{\infty} + |l_{31}| \|u_1^T\|_{\infty} + |l_{32}| \|u_2^T\|_{\infty} \leq 2^2 \|A\|_{\infty},$$

.....

$$\|u_n^T\|_{\infty} \leq \|a_n^T\|_{\infty} + \sum_{j=1}^{n-1} |l_{nj}| \|u_j^T\|_{\infty} \leq (1 + 1 + 2^1 + 2^2 + \dots + 2^{n-2}) \|A\|_{\infty} = 2^{n-1} \|A\|_{\infty}.$$

2.11 设

$$A = \begin{bmatrix} 375 & 374 \\ 752 & 750 \end{bmatrix}.$$

(1) 计算 A^{-1} 和 $\kappa_{\infty}(A)$;

(2) 选择 $b, \delta b, x$ 和 δx , 使得

$$Ax = b, A(x + \delta x) = b + \delta b,$$

而且 $\|\delta b\|_{\infty} / \|b\|_{\infty}$ 很小, 但 $\|\delta x\|_{\infty} / \|x\|_{\infty}$ 却很大;

(3) 选择 $b, \delta b, x$ 和 δx , 使得

$$Ax = b, A(x + \delta x) = b + \delta b,$$

而且 $\|\delta x\|_{\infty} / \|x\|_{\infty}$ 很小, 但 $\|\delta b\|_{\infty} / \|b\|_{\infty}$ 却很大。

[解] (1) 显然

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} 750 & -374 \\ -752 & 375 \end{bmatrix} = \begin{bmatrix} 375 & -187 \\ -376 & 187.5 \end{bmatrix};$$

从而, $\|A^{-1}\|_{\infty} = 563.5$, $\|A\|_{\infty} = 1502$, 于是

$$\kappa_{\infty}(A) = \|A^{-1}\|_{\infty} \|A\|_{\infty} = 846377.$$

选取: $b = (749, 1502)^T$, $\delta b = (0.001, -0.001)^T$, 则可计算得

$$x = (1, 1)^T, \quad \delta x = (0.562, -0.5635)^T$$

$$\|\delta b\|_{\infty} / \|b\|_{\infty} = 0.001/1502 = 1/1502000, \quad \|\delta x\|_{\infty} / \|x\|_{\infty} = 0.5635.$$

$$\|\delta x\|_{\infty} / \|x\|_{\infty} = 846377 \|\delta b\|_{\infty} / \|b\|_{\infty}$$

选取: $x = (562, 563.5)^T$, $\delta x = (0.001, 0.001)^T$, 则可计算得

$$b = (-1, 1)^T, \quad \delta b = (0.749, 1.502)^T,$$

$$\|\delta x\|_{\infty} / \|x\|_{\infty} = 0.001/563.5 = 1/563500, \quad \|\delta b\|_{\infty} / \|b\|_{\infty} = 1.502.$$

$$\|\delta b\|_{\infty} / \|b\|_{\infty} = 846377 \|\delta x\|_{\infty} / \|x\|_{\infty}.$$

2.12 证明对任意的矩阵范数都有 $\|I\| \geq 1$, 并由此导出 $\kappa(A) \geq 1$

[证明] 由定理 2.1.6(1)可知, 对任意矩阵范数都有 $\rho(A) \leq \|A\|$, 而 $\rho(I) = 1$, 于是

$$\|I\| \geq 1,$$

从而

$$\kappa(A) = \|A^{-1}\| \|A\| \geq \|A^{-1}A\| = \|I\| \geq 1.$$

2.13 若 A 和 $A+E$ 都是非奇异的, 证明

$$\|(A+E)^{-1} - A^{-1}\| \leq \|E\| \|A^{-1}\| \|(A+E)^{-1}\|.$$

[证明] 因为

$$(A+E)^{-1} - A^{-1} = (A+E)^{-1}(I - (A+E)A^{-1}) = (A+E)^{-1}EA^{-1}$$

所以, 根据矩阵范数的相容性可得

$$\|(A+E)^{-1} - A^{-1}\| \leq \|(A+E)^{-1}\| \|E\| \|A^{-1}\|.$$

2.14 估计连乘 $f(x_1 \cdots x_n) = x_1 \cdots x_n(1+\varepsilon)$ 中 ε 的上界.

[解] 假定 $f(x_i) = x_i(1+\delta_i)$, $|\delta_i| \leq \mathbf{u}$. 那么

$$f(x_1 \cdots x_n) = \prod_{i=1}^n x_i(1+\delta_i) = \prod_{i=1}^n x_i \cdot \prod_{i=1}^n (1+\delta_i) = (1+\varepsilon) \prod_{i=1}^n x_i.$$

则

$$1+\varepsilon = \prod_{i=1}^n (1+\delta_i)$$

由定理 2.3.3, 若假定 $n\mathbf{u} \leq 0.01$, 则

$$\prod_{i=1}^n (1+\delta_i) \leq 1 + 1.01n\mathbf{u},$$

从而

$$\varepsilon \leq 1.01n\mathbf{u}.$$

2.15 证明: 若 $n\mathbf{u} \leq 0.01$, 则

$$f\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n x_i(1+\eta_i),$$

其中

$$|\eta_1| \leq 1.01(n-1)\mathbf{u}, \quad |\eta_i| \leq 1.01(n-i+1)\mathbf{u}, \quad (i \geq 2).$$

[证明] 由定理 23.2 得

$$f(x_1 + x_2) = (x_1 + x_2)(1+\delta_2) = x_1(1+\delta_1)(1+\delta_2) + x_2(1+\delta_2),$$

$$f(x_1 + x_2 + x_3) = f(f(x_1 + x_2) + x_3)$$

$$= x_1(1+\delta_1)(1+\delta_2)(1+\delta_3) + x_2(1+\delta_2)(1+\delta_3) + x_3(1+\delta_3),$$

以此类推, 我们有

$$f\left(\sum_{i=1}^n x_i\right) = f\left(\sum_{i=1}^{n-1} x_i + x_n\right) = \sum_{i=1}^n x_i \prod_{j=i}^n (1+\delta_j),$$

其中: $|\delta_1| = 0$; $|\delta_i| \leq \mathbf{u}$, $i = 2, \dots, n$. 令 $\prod_{j=i}^n (1+\delta_j) = 1+\eta_i$, 那么

$$fl\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n x_i(1+\eta_i),$$

再由定理 2.3.3 知

$$|\eta_1| \leq 1.01(n-1)\mathbf{u}, \quad |\eta_i| \leq 1.01(n-i+1)\mathbf{u}, \quad (i \geq 2).$$

2.16 设 $A \in \mathbf{R}^{n \times n}$, $x \in \mathbf{R}^n$, 而且 $n\mathbf{u} \leq 0.01$. 证明:

$$fl(Ax) = (A+E)x,$$

其中 $E = [e_{ij}]$ 的元素满足

$$|e_{ij}| \leq 1.01n|a_{i1}|\mathbf{u} \quad (i=1, \dots, n),$$

$$|e_{ij}| \leq 1.01(n-j+2)|a_{ij}|\mathbf{u} \quad (i=1, \dots, n, j=2, \dots, n).$$

[证明] 因为

$$Ax = \left(\sum_{j=1}^n a_{1j}x_j, \dots, \sum_{j=1}^n a_{nj}x_j \right)^T$$

由例 2.3.1 的结果我们可以得到

$$fl\left(\sum_{j=1}^n a_{ij}x_j\right) = \sum_{j=1}^n (1+\varepsilon_{ij})a_{ij}x_j, \quad i=1, \dots, n.$$

其中

$$1+\varepsilon_{ij} = (1+\gamma_{ij}) \prod_{k=j}^n (1+\delta_{ik}), \quad |\gamma_{ij}| \leq \mathbf{u}, |\delta_{ik}| \leq \mathbf{u}, \delta_{i1} = 0.$$

再由定理 2.3.3 得

$$|\varepsilon_{i1}| \leq 1.01n\mathbf{u}, \quad (i=1, \dots, n).$$

$$|\varepsilon_{ij}| \leq 1.01(n-j+2)\mathbf{u}, \quad (i=1, \dots, n, j=2, \dots, n).$$

令 $e_{ij} = \varepsilon_{ij}a_{ij}$, 则

$$|e_{i1}| \leq 1.01n|a_{i1}|\mathbf{u}, \quad (i=1, \dots, n).$$

$$|e_{ij}| \leq 1.01(n-j+2)|a_{ij}|\mathbf{u}, \quad (i=1, \dots, n, j=2, \dots, n).$$

注意到

$$\sum_{j=1}^n (1+\varepsilon_{ij})a_{ij}x_j = \sum_{j=1}^n a_{ij}x_j + \sum_{j=1}^n \varepsilon_{ij}a_{ij}x_j, \quad i=1, \dots, n.$$

从而得到

$$\begin{aligned} fl(Ax) &= \left(\sum_{j=1}^n a_{1j}x_j, \dots, \sum_{j=1}^n a_{nj}x_j \right)^T + \left(\sum_{j=1}^n e_{1j}x_j, \dots, \sum_{j=1}^n e_{nj}x_j \right)^T \\ &= (A+E)x, \end{aligned}$$

其中 $E = [e_{ij}]$.

2.17 证明: 若 x 是 n 维向量, 则 $fl(x^T x) = x^T x(1+\alpha)$, 其中 $|\alpha| \leq n\mathbf{u} + O(\mathbf{u}^2)$.

[证明] 由定理 2.3.2 可知, 对一切 $i=1, \dots, n$, 有

$$fl(x_i^2) = x_i^2(1+\delta_i), \quad |\delta_i| \leq \mathbf{u}$$

下面对 n 用数学归纳法证明。当 $n=1$ 时, 命题显然成立。

假设当 $n \leq k$ 时, 命题仍然成立, 即有

$$f_l(x^T x) = f_l\left(\sum_{i=1}^k x_i^2\right) = (1 + \alpha_k) \sum_{i=1}^k x_i^2, \quad |\alpha_k| \leq k\mathbf{u} + O(\mathbf{u}^2).$$

那么当 $n = k + 1$ 时, 我们有

$$f_l\left(\sum_{i=1}^{k+1} x_i^2\right) = f_l\left(f_l\left(\sum_{i=1}^k x_i^2\right) + f_l(x_{k+1}^2)\right) = \left((1 + \alpha_k) \sum_{i=1}^k x_i^2 + (1 + \delta_{k+1}) x_{k+1}^2\right)(1 + \gamma),$$

其中, $|\delta_{k+1}| \leq \mathbf{u}$, $|\alpha_k| \leq k\mathbf{u} + O(\mathbf{u}^2)$, $|\gamma| \leq \mathbf{u}$. 于是

$$\begin{cases} f_l\left(\sum_{i=1}^{k+1} x_i^2\right) \geq (1 - \mathbf{u}) \left((1 - k\mathbf{u} - O(\mathbf{u}^2)) \sum_{i=1}^k x_i^2 + (1 - \mathbf{u}) x_{k+1}^2\right) \\ f_l\left(\sum_{i=1}^{k+1} x_i^2\right) \leq (1 + \mathbf{u}) \left((1 + k\mathbf{u} + O(\mathbf{u}^2)) \sum_{i=1}^k x_i^2 + (1 + \mathbf{u}) x_{k+1}^2\right) \end{cases},$$

从而

$$\begin{cases} f_l\left(\sum_{i=1}^{k+1} x_i^2\right) \geq (1 - \mathbf{u}) (1 - k\mathbf{u} - O(\mathbf{u}^2)) \sum_{i=1}^{k+1} x_i^2 = (1 - (k+1)\mathbf{u} - O(\mathbf{u}^2)) \sum_{i=1}^{k+1} x_i^2 \\ f_l\left(\sum_{i=1}^{k+1} x_i^2\right) \leq (1 + \mathbf{u}) (1 + k\mathbf{u} + O(\mathbf{u}^2)) \sum_{i=1}^{k+1} x_i^2 = (1 + (k+1)\mathbf{u} + O(\mathbf{u}^2)) \sum_{i=1}^{k+1} x_i^2 \end{cases},$$

由介值定理显然存在 $|\alpha_{k+1}| \leq (k+1)\mathbf{u} + O(\mathbf{u}^2)$, 使

$$f_l\left(\sum_{i=1}^{k+1} x_i^2\right) = (1 + \alpha_{k+1}) \sum_{i=1}^{k+1} x_i^2,$$

即当 $n = k + 1$ 时, 命题亦成立。

习题 3 最小二乘问题的解法

1. 设

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

用正则化方法求对应的 LS 问题的解.

[解] 由定理 3.1.4 可知, LS 问题的解就是下列正则化方程组解:

$$A^T A x = A^T b$$

即

$$\begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 9 \\ 12 \end{bmatrix}$$

解得: $x = (-1, 1)^T$.

2. 设

$$A = \begin{bmatrix} 1 & 3 & 1 & 1 \\ 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

求对应的 LS 问题的全部解.

[解] 由定理 3.1.4 可知, LS 问题的解就是下列正则化方程组解:

$$\begin{bmatrix} 6 & 3 & 1 & 1 \\ 3 & 9 & 3 & 3 \\ 1 & 3 & 1 & 1 \\ 1 & 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 1 \\ 1 \end{bmatrix}$$

经初等行变换得其同解方程组

$$\begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 15 & 5 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

从而

$$\begin{bmatrix} 1 & 3 \\ 0 & 15 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 - x_3 - x_4 \\ 2 - 5x_3 - 5x_4 \end{bmatrix}$$

即

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ \frac{2}{15} - \frac{1}{3}x_3 - \frac{1}{3}x_4 \end{bmatrix},$$

其中 $x_3, x_4 \in \mathbb{R}$.

设 $x = (1 \ 0 \ 4 \ 6 \ 3 \ 4)^T$, 求一个 Householder 变换 H 和一个正数 α 使得

$$Hx = (1 \ \alpha \ 4 \ 6 \ 0 \ 0)^T.$$

[解] 由于 2 范数具有正交不变性, 故 $\|x\|_2 = \|Hx\|_2$. 于是

$$\alpha = \sqrt{1^2 + 4^2 + 6^2 + 3^2 + 4^2 - 1^2 - 4^2 - 6^2} = 5.$$

于是 $y = Hx = (1 \ 5 \ 4 \ 6 \ 0 \ 0)^T$, 令

$$v = x - y = (0 \ -1 \ 0 \ 0 \ 3 \ 4)^T$$

$$w = \frac{v}{\|v\|} = \frac{1}{\sqrt{26}} (0 \ -1 \ 0 \ 0 \ 3 \ 4)^T.$$

那么, $H = I - 2ww^T$ 可以验证满足该题的要求.

4. 确定 $c = \cos \theta$ 和 $s = \sin \theta$ 使得

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} 5 \\ 12 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \alpha \in \mathbf{R}$$

[解] 由 2 范数具有正交不变性, 故

$$\alpha = \frac{\sqrt{5^2 + 12^2}}{\sqrt{2}} = \frac{13}{2}\sqrt{2}.$$

于是

$$\begin{bmatrix} 5 & 12 \\ 12 & -5 \end{bmatrix} \begin{bmatrix} c \\ s \end{bmatrix} = \begin{bmatrix} \frac{13}{2}\sqrt{2} \\ \frac{13}{2}\sqrt{2} \end{bmatrix}$$

从而

$$\begin{bmatrix} c \\ s \end{bmatrix} = \begin{bmatrix} 5 & 12 \\ 12 & -5 \end{bmatrix}^{-1} \begin{bmatrix} \frac{13}{2}\sqrt{2} \\ \frac{13}{2}\sqrt{2} \end{bmatrix} = \frac{1}{169} \begin{bmatrix} 5 & 12 \\ 12 & -5 \end{bmatrix} \begin{bmatrix} \frac{13}{2}\sqrt{2} \\ \frac{13}{2}\sqrt{2} \end{bmatrix} = \begin{bmatrix} \frac{17}{26}\sqrt{2} \\ \frac{7}{26}\sqrt{2} \end{bmatrix}.$$

5. 假定 $x = (x_1, x_2)^T$ 是一个二维复向量, 给出一种算法计算一个如下形式的酉矩阵

$$Q = \begin{bmatrix} c & \bar{s} \\ -s & c \end{bmatrix}, \quad c \in \mathbf{R}, c^2 + |s|^2 = 1.$$

使得 Qx 的第二个分量为零.

[解] 对于复向量 x 的 2 范数定义如下:

$$\|x\|_2 = \sqrt{x^H x} = \sqrt{\bar{x}_1 x_1 + \bar{x}_2 x_2} = \sqrt{|x_1|^2 + |x_2|^2}$$

显然, 在复数空间中, 2 范数仍然保持着正交不变性. 即对酉矩阵 Q 有

$$\|x\|_2 = \|Qx\|_2$$

根据题意, 不妨设 $y = Qx = (y_1 \ 0)^T$, 从而

$$\begin{bmatrix} c & \bar{s} \\ -s & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ 0 \end{bmatrix}, \quad |x_1|^2 + |x_2|^2 = |y_1|^2.$$

注意到

$$Q^H = \begin{bmatrix} c & -\bar{s} \\ s & c \end{bmatrix}$$

于是

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} c & -\bar{s} \\ s & c \end{bmatrix} \begin{bmatrix} y_1 \\ 0 \end{bmatrix} = \begin{bmatrix} cy_1 \\ sy_1 \end{bmatrix}$$

由 $x_1 = cy_1$, 从而

$$|c| = \frac{|x_1|}{|y_1|} = \frac{|x_1|}{\sqrt{|x_1|^2 + |x_2|^2}}$$

不妨设 $c \geq 0$, 即

$$c = \frac{|x_1|}{\sqrt{|x_1|^2 + |x_2|^2}} = \frac{|x_1|}{\|x\|_2},$$

又因 $sy_1 = x_2$, 所以

$$s = \frac{x_2 \bar{y}_1}{y_1 \bar{y}_1} = \frac{x_2 \bar{x}_1}{|x_1| \sqrt{|x_1|^2 + |x_2|^2}}.$$

6. 假定 x 和 y 是 \mathbf{R}^n 中的两个单位向量, 给出一种使用 Givens 变换的算法, 计算一个正交阵 Q , 使得 $Qx = y$.

[解] 首先考虑对指定的一个二维非零向量 $x = (x_1, x_2)^T$ 和一个实数 y_1 , 如何构造 Givens 变换

$$G = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \quad c = \cos \theta, \quad s = \sin \theta.$$

使 $Gx = (y, \alpha)^T$. 注意 2 范数的正交不变性, 则

$$\alpha = \sqrt{x_1^2 + x_2^2 - y^2} \geq 0. \quad (\text{这里我们假定了 } y^2 \leq x_1^2 + x_2^2, \text{ 稍后对此加以处理})$$

那么, G 应满足

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y \\ \alpha \end{bmatrix}$$

即

$$\begin{bmatrix} x_1 & x_2 \\ x_2 & -x_1 \end{bmatrix} \begin{bmatrix} c \\ s \end{bmatrix} = \begin{bmatrix} y \\ \alpha \end{bmatrix}$$

注意 $x \neq 0$, 则矩阵

$$X = \begin{bmatrix} x_1 & x_2 \\ x_2 & -x_1 \end{bmatrix} \text{ 非奇异, 且 } X^{-1} = \frac{1}{\|x\|_2^2} \begin{bmatrix} x_1 & x_2 \\ x_2 & -x_1 \end{bmatrix}$$

于是

$$\begin{bmatrix} c \\ s \end{bmatrix} = X^{-1} \begin{bmatrix} y \\ \alpha \end{bmatrix} = \begin{bmatrix} \frac{x_1 y + \alpha x_2}{\|x\|_2^2} \\ \frac{x_2 y - \alpha x_1}{\|x\|_2^2} \end{bmatrix}$$

这样，我们便可考虑从 x 的前两个分量 $(x_1, x_2)^T$ 开始，施以 Givens 变换，使其第一个分量变换为 y_1 。然后对 $(x_2, x_3)^T$ 施以 Givens 变换，使其首分量变换为 y_2 ；这样一直继续 $n-1$ 次变换，最后使得 $(x_{n-1}, x_n)^T$ 变换为 $(y_{n-1}, y_n)^T$ 。

几点说明：

- 为使算法能一步步正常进行，需要首先对单位向量 x 用一组 Givens 变换进行规范化处理，使其成为标准单位向量 e_1 。这样在接下来的 $n-1$ 步的 Givens 变换中就能保证 $y^2 \leq x_1^2 + x_2^2$ 。
- 在规范化 x 后，对其实施正交变换的每一步中，可以通过逐次计算向量 y 的范数，当其等于 1 时，即可结束算法。因为此时， x 和 y 的剩余分量均以为零。

算法总结：

算法 1（用 Givens 变换求正交矩阵 G 使单位向量 x 满足： $Gx = e_1$ 。）

```
void standard(double **g, double *x, int n)
{
    int i, j;
    for(i=0; i<n; i++)
        for(j=0; j<n; j++)
            if(j==i) g[i][j]=1; else g[i][j]=0;

    double c, s;
    double a, b, t;
    for(i=n-2; i>=0; i--)
    {
        if(x[i+1]==0)
            continue;
        else if(fabs(x[i+1])>fabs(x[i]))
        {
            t=x[i]/x[i+1];
            s=1.0/sqrt(1.0+t*t);
            c=s*t;
        }
        else
        {
            t=x[i+1]/x[i];
            c=1.0/sqrt(1.0+t*t);
            s=c*t;
        }
        x[i]=c*x[i]+s*x[i+1];
        x[i+1]=0;
        for(j=0; j<n; j++)
        {
            a=g[i][j]; b=g[i+1][j];
            g[i][j]=c*a+s*b;
```

```

        g[i+1][j]=c*b-s*a;
    }
}

```

算法 2 (计算 Givens 变换, $(y, \alpha)^T = G(x_1, x_2)^T$, 其中 y, x_1, x_2 已知)

```

void GetCS(double *g, double *x, double y)
{
    double a;
    a=sqrt(x[0]*x[0]+x[1]*x[1]-y*y);
    if(a==0)
    {
        g[0]=1;
        g[1]=0;
    }
    else
    {
        g[0]=(x[0]*y+a*x[1])/(x[0]*x[0]+x[1]*x[1]);
        g[1]=(x[1]*y-a*x[0])/(x[0]*x[0]+x[1]*x[1]);
        x[0]=y;
        x[1]=a;
    }
}

```

算法 3 (使用 Givens 变换, 求正交矩阵 G 使单位向量 x, y 满足: $Gx = y$)

```

void XtoY(double **g, double *x, double *y, int n)
{
    standard(g, x, n);
    double c, s, t;
    double cs[2];
    t=0.0;
    for(int i=0; i<n-1; i++)
    {
        GetCS(cs, x+i, y[i]);
        for(int j=0; j<n; j++)
        {
            c=g[i][j]; s=g[i+1][j];
            g[i][j]=cs[0]*c+cs[1]*s;
            g[i+1][j]=cs[0]*s-cs[1]*c;
        }
        t+=y[i]*y[i];
        if(t==1)
            break;
    }
}

```

7. 设 x, y 是 \mathbf{R}^n 中的两个非零向量, 给出一种算法来确定一个 householder 矩阵 H , 使 $Hx = \alpha y$, 其中 $\alpha \in \mathbf{R}$.

[解] (1) 当 x, y 线性相关时, $H = I$.

(2) 当 x, y 线性无关时, 令: $w = (x - \alpha y) / \|x - \alpha y\|_2$, 则

$$H = I - 2ww^T.$$

即为所求。

8. 假定 $L \in \mathbf{R}^{m \times n} (m \geq n)$ 是下三角阵, 说明如何确定 Householder 矩阵 H_1, \dots, H_n , 使得

$$H_n \cdots H_1 L = \begin{bmatrix} L_1 \\ 0 \end{bmatrix},$$

其中 $L_1 \in \mathbf{R}^{n \times n}$ 是下三角阵。

[解] 为讨论方便, 我们记

$$L = L^{(0)} = \begin{bmatrix} l_{11}^{(0)} & & \\ \vdots & \ddots & \\ l_{n1}^{(0)} & \cdots & l_{nn}^{(0)} \\ \vdots & \vdots & \vdots \\ l_{m1}^{(0)} & \cdots & l_{mn}^{(0)} \end{bmatrix}.$$

第 1 步: 令

$$v_1 = (0, \dots, 0, l_{nn}^{(0)} - \sigma_1, l_{n+1,n}^{(0)}, \dots, l_{mn}^{(0)})^T,$$

其中:

$$\sigma_1 = (l_{nn}^{(0)})^2 + \sum_{p=n+1}^m (l_{pn}^{(0)})^2,$$

构造 $H_1 = I - 2v_1v_1^T / v_1^T v_1$, 则由 householder 变换的性质得

$$L^{(1)} = H_1 L^{(0)} = \begin{bmatrix} l_{11}^{(1)} & & \\ \vdots & \ddots & \\ l_{n-1,1}^{(1)} & \cdots & l_{n-1,n-1}^{(1)} \\ l_{n,1}^{(1)} & \cdots & l_{n,n-1}^{(1)} & l_{n,n}^{(1)} \\ l_{n+1,1}^{(1)} & \cdots & l_{n+1,n-1}^{(1)} & 0 \\ \vdots & \cdots & \vdots & \vdots \\ l_{m,1}^{(1)} & \cdots & l_{m,n-1}^{(1)} & 0 \end{bmatrix}.$$

第 2 步: 令 $v_2 = (0, \dots, 0, l_{n-1,n-1}^{(1)} - \sigma_2, l_{n+1,n-1}^{(1)}, \dots, l_{m,n-1}^{(1)})^T$, 其中: $\sigma_2 = (l_{n-1,n-1}^{(1)})^2$

+ $\sum_{p=n+1}^m (l_{pn,n-1}^{(1)})^2$, 构造: $H_2 = I - 2v_2v_2^T / v_2^T v_2$, 则

$$L^{(2)} = H_2 H_1 L^{(0)} = \begin{bmatrix} l_{11}^{(2)} & & & & \\ \vdots & \ddots & & & \\ l_{n-2,1}^{(2)} & \cdots & l_{n-2,n-2}^{(2)} & & \\ l_{n-1,1}^{(2)} & \cdots & l_{n-1,n-2}^{(2)} & l_{n-1,n-1}^{(2)} & \\ l_{n,1}^{(2)} & \cdots & l_{n,n-2}^{(2)} & l_{n,n-1}^{(2)} & l_{n,n}^{(2)} \\ l_{n+1,1}^{(2)} & \cdots & l_{n+1,n-2}^{(2)} & 0 & 0 \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ l_{n,l}^{(2)} & \cdots & l_{m,n-2}^{(2)} & 0 & 0 \end{bmatrix}.$$

.....;

第 n 步, 令: $v_n = (l_{11}^{(n-1)} - \sigma_n, 0, \dots, 0, l_{n+1,1}^{(n-1)}, \dots, l_{m,1}^{(n-1)})^T$, $\sigma_n = (l_{11}^{(n-1)})^2 + \sum_{p=n+1}^m (l_{p,1}^{(n-1)})^2$, 构造

$H_n = I - 2v_n v_n^T / v_n^T v_n$, 则有

$$L^{(n)} = H_n \cdots H_1 L = \begin{bmatrix} l_{11}^{(n)} & & \\ \vdots & \ddots & \\ l_{n,1}^{(n)} & \cdots & l_{n,n}^{(n)} \\ 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}.$$

9. 假定 $A \in \mathbf{R}^{m \times n}$ 的秩为 n , 并假定已经用部分主元 Gauss 消去法计算好了 LU 分解 $PA = LU$, 其中 $L \in \mathbf{R}^{m \times n}$ 是单位下三角阵, $U \in \mathbf{R}^{n \times n}$ 是上三角阵, $P \in \mathbf{R}^{m \times m}$ 是排列方阵. 说明怎样用上题中的分解方法去找向量 $z \in \mathbf{R}^n$ 使得

$$\|Lz - Pb\|_2 = \min.$$

并证明: 如果 $Ux = z$, 那么 $\|Ax - b\|_2 = \min$.

[解] 由上题的结果可知, 存在正交矩阵

$$Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}, \quad Q_1 \in \mathbf{R}^{n \times m}, Q_2 \in \mathbf{R}^{(m-n) \times m}.$$

使

$$QL = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} L = \begin{bmatrix} Q_1 L \\ Q_2 L \end{bmatrix} = \begin{bmatrix} L_1 \\ 0 \end{bmatrix}, \text{ 其中 } L_1 \in \mathbf{R}^{n \times n}.$$

由 2 范数的正交不变性可知,

$$\|Lz - Pb\|_2 = \|QLz - QPb\|_2$$

记

$$QPb = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \quad c_1 \in \mathbf{R}^n, \quad c_2 \in \mathbf{R}^{m-n}.$$

则有

$$\|Lz - Pb\|_2 = \|L_1 z - c_1\|_2 + \|c_2\|$$

从而最小二乘问题: $\|L_1 z - c_1\|_2 = \min$ 的求解算法可按下列计算过程实现:

- 利用上题的算法分解下三角阵, 即求正交阵 Q :

$$QL = \begin{bmatrix} Q_1 L \\ 0 \end{bmatrix} = \begin{bmatrix} L_1 \\ 0 \end{bmatrix};$$

- 计算: $c_1 = Q_1^T P b$;
- 求解下三角方程组 $L_1 z = c_1$.

现在来证明: 如果 $Ux = z$, 那么 $\|Ax - b\|_2 = \min$. 结合上面讨论的算法, 我们只须证明:

$$\|Ax - b\|_2 = \min \text{ 等价于 } \|L_1 Ux - c_1\|_2 = \min$$

事实上,

$$\|Ax - b\|_2^2 = \|QPAx - QPb\|_2^2 = \|QLUx - QPb\|_2^2 = \|L_1 Ux - c_1\|_2^2 + \|c_2\|_2^2.$$

$\|Ax - b\|_2 = \min$ 与 $\|L_1 Ux - c_1\|_2 = \min$ 是等价的.

10. 设 $A \in \mathbf{R}^{m \times n}$ 且存在 $X \in \mathbf{R}^{n \times m}$ 使得对每一个 $b \in \mathbf{R}^m, x = Xb$ 均极小化 $\|Ax - b\|_2$.

证明: $AXA = A$ 和 $(AX)^T = AX$.

[解] 由矩阵奇异值分解定理知, 设 $A \in \mathbf{R}^{m \times n}$ 的秩 $r = \text{rank}(A)$, 则存在 m 阶正交阵 Q 和 n 阶正交阵 P , 使

$$A = QDP^T$$

$$D = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}$$

其中: $\sigma_i = \sqrt{\lambda_i}, \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ 是 $A^T A$ 的非零特征值全体.

可以证明矩阵 $X \in \mathbf{R}^{n \times m}$, 且

$$X = P \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q^T.$$

事实上,

$$\begin{aligned} A^T A X b &= P D Q^T Q D P^T P \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q^T b = P \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q^T b \\ &= P \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q^T b = A^T b \end{aligned}$$

由定理 3.1.4 可知, 对任一 $b \in \mathbf{R}^m, x = Xb$ 是 $\|Ax - b\|_2 = \min$ 的解.

另外,

$$AXA = QDP^T P \begin{bmatrix} \Sigma^{-1} & \\ & 0 \end{bmatrix} Q^T QDP^T = QDP^T = A$$

$$AX = QDP^T P \begin{bmatrix} \Sigma^{-1} & \\ & 0 \end{bmatrix} Q^T = Q \begin{bmatrix} I_r & \\ & 0 \end{bmatrix} Q^T.$$

于是我们有

$$AXA = A, \quad (AX)^T = AX.$$

11. 设 $A \in \mathbf{R}^{n \times n}$ 是一个对角加边矩阵, 即

$$A = \begin{bmatrix} \alpha_1 & \rho_2 & \rho_3 & \cdots & \cdots & \rho_n \\ \beta_2 & \alpha_2 & 0 & \cdots & \cdots & 0 \\ \beta_3 & 0 & \alpha_3 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \alpha_{n-1} & 0 \\ \beta_n & 0 & \cdots & \cdots & 0 & \alpha_n \end{bmatrix}$$

试给出用 Givens 变换求 A 的 QR 分解的详细算法。

[解] 算法可分两部分: 第一部分通过 $n-1$ 次 Givens 变换, 将 A 变换成如下形式

$$A^{(1)} = \begin{bmatrix} \alpha_1^{(1)} & 0 & \cdots & \cdots & \cdots & 0 \\ \beta_2^{(1)} & \alpha_2^{(1)} & 0 & \cdots & \cdots & 0 \\ \beta_3^{(1)} & 0 & \alpha_3^{(1)} & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \alpha_{n-1}^{(1)} & 0 \\ \beta_n^{(1)} & 0 & \cdots & \cdots & 0 & \alpha_n^{(1)} \end{bmatrix}$$

第二部分通过 $n-1$ 次 Givens 变换, 将 $A^{(1)}$ 变换成如下形式

$$A^{(2)} = \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{(2)} \end{bmatrix}$$

下面我们详细描述这两部分算法过程:

第一部分: 对于 $k = 2, \dots, n$. 构造

$$G_k = G(1, k, \theta) = \begin{bmatrix} c & & s & & \\ & \ddots & & & \\ & & & & \\ -s & & c & & \\ & & & \ddots & \\ 1 & & & & k & 1 \end{bmatrix}$$

其中

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} \rho_k \\ \alpha_k \end{bmatrix} = \begin{bmatrix} 0 \\ r \end{bmatrix}$$

从而

$$A^{(1)} = G_n G_{n-1} \cdots G_2 A.$$

算法 1 (计算 Givens 变换 1)

```
function [c, s] = givens1(a, b)
    if(a == 0)
        c = 1; s = 0;
    else if(|a| > |b|)
        r = b/a; s = -1/sqrt(1+r*r); c = -sr;
    else
        r = a/b; c = 1/sqrt(1+r*r); s = -cr;
    end
```

算法 2: (完成本题中的第一部分计算)

```
for k = 2:n
    [c, s] = givens1(A(1, k), A(k, k))
    A(1, 1:n) = c*A(1, 1:n) + s*A(k, 1:n)
    A(k, 1:n) = -s*A(1, 1:n) + c*A(k, 1:n)
end.
```

第二部分, 也是需要 $n-1$ 个 Givens 变换

$$\tilde{G}_k = \tilde{G}(1, k, \theta) = \begin{bmatrix} c & & s & & \\ & \ddots & & & \\ & & c & & \\ -s & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{matrix} 1 \\ \\ k \\ \\ 1 \end{matrix}$$

其中

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \beta_k \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

从而

$$A^{(2)} = G_n G_{n-1} \dots G_2 A^{(1)}$$

算法 3 (计算 Givens 变换 2)

```
function [c, s] = givens2(a, b)
    if(b == 0)
        c = 1; s = 0;
    else if(|b| > |a|)
        r = a/b; s = 1/sqrt(1+r*r); c = sr;
    else
        r = b/a; c = 1/sqrt(1+r*r); s = cr;
    end
```

算法 4: (完成本题中的第二部分计算)

```

for k = 2:n
    [c,s] = givens1(A(1,1), A(k,1))
    A(1,1:n) = c * A(1,1:n) + s * A(k,1:n)
    A(k,1:n) = -s * A(1,1:n) + c * A(k,1:n)
end.

```

12. 利用等于

$$\|A(x + \alpha w) - b\|_2^2 = \|Ax - b\|_2^2 + 2\alpha w^T A^T (Ax - b) + \alpha^2 \|Aw\|_2^2.$$

证明：如果 $x \in X_{LS}$ ，那么 $A^T Ax = A^T b$.

[证明] 令泛函

$$f(x) = \|Ax - b\|_2^2,$$

如果 $x \in X_{LS}$ ，那么对 $w \in R^n$ 且 $\|w\|_2 = 1$ ，当 $\alpha \in R$ 且 $|\alpha|$ 充分小时， $f(x + \alpha w) > f(x)$ ，从而由 $f(x)$ 连续性有

$$\lim_{\alpha \rightarrow 0} \frac{f(x + \alpha w) - f(x)}{\alpha} = 2w^T A^T (Ax - b) = 0,$$

由 $w \in R^n$ 的任意性，则必有 $A^T (Ax - b) = 0$ ，即 $A^T Ax = A^T b$.

习题四 线性方程组的古典迭代解法

1. 设方程组 $Ax = b$ 的系数矩阵为

$$A_1 = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}$$

证明 对 A_1 来说, Jacobi 迭代不收敛, 而 G-S 迭代收敛; 而对 A_2 来说, Jacobi 迭代收敛, 而 G-S 迭代不收敛。

[解] 对于 A_1 , 则有

$$D = \begin{bmatrix} 2 & & \\ & 1 & \\ & & -2 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & & \\ -1 & 0 & \\ -1 & -1 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} 0 & 1 & -1 \\ & 0 & -1 \\ & & 0 \end{bmatrix}.$$

从而,

$$B = D^{-1}(L+U) = \begin{bmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ -1 & 0 & -1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix},$$
$$G = (D-L)^{-1}U = \begin{bmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

于是

$$\det(\lambda I - B) = \begin{vmatrix} \lambda & -\frac{1}{2} & \frac{1}{2} \\ 1 & \lambda & 1 \\ -\frac{1}{2} & -\frac{1}{2} & \lambda \end{vmatrix} = \lambda \left(\lambda^2 + \frac{5}{4} \right)$$
$$\det(\lambda I - G) = \begin{vmatrix} \lambda & -\frac{1}{2} & \frac{1}{2} \\ 0 & \lambda + \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \lambda + \frac{1}{2} \end{vmatrix} = \lambda \left(\lambda + \frac{1}{2} \right)^2$$

从而,

$$\lambda(B) = \left\{ 0, \frac{\sqrt{5}}{2}i, -\frac{\sqrt{5}}{2}i \right\}, \quad \lambda(G) = \left\{ 0, -\frac{1}{2}, -\frac{1}{2} \right\}.$$

即有

$$\rho(B) = \frac{\sqrt{5}}{2} > 1, \quad \rho(G) = \frac{1}{2}.$$

由定理 4.2.1 知, Jacobi 迭代法不收敛; G-S 迭代收敛。

对于 A_2 ,

$$D = \text{diag}(1,1,1), \quad L = \begin{bmatrix} 0 & & \\ -1 & 0 & \\ -2 & -2 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} 0 & -2 & 2 \\ & 0 & -1 \\ & & 0 \end{bmatrix},$$

从而

$$B = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 & -2 & 2 \\ & 2 & -3 \\ & & 2 \end{bmatrix}$$

进而

$$\det(\lambda I - B) = \lambda^3, \quad \det(G) = \lambda(\lambda - 2)^2$$

显然, $\rho(B) = 0$, $\rho(G) = 2$. 故由定理 4.2.1 知, Jacobi 迭代法收敛; G-S 迭代不收敛。

2. 设 $B \in R^{n \times n}$ 满足 $\rho(B) = 0$, 证明对任意的 $g, x_0 \in R^n$, 迭代格式

$$x^{(k+1)} = Bx^{(k)} + g, \quad k = 0, 1, \dots$$

最多迭代 n 次就可得方程组 $x = Bx + g$ 的精确解。

[证明] 由于 $\rho(B) = 0$, 故 B 的所有特征值均为零。于是存在正交矩阵 P 及矩阵

$$J = \begin{bmatrix} 0 & \alpha_1 & & & \\ & 0 & \alpha_2 & & \\ & & \ddots & \ddots & \\ & & & 0 & \alpha_{n-1} \\ & & & & 0 \end{bmatrix}$$

使 $B = P^{-1}JP$, 注意到 $J^n = 0$ 于是: $B^n = P^{-1}J^nP = 0$.

另一方面, 记: $y^{(k)} = x^{(k)} - x^*$, $k = 0, 1, \dots, n$. 从而,
 $y^{(n)} = By^{(n-1)} = \dots = B^n y^{(0)} = 0$, 即 $x^{(n)} = x^*$.

3. 考虑线性代数方程组

$$Ax = b$$

这里

$$A = \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{bmatrix}$$

- (1) α 为何值时, A 是正定的?
- (2) α 为何值时, Jacobi 迭代收敛?
- (3) α 为何值时, G-S 迭代收敛?

[解] (1) 对称矩阵 A 正定的充分必要条件是其特征值均为正数。而 A 的特征多项式为

$$\det(\lambda I - A) = (\lambda - 1)^3 - \alpha^2(\lambda - 1) = (\lambda - 1)(\lambda - 1 - \alpha)(\lambda - 1 + \alpha)$$

于是 A 的特征值为: $\lambda_1 = 1$, $\lambda_2 = 1 + \alpha$, $\lambda_3 = 1 - \alpha$. 欲使它们均大于零, 则
 $-1 < \alpha < 1$.

(2) 由于 Jacobi 迭代矩阵为

$$B = \begin{bmatrix} 0 & 0 & -\alpha \\ 0 & 0 & 0 \\ -\alpha & 0 & 0 \end{bmatrix}$$

B 的特征多项式为

$$\det(\lambda I - B) = \lambda^3 - \alpha^2 \lambda = \lambda(\lambda - \alpha)(\lambda + \alpha)$$

其特征值为: $\lambda_1 = 0$, $\lambda_2 = \alpha$, $\lambda_3 = -\alpha$, 于是 B 谱半径 $\rho(B) = |\alpha|$. 由定理

4.2.1 可知, Jacobi 迭代收敛当且仅当 $\rho(B) = |\alpha| < 1$. 从而当 $|\alpha| < 1$ 时, Jacobi 迭代收敛。

(3) 由于 G-S 迭代矩阵为

$$G = \begin{bmatrix} 0 & 0 & \alpha \\ 0 & 0 & 0 \\ 0 & 0 & -\alpha^2 \end{bmatrix}$$

其特征多项式为

$$\det(\lambda I - G) = \lambda^2(\lambda - \alpha^2)$$

特征值为: $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = \alpha^2$. 从而 $\rho(G) = |\alpha^2|$ 故由定理 4.2.1 可知, 当 $|\alpha^2| < 1$ 时, G-S 迭代收敛。

注意: (2) 和 (3) 中的 α 可以是复数。

4. 证明: 若 $A \in R^{n \times n}$ 非奇异, 则必可找到一个排列方阵 P 使得 PA 的对角元素均不为零。

[证明] 用数学归纳法证明。

$n = 1, 2$ 时, 结论显然成立。假定对 $(n-1)$ 阶非奇异矩阵, 结论也成立。那么对于 n 阶非奇异矩阵我们来证明结论也是成立的。将 A 按第 1 列作拉普拉斯展开, 即有

$$\det(A) = a_{11}A_{11} + a_{21}A_{21} + \cdots + a_{n1}A_{n1} \neq 0$$

从而必存在 $1 \leq i \leq n$, 使 $a_{i1}A_{i1} \neq 0$, 即 $a_{i1} \neq 0$, $A_{i1} \neq 0$. 排列阵 $P(1, i)$ 能使

$$A^{(1)} = P(1, i)A = \begin{bmatrix} a_{i1} & a \\ b & A_1 \end{bmatrix}$$

其中 $a^T \in R^{n-1}$, $b \in R^{n-1}$, $A_1 \in R^{(n-1) \times (n-1)}$, 且 (a_{i1}, a) 是 A 的第 i 行, $(a_{i1}, b^T)^T$ 是 A 的第 1 列左乘排列阵 $P(1, i)$ 后的向量, A_1 是 a_{i1} 的余子式。显然, $\det(A_1) = A_{i1} \neq 0$, 即 $A_1 \in R^{(n-1) \times (n-1)}$ 非奇异。由归纳假设, 则存在 $(n-1)$ 阶排列阵 P_1 使 $P_1 A_1$ 的对角元素均不为零。作 n 阶排列阵

$$\tilde{P} = \begin{bmatrix} 1 & 0 \\ 0 & P_1 \end{bmatrix}$$

则 $\tilde{P}A^{(1)}$ 将只更换 $A^{(1)}$ 的后 $n-1$ 行, 且使其右下角的 $(n-1)$ 阶子矩阵的对角线元素非零。

令 $P = \tilde{P}P(1, i)$, 则 PA 的各对角线元素均非零。

注意: 本题结论为使用古典迭代法求解线性方程组奠定了可行性。对于非奇异线性方程组, 完全可以假定其系数矩阵对角线元素均不为零, 即 D 非奇异。

5. 若 A 是严格对角占优的或不可约对角占优的, 则 G-S 迭代法收敛。

[证明] 若 A 是严格对角占优的或不可约对角占优的, 则必有 $a_{ii} \neq 0$, 因此 $D-L$ 非奇异。

现在来证明：G-S 迭代矩阵 $G = (D - L)^{-1}U$ 的谱半径小于 1。假设 $|\lambda| \geq 1$ ，则

由 A 的假设知， $D - L - \frac{1}{\lambda}U$ 也是严格对角占优或不可约对角占优的，因此

$$\det(D - L - \frac{1}{\lambda}U) \neq 0, \quad \det((D - L)^{-1}) \neq 0$$

而由于

$$\begin{aligned} \det(\lambda I - G) &= \det(\lambda I - (D - L)^{-1}U) = \det\left(\lambda(D - L)^{-1}\left(D - L - \frac{1}{\lambda}U\right)\right) \\ &= \lambda^n \det((D - L)^{-1}) \det(D - L - \frac{1}{\lambda}U) \neq 0 \end{aligned}$$

这说明迭代矩阵 $G = (D - L)^{-1}U$ 不存在模大于等于 1 的特征值。因此 $\rho(G) < 1$ ，从而 G-S 迭代收敛。

6. 设 $A = (a_{ij}) \in R^{n \times n}$ 是严格对角占优的。试证：

$$|\det(A)| \geq \prod_{i=1}^n \left(|a_{ii}| - \sum_{j \neq i} |a_{ij}| \right)$$

[证明] 由 Gerschgorin 定理知，对任意 n 阶复矩阵，其特征值都在复平面上的 n 个圆

$$|z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|, \quad i = 1, \dots, n.$$

的和集内。从而对任意一个特征值均有

$$\begin{aligned} \|\lambda\| - |a_{ii}| &\leq |\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \\ |a_{ii}| - \sum_{j \neq i} |a_{ij}| &\leq |\lambda| \leq |a_{ii}| + \sum_{j \neq i} |a_{ij}| \end{aligned}$$

从而

$$|\det(A)| = |\lambda_1 \lambda_2 \cdots \lambda_n| \geq \prod_{i=1}^n \left(|a_{ii}| - \sum_{j \neq i} |a_{ij}| \right)$$

7. 设 A 是具有正对角元素的非奇异对称矩阵。证明：若求解 $Ax = b$ 的 G-S 迭代方法对任意近似 $x^{(0)}$ 皆收敛，则 A 必为正定的。

[证明] 方程组 $Ax = b$ 的 G-S 迭代如下：

$$x_{k+1} = (D - L)^{-1}Ux_k + (D - L)^{-1}b,$$

由 A 的对称性可知, $U = L^T$, 即 G-S 迭代又可写成

$$x_{k+1} = (D - L)^{-1} L^T x_k + (D - L)^{-1} b$$

若令方程组的精确解为 x^* , 并记 $y_k = x_k - x^*$, 则有

$$(D - L)y_{k+1} = L^T y_k$$

若令 $\varepsilon_k = y_k - y_{k+1}, k \geq 0$. 注意到: $(D - L) = A + L^T$, 则有

$$(D - L)\varepsilon_k = Ay_k, \quad Ay_{k+1} = L^T \varepsilon_k$$

用 y_k^T 和 y_{k+1}^T 分别左乘上两式, 并做差得

$$\begin{aligned} y_k^T Ay_k - y_{k+1}^T Ay_{k+1} &= y_k^T (D - L)\varepsilon_k - y_{k+1}^T L^T \varepsilon_k \\ &= \varepsilon_k^T D y_k - \varepsilon_k^T L^T y_k - \varepsilon_k^T L y_{k+1} \\ &= \varepsilon_k^T D \varepsilon_k + \varepsilon_k^T (D - L)y_{k+1} - \varepsilon_k^T L^T y_k \end{aligned}$$

注意到: $(D - L)y_{k+1} = L^T y_k$, 故得

$$y_k^T Ay_k - y_{k+1}^T Ay_{k+1} = \varepsilon_k^T D \varepsilon_k$$

由题设可知 D 是正定的, 因此, 上式表明误差向量 $\{y_k\}$ 具有某种“减小性”, 即

$$y_k^T Ay_k > y_{k+1}^T Ay_{k+1}, \quad k \geq 0$$

现在来证明若 G-S 迭代收敛, 则 A 是正定的。用反证法: 设 A 不正定, 则可找到一个 $\varepsilon_0 \neq 0$, 使 $\varepsilon_0^T A \varepsilon_0 \leq 0$ 。由题设知: $\rho(L_1) < 1$, 所以 $(I - L_1)$ 非奇异, 于是方程组

$$(I - L_1)y = \varepsilon_0$$

有非零解, 记为 y_0 , 则

$$\varepsilon_0 = y_0 - y_1 = (I - L_1)y_0 \neq 0$$

由假设知

$$0 \geq \varepsilon_0^T A \varepsilon_0 = y_0^T Ay_0 - y_1^T Ay_1$$

于是有

$$y_1^T Ay_1 \geq y_0^T Ay_0$$

这与 $\{y_k\}$ 的减小性相抵触。该矛盾说明 A 是正定的。

8. 若存在对称正定阵 P , 使

$$B = P - H^T P H$$

为对称正定阵, 试证迭代法

$$x_{k+1} = Hx_k + b, \quad k = 0, 1, 2, \dots$$

收敛。

[证明] 设 λ 是 H 的任一特征值, $v \neq 0$ 是 H 关于 λ 的特征向量, 于是

$$v^T B v = v^T P v - v^T H^T P H v = (1 - \lambda^2) v^T P v$$

因 B, P 都是正定阵, 故 $(1 - \lambda^2) > 0$, 即 $|\lambda| < 1$. 由 λ 的任意性得知 $\rho(H) < 1$, 故迭代法收敛。

9. 对 Jacobi 方法引进迭代参数 $\omega > 0$, 即

$$x_{k+1} = x_k - \omega D^{-1}(Ax_k - b)$$

或者

$$x_{k+1} = (I - \omega D^{-1}A)x_k + \omega D^{-1}b$$

称为 Jacobi 松弛法 (简称 JOR 方法). 证明: 当 $Ax = b$ 的 Jacobi 方法收敛时, JOR 方法对 $0 < \omega \leq 1$ 收敛.

[证明] 对于 $Ax = b$, $A = D - L - U$, 则 Jacobi 迭代矩阵 B 和 JOR 迭代矩阵 B_ω 分别是

$$B = D^{-1}(L + U) = I - D^{-1}A,$$

$$B_\omega = I - \omega D^{-1}A.$$

由于 Jacobi 迭代收敛当且仅当 $\rho(B) < 1$, 即 B 的任一特征值 $|\lambda| < 1$. 现设 λ 是 Jacobi 迭代矩阵 B 的一个特征值, 非零向量 v 是其对应的特征向量, 则有

$$\lambda v = Bv = V - D^{-1}Av$$

即有

$$(1 - \lambda)v = D^{-1}Av$$

进而

$$B_\omega v = V - \omega D^{-1}Av = (1 - \omega(1 - \lambda))v$$

即若 λ 是 Jacobi 迭代矩阵 B 的一个特征值, 则 $(1 - \omega(1 - \lambda))$ 便是 B_ω 的一个特征值.

当取定: $0 < \omega \leq 1$, 并假定 $\lambda = \alpha + \beta i$, $\alpha^2 + \beta^2 < 1$, 注意到

$$\begin{aligned} 1 - |1 - \omega(1 - \lambda)|^2 &= 1 - [(1 - \omega(1 - \alpha))^2 + \omega^2 \beta^2] \\ &= 1 - [1 - 2\omega(1 - \alpha) + \omega^2(1 - \alpha)^2 + \omega^2 \beta^2] \\ &= 2\omega(1 - \alpha) - \omega^2 + 2\alpha\omega - \omega^2(\alpha^2 + \beta^2) \\ &> 2\omega(1 - \alpha) - 2\omega^2 + 2\alpha\omega = 2\omega(1 - \omega) \geq 0 \end{aligned}$$

即 B_ω 的所有特征值模小于 1, 从而 $\rho(B_\omega) < 1$, 即 JOR 迭代收敛.

10. 证明: 若 A 是具有正对角元的实对称矩阵, 则 JOR 方法收敛的充分必要条件是 A 及 $2\omega^{-1}D - A$ 均为正定对称矩阵.

[证明] 由于 A 的对角元都是正数, 故 $D = \text{diag}(a_{11}, \dots, a_{nn})$ 的对角元为正数, 故

$$B_\omega = I - \omega D^{-1}A = D^{-\frac{1}{2}}(I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}$$

显然, 矩阵 $(I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}})$ 与 B_ω 相似, 两者有相同的特征值. 同时, 它与 A 有着相同的实对称性. 因此, 两个矩阵的特征值都是实数.

必要性. 设 JOR 迭代收敛, 即 $\rho(B_\omega) < 1$. 那么, 矩阵 $(I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}})$ 的特征值在区间 $(-1, 1)$ 内, 于是得出 $\omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 的特征值位于区间 $(0, 2)$ 内, 这就是说 $\omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 是正定的, 而它与 A 具有相同的正定性, 因此 A 也是正定的.

另外, 实对称矩阵 $(2I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}})$ 的特征值完全由 $(I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}})$ 的特征值所生成, 所以 $(2I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}})$ 的特征值将全部位于区间 $(0, 2)$ 内, 因此是正定的. 注意到

$$2I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}} = \omega D^{-\frac{1}{2}}(2\omega^{-1}D - A)D^{-\frac{1}{2}}, \quad \omega > 0$$

因此矩阵 $2\omega^{-1}D - A$ 也是正定的.

充分性. 一方面, 因为

$$D^{\frac{1}{2}}(I - B_\omega)D^{-\frac{1}{2}} = \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$$

所以 $(I - B_\omega)$ 与 A 一样是正定矩阵. 即 $(I - B_\omega)$ 的特征值均大于 0. 即 B_ω

的特征值均小于 1. 另一方面, 由于 $2\omega^{-1}D - A$ 正定, 而且

$$\omega D^{-\frac{1}{2}}(2\omega^{-1}D - A)D^{-\frac{1}{2}} = D^{\frac{1}{2}}(I + B_{\omega})D^{-\frac{1}{2}}$$

所以, 矩阵 $(I + B_{\omega})$ 是正定的, 即特征值全部为正数, 即 B_{ω} 的特征值均大于 -1 .

结合两方面的结果, 得知: $\rho(B_{\omega}) < 1$, 即 JOR 迭代收敛.

11. 证明: 若系数矩阵 A 是严格对角占优的或不可约对角占优的, 且松弛因子 $\omega \in (0, 1)$, 则 SOR 收敛.

[证明] 若矩阵 A 是严格对角占优的或不可约对角占优的, 则必有 $|a_{ii}| > 0$, 因此 D 非奇异. 现假定某个复数 $|\lambda| \geq 1$, 则矩阵 $(\lambda + \omega - 1)D - \lambda\omega L - \omega U$ 也是严格对角占优的或不可约对角占优的. 不妨假设 $\lambda = \alpha + \beta i$, 且 $\alpha^2 + \beta^2 \geq 1$, 于是就有

$$\begin{aligned} |\lambda + \omega - 1|^2 - |\lambda|^2 \omega^2 &= (\alpha + \omega - 1)^2 + \beta^2 - \omega^2(\alpha^2 + \beta^2) \\ &= \alpha^2 - 2\alpha(1 - \omega) + (1 - \omega)^2 + \beta^2 - \omega^2(\alpha^2 + \beta^2) \\ &= (1 - \omega^2)(\alpha^2 + \beta^2) - 2\alpha(1 - \omega) + (1 - \omega)^2 \\ &= (1 - \omega)[\alpha^2 + \beta^2 + \omega(\alpha^2 + \beta^2) - 2\alpha + 1 - \omega] \\ &= (1 - \omega)[(\alpha - 1)^2 + \beta^2 + \omega(\alpha^2 + \beta^2 - 1)] \\ &\geq 0. \end{aligned}$$

从而

$$|\lambda + \omega - 1|^2 \geq |\lambda|^2 \omega^2 \geq \omega^2$$

因此得到

$$|\lambda + \omega - 1| \geq |\lambda| \omega \geq \omega$$

于是由 A 的严格对角占优或不可约严格对角占优可知

$(\lambda + \omega - 1)D - \lambda\omega L - \omega U$ 也是严格对角占优或不可约对角占优的. 因此,

$(\lambda + \omega - 1)D - \lambda\omega L - \omega U$ 是非奇异的. 而

$$\begin{aligned} \det(\lambda I - L_{\omega}) &= \det(\lambda I - (D - \omega L)^{-1}[(1 + \omega)D + \omega U]) \\ &= \det((D - \omega L)^{-1}) \det((\lambda + \omega - 1)D - \lambda\omega L - \omega U) \\ &\neq 0. \end{aligned}$$

因此, λ 不是 SOR 迭代矩阵 L_ω 的特征值。由 λ 的任意可知, L_ω 的特征值都将满足 $|\lambda| < 1$, 于是 $\rho(L_\omega) < 1$, 从而 SOR 迭代收敛。

12. 证明矩阵

$$A = \begin{bmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{bmatrix}$$

是具有相容次序的。

[证明] 只需按定义验证, 取 $S_1 = \{1\}$, $S_2 = \{2, 3\}$, $S_3 = \{4\}$. 矩阵 A 符合相容次序的定义。

习题 5 共轭梯度法

1. 证明等式 (5.1.4).

[证明] 考虑 $\phi(x) = x^T Ax - 2b^T x$ 在方程组 $Ax = b$ 的解向量 x_* 处的 Taylor 展式, 则有

$$\begin{aligned}\phi(x) &= \phi(x_*) + \nabla \phi(x_*)(x - x_*) + \frac{(x - x_*)^T \nabla^2 \phi(x_*)(x - x_*)}{2} \\ &= x_*^T Ax_* - 2b^T x_* + (x - x_*)^T A(x - x_*)\end{aligned},$$

注意到: $Ax_* = b$, 于是上式可写为

$$\phi(x) + x_*^T Ax_* = (x - x_*)^T A(x - x_*).$$

2. 设 x_k 是由最速下降法产生的. 证明:

$$\phi(x_k) \leq \left(1 - \frac{1}{\kappa_2(A)}\right) \phi(x_{k-1}),$$

其中 $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2$.

[证明] 由 Taylor 展式易知

$$\phi(x_k) = \phi(x_{k-1}) - \frac{(r_{k-1}^T r_{k-1})^2}{r_{k-1}^T A r_{k-1}}, \quad r_{k-1} = b - Ax_{k-1}.$$

注意到:

$$\begin{aligned}r_{k-1}^T A^{-1} r_{k-1} &= (b - Ax_{k-1})^T A^{-1} (b - Ax_{k-1}) \\ &= b^T A^{-1} b + x_{k-1}^T Ax_{k-1} - 2b^T x_{k-1} \\ &= b^T A^{-1} b + \phi(x_{k-1})\end{aligned},$$

由 A 的正定性可知 A^{-1} 是正定的, 因此 $b^T A^{-1} b > 0$, 于是

$$\phi(x_{k-1}) < r_{k-1}^T A^{-1} r_{k-1},$$

从而

$$\begin{aligned}\phi(x_k) &< \left(1 - \frac{(r_{k-1}^T r_{k-1})^2}{(r_{k-1}^T A r_{k-1})(r_{k-1}^T A^{-1} r_{k-1})}\right) \phi(x_{k-1}) \\ &< \left(1 - \frac{1}{\|A\|_2 \|A^{-1}\|_2}\right) \phi(x_{k-1}) = \left(1 - \frac{1}{\kappa_2(A)}\right) \phi(x_{k-1}).\end{aligned}$$

3. 试证明当最速下降法在有限步求得极小值时, 最后一步迭代的下降方向必是 A 的一个特征向量.

[证明] 假定在 $k+1$ 步迭代后, 得到了精确解 $x_{k+1} = x_*$, 即

$$x_* = x_k + \frac{r_k^T r_k}{r_k^T A r_k} r_k,$$

从而有

$$b = Ax_* = Ax_k + \frac{r_k^T r_k}{r_k^T A r_k} A r_k,$$

记: $\lambda = \frac{r_k^T A r_k}{r_k^T r_k}$, 整理可得

$$A r_k = \lambda r_k,$$

即是说 λ 是 A 的一个特征值, r_k 是其对应的特征向量.

4. 证明线性方程组 (5.2.1) 的解存在唯一.

[证明] 为证明 (5.2.1) 的解存在唯一, 只需证明其系数矩阵的行列式不为零. 注意到:

$$r_k = r_{k-1} - \alpha_{k-1} A p_{k-1}, \quad p_k = r_k + \beta_{k-1} p_{k-1},$$

其中

$$\alpha_{k-1} = \frac{r_{k-1}^T p_{k-1}}{p_{k-1}^T A p_{k-1}}, \quad \beta_{k-1} = -\frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}$$

由定理 5.2.1 可得

$$\begin{aligned} p_{k-1}^T A p_{k-1} &= \frac{1}{\alpha_{k-1}} r_{k-1}^T r_{k-1}, \quad \alpha_{k-1} = \frac{r_{k-1}^T r_{k-1}}{p_{k-1}^T A p_{k-1}}, \\ r_k^T A p_{k-1} &= \frac{1}{\alpha_{k-1}} r_k^T (r_{k-1} - r_k) = -\frac{1}{\alpha_{k-1}} r_k^T r_k, \quad \beta_{k-1} = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}}, \\ r_k^T A r_k &= (p_k - \beta_{k-1} p_{k-1})^T A (p_k - \beta_{k-1} p_{k-1}) \\ &= \frac{1}{\alpha_k} r_k^T r_k + \frac{\beta_{k-1}^2}{\alpha_{k-1}} r_{k-1}^T r_{k-1} = \frac{1}{\alpha_k} r_k^T r_k + \frac{(r_k^T r_k)^2}{\alpha_{k-1} r_{k-1}^T r_{k-1}} \end{aligned}$$

为了讨论方便, 我们引入记号 $a = r_k^T r_k$, $b = r_{k-1}^T r_{k-1}$, 则

$$\begin{aligned}\det &= \begin{vmatrix} r_k^T A r_k & r_k^T A p_{k-1} \\ r_k^T A p_{k-1} & p_{k-1}^T A p_{k-1} \end{vmatrix} = \begin{vmatrix} \frac{a}{\alpha_k} + \frac{a^2}{\alpha_{k-1}b} & -\frac{a}{\alpha_{k-1}} \\ -\frac{a}{\alpha_{k-1}} & \frac{b}{\alpha_{k-1}} \end{vmatrix} \\ &= \frac{ab}{\alpha_k \alpha_{k-1}} + \frac{a^2}{\alpha_{k-1}^2} - \frac{a^2}{\alpha_{k-1}^2} \\ &= \frac{ab}{\alpha_k \alpha_{k-1}}\end{aligned}$$

将 $a, b, \alpha_k, \alpha_{k-1}$ 代入后, 得

$$\det = p_k^T A p_k \cdot p_{k-1}^T A p_{k-1} \neq 0.$$

5. 设 $A \in R^{n \times n}$ 对称正定的, $p_1, \dots, p_k \in R^n$ 是互相共轭正交的, 即 $p_i^T A p_j = 0, i \neq j$. 证明 p_1, \dots, p_k 是线性无关的.

〔证明〕若有一组数 $\alpha_1, \dots, \alpha_k$ 满足

$$\alpha_1 p_1 + \dots + \alpha_k p_k = 0,$$

则对一切 $i = 1, \dots, k$ 一定有

$$0 = p_i^T A (\alpha_1 p_1 + \dots + \alpha_k p_k) = \alpha_i p_i^T A p_i,$$

注意到 $p_i^T A p_i \neq 0$, 由此得出: $\alpha_i = 0$. 即所有的 $\alpha_i = 0$. 因此, p_1, p_2, \dots, p_k 是线性无关的.

6. 设 A 为对称正定矩阵, 从方程组的近似解 $y_0 = x_k$ 出发, 依次求 y_i 使得

$$\phi(y_i) = \min_t \phi(y_{i-1} + t e_i),$$

其中 e_i 是 n 阶单位矩阵的第 i 列, $i = 1, 2, \dots, n$. 然后令 $x_{k+1} = y_n$. 验证这样得到的迭代算法就是G-S迭代法.

〔证明〕在下面的讨论中, 我们用 y_i 表示第 i 迭代的向量, $y_i[j]$ 表示 y_i 的第 j 个分量.

第一步: 从 $y_0 = x_k$ 出发, 沿 e_1 方向搜索得新的极小值点 y_1 , 则

$$y_1 = y_0 + \alpha_0 e_1,$$

其中

$$\begin{aligned}\alpha_0 &= r_0^T e_1 / e_1^T A e_1 = (b - Ay_0)^T e_1 / a_{11} \\ &= \left(b[1] - \sum_{i=1}^n a_{1i} y_0(i) \right) / a_{11},\end{aligned}$$

从而

$$\begin{aligned}y_1[1] &= \left(b[1] - \sum_{i=2}^n a_{1i} y_0[i] \right) / a_{11} \\ y_1[j] &= y_0[j], \quad j = 2, \dots, n;\end{aligned}$$

完成第一步后, 可以看出 $y_1[1]$ 与直接从 $y_0 = x_k$ 做 G - S 迭代一步所得的第一个分量相同.

现在考虑第 m 迭代. 假定 y_{m-1} 的前 $m-1$ 个分量符合 G - S 迭代形式, 现从 y_{m-1} 出发, 沿 e_m 进行极小化搜索, 得极小值点

$$y_m = y_{m-1} + \alpha_{m-1} e_m,$$

其中

$$\alpha_{m-1} = (b - Ay_{m-1})^T e_m / e_m^T A e_m = \left(b[m] - \sum_{i=1}^n a_{mi} y_{m-1}[i] \right) / a_{mm}.$$

从而

$$\begin{aligned}y_m[m] &= y_{m-1}[m] + \left(b[m] - \sum_{\substack{i=1 \\ i \neq m}}^n a_{mi} y_{m-1}[i] \right) / a_{mm}, \\ y_m[j] &= y_{m-1}[j], \quad j = 1, \dots, m-1, m+1, \dots, n.\end{aligned}$$

这里 $m = 2, \dots, n$.

显然令 $x_{k+1} = y_n$, 则 x_{k+1} 恰与对 x_k 经一次 G - S 迭代后的近似解完全一致.

7. 设 A 是一个只有 k 个互不相同的特征值的 $n \times n$ 实对称矩阵, r 是任一 n 维实向量. 证明: 子空间

$$\text{span}\{r, Ar, \dots, A^{n-1}r\}$$

的维数至多是 k .

[证明] 由于 A 是实对称矩阵, 因此存在一个完全实特征向量系. 不妨设 A 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_k$ 其重数为 m_1, m_2, \dots, m_k , $m_1 + \dots + m_k = n$, 且设 A 关于 λ_i 的特

征向量为 $x_1^{(i)}, \dots, x_{m_i}^{(i)}$, $i = 1, \dots, k$ 是 A 的单位正交实特征向量系, 于是对任一 n 维实向量 r 有

$$r = \sum_{i=1}^k \sum_{j=1}^{m_i} \alpha_{ij} x_j^{(i)}, \quad \alpha_{ij} \in \mathbb{R}.$$

从而

$$A^p r = \sum_{i=1}^k \sum_{j=1}^{m_i} \alpha_{ij} \lambda_i^p x_j^{(i)}, \quad p = 1, \dots, n-1.$$

现在用归纳法证明 $\text{span}\{r, Ar, \dots, A^{n-1}r\}$ 的维数至多是 k .

当 $k=1$, 假定存在 n 个实数 β_1, \dots, β_n 使

$$0 = \beta_1 r + \beta_2 Ar + \dots + \beta_n A^{n-1}r = \left(\alpha_{11} \sum_{i=1}^n \beta_i \lambda_1^{i-1} \right) x_1^{(1)} + \dots + \left(\alpha_{1n} \sum_{i=1}^n \beta_i \lambda_1^{i-1} \right) x_n^{(1)}.$$

在上式两边同乘以 $(x_j^{(1)})^T$, 则得到

$$\alpha_{1j} \sum_{i=1}^{n-1} \lambda_1^{i-1} \beta_i = 0, \quad j = 1, \dots, n.$$

将其看作是 β_1, \dots, β_n 的方程, 则系数矩阵为

$$\begin{bmatrix} \alpha_{11} & \alpha_{11}\lambda_1 & \dots & \alpha_{11}\lambda_1^{n-1} \\ \alpha_{12} & \alpha_{12}\lambda_1 & \dots & \alpha_{12}\lambda_1^{n-1} \\ \vdots & \vdots & \dots & \vdots \\ \alpha_{1n} & \alpha_{1n}\lambda_1 & \dots & \alpha_{1n}\lambda_1^{n-1} \end{bmatrix},$$

显然, 它是一个秩不超过 1 的矩阵. 因此, 该方程的解系的自由变量至少有 $n-1$

个. 这说明 $\text{span}\{r, Ar, \dots, A^{n-1}r\}$ 的维数为至多为 1.

习题 6 非对称特征问题的计算方法

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

$$\lambda(BA) = \lambda(AB) \cup \underbrace{\{0, \dots, 0\}}_{m-n}$$

[证明] 令:

$$\tilde{A} = \begin{bmatrix} A \\ 0 \end{bmatrix} \in R^{m \times m}, \quad \tilde{B} = [B, 0] \in R^{m \times m}.$$

则

$$\tilde{A}\tilde{B} = \begin{bmatrix} AB & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{B}\tilde{A} = BA.$$

显然, 欲证本题结论, 只需证明:

$$\lambda(\tilde{A}\tilde{B}) = \lambda(\tilde{B}\tilde{A}).$$

假设 $\lambda \in \lambda(\tilde{A}\tilde{B})$, 分两种情况讨论:

(1) $\lambda \neq 0$ 是 $\tilde{A}\tilde{B}$ 的一个特征值, 非零 m 维向量 x 是 $\tilde{A}\tilde{B}$ 关于 λ 的特征向量, 即有 $\tilde{A}\tilde{B}x = \lambda x$, 显然 $\tilde{B}x \neq 0$, 于是有 $\tilde{B}\tilde{A}\tilde{B}x = \lambda\tilde{B}x$, 此式表明 λ 是 $\tilde{B}\tilde{A}$ 的特征值, 对应的特征向量为 $\tilde{B}x$. 从而有 $\lambda \in \lambda(\tilde{B}\tilde{A})$.

(2) $\lambda = 0$, 则

$$\det(\tilde{A}\tilde{B}) = \det(\tilde{B}\tilde{A}) = 0.$$

即

$$\lambda \in \lambda(\tilde{B}\tilde{A}).$$

于是得知:

$$\lambda(\tilde{A}\tilde{B}) \subset \lambda(\tilde{B}\tilde{A}).$$

同理可证:

$$\lambda(\tilde{A}\tilde{B}) \supset \lambda(\tilde{B}\tilde{A}).$$

从而

$$\lambda(\tilde{A}\tilde{B}) = \lambda(\tilde{B}\tilde{A}).$$

由 $\tilde{A}\tilde{B}$ 和 $\tilde{B}\tilde{A}$ 的定义知:

$$\lambda(\tilde{A}\tilde{B}) = \lambda(AB) \cup \underbrace{\{0, \dots, 0\}}_{n-n}.$$

至此，命题得证.

2. 设 $\lim_{k \rightarrow \infty} A_k = A$, $A_k = Q_k T_k Q_k^*$ 是 A_k 的 Schur 分解. 证明: $\{Q_k\}$ 有收敛的子序列 $\{Q_{k_j}\}$; 若记 $Q = \lim_{k \rightarrow \infty} Q_{k_j}$, 则有 $Q^* A Q$ 是上三角矩阵.

[证明] 由于 $A_k = Q_k T_k Q_k^*$ 是 A_k 的 Schur 分解, 因此 $T_k = Q_k^* A_k Q_k$ 是上三角矩阵, 因此, $\lim_{k \rightarrow \infty} T_k = \lim_{k \rightarrow \infty} Q_k^* A_k Q_k = Q^* A Q$ 仍为上三角矩阵.

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

[证明] 根据题设, 既然 A 没有重特征值, 不妨设 A 的特征值如下排列

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| \geq 0.$$

再由 Schur 分解定理, 不妨设定

$$T = \begin{bmatrix} \Sigma & t \\ 0 & \lambda_n \end{bmatrix}, \quad \Sigma \in C^{(n-1) \times (n-1)} \text{ 非奇异}, \quad t \in C^{n-1}$$

再将 $A = Q T Q^*$ 代入 $AB = BA$ 中, 整理可得

$$Q^* B Q T = T Q^* A Q$$

为叙述方便, 记 $X = Q^* B Q$. 至此我们只需证明: 对于一个 n 阶矩阵 X , 如果 $TX = XT$, 则 X 必是上三角矩阵.

用归纳法证明. 当 $n = 2$ 时, 设

$$T = \begin{bmatrix} \lambda_1 & t \\ 0 & \lambda_2 \end{bmatrix}, \quad X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$

于是有

$$TX = \begin{bmatrix} \lambda_1 x_{11} + t x_{21} & \lambda_1 x_{12} + t x_{22} \\ \lambda_2 x_{21} & \lambda_2 x_{22} \end{bmatrix}, \quad XT = \begin{bmatrix} x_{11} \lambda_1 & x_{11} t + x_{12} \lambda_2 \\ x_{21} \lambda_1 & x_{21} t + x_{22} \lambda_2 \end{bmatrix}$$

由于 $TX = XT$, 则对应地有: $\lambda_2 x_{21} = x_{21} \lambda_1$, 又因 $\lambda_1 \neq \lambda_2$, 故得 $x_{21} = 0$. 即 X 是上三角矩阵.

归纳假设, 当 T 和 X 为 $n-1$ 阶矩阵时结论成立. 现在来证明 n 阶的情况, 我们将

X 写成如下块形状

$$X = \begin{bmatrix} X_{11} & x_{12} \\ x_{21} & x_{nn} \end{bmatrix}, \quad \text{其中 } X_{11} \in C^{(n-1) \times (n-1)}, \quad x_{12} \in C^{n-1}, \\ x_{21}^T \in C^{n-1}, \quad x_{nn} \in C.$$

于是便有

$$TX = \begin{bmatrix} \Sigma & t \\ 0 & \lambda_{nn} \end{bmatrix} \begin{bmatrix} X_{11} & x_{12} \\ x_{21} & x_{nn} \end{bmatrix} = \begin{bmatrix} \Sigma X_{11} + tx_{21} & \Sigma x_{12} + tx_{nn} \\ \lambda_{nn} x_{21} & \lambda_{nn} x_{nn} \end{bmatrix}, \\ XT = \begin{bmatrix} X_{11} & x_{12} \\ x_{21} & x_{nn} \end{bmatrix} \begin{bmatrix} \Sigma & t \\ 0 & \lambda_{nn} \end{bmatrix} = \begin{bmatrix} X_{11}\Sigma & X_{11}t + x_{12}\lambda_{nn} \\ x_{21}\Sigma & x_{nn}\lambda_{nn} \end{bmatrix}.$$

因 $TX = XT$, 故应有 $\lambda_{nn} x_{21} = x_{21}\Sigma$, 于是

$$(\lambda_{nn} I - \Sigma^T) x_{21}^T = 0,$$

由题设不难看出 $(\lambda_{nn} I - \Sigma^T)$ 非奇异, 因此 $x_{21} = 0$. 从而 $X = \begin{bmatrix} X_{11} & x_{12} \\ 0 & x_{nn} \end{bmatrix}$, 进而得知

$$X_{11}\Sigma = \Sigma X_{11}$$

由归纳假设知 X_{11} 是上三角矩阵. 从而证得 X 是上三角矩阵.

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

$$R(x) = x^* Ax / x^* x,$$

称之为 x 对 A 的 Rayleigh 商. 证明对任意的 $x \in C^n (x \neq 0)$ 有

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

即 Rayleigh 商有极小剩余性.

[证明] 使 $\|Ax - \mu x\|_2$ 达到极小的问题是一个关于未知量 μ 的线性最小二乘问题. 它的正规方程即是

$$x^* x \mu = x^* Ax,$$

其解为

$$\mu_* = R(x) = x^* Ax / x^* x.$$

从而

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

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

[解] 显然 α 和 β 都是单特征值.

对于 α 来说, 显然 $x = (1, 0)^T$ 是 A 关于 α 的一个模 1 特征向量. 同时, 容易求得 $y = (1, 0)^T$ 是 A 关于 α 的满足 $y^T x = 1$ 的左特征向量, 故由特征值条件数的定义得知

$$\text{cond}(\alpha) = 1.$$

对于 β 来说, 解方程

$$\begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \beta \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

得到 A 关于 β 的特征向量

$$x = \left(\frac{k\gamma}{\beta - \alpha}, k \right)^T, \quad k \in C, k \neq 0.$$

当 $k = |\beta - \alpha| / \sqrt{(\beta - \alpha)^2 + \gamma^2}$ 时, $\|x\|_2 = 1$.

再由方程 $y^T A = \beta y^T$ 可解得 A 关于 β 的左特征向量 $y = (0, t)^T, t \neq 0$, 令 $y^T x = 1$, 则得出

$$t = \frac{1}{k} = \frac{\sqrt{(\beta - \alpha)^2 + \gamma^2}}{|\beta - \alpha|}$$

从而由特征值条件数的定义知

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

7. 分别应用幂法于矩阵

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

并考察所得序列的特性.

[解] 我们不妨设 $\lambda > 0$. 对于矩阵

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

即初始向量: $u_0 = (1, 1)^T$, 迭代如下:

$$y_1 = Au_0 = \begin{bmatrix} \lambda + 1 \\ \lambda \end{bmatrix}, \quad m_1 = \max(y_1) = \lambda + 1, \quad u_1 = y_1 / m_1 = \begin{bmatrix} 1 \\ \frac{\lambda}{\lambda + 1} \end{bmatrix};$$

$$y_2 = Au_1 = \begin{bmatrix} \lambda + \frac{\lambda}{\lambda + 1} \\ \frac{\lambda^2}{\lambda + 1} \end{bmatrix}, \quad m_2 = \max(y_2) = \lambda + \frac{\lambda}{\lambda + 1}, \quad u_2 = y_2 / m_2 = \begin{bmatrix} 1 \\ \frac{\lambda}{\lambda + 2} \end{bmatrix};$$

$$y_3 = Au_2 = \begin{bmatrix} \lambda + \frac{\lambda}{\lambda + 2} \\ \frac{\lambda^2}{\lambda + 2} \end{bmatrix}, \quad m_3 = \max(y_3) = \lambda + \frac{\lambda}{\lambda + 2}, \quad u_3 = y_3 / m_3 = \begin{bmatrix} 1 \\ \frac{\lambda}{\lambda + 3} \end{bmatrix};$$

一般地,

$$m_k = \lambda + \frac{\lambda}{\lambda + k - 1}, \quad u_k = \begin{bmatrix} 1 \\ \frac{\lambda}{\lambda + k} \end{bmatrix}, \quad k = 1, 2, \dots.$$

显然

$$m_k \rightarrow \lambda, \quad u_k \rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad k \rightarrow \infty.$$

对于矩阵

$$B = \begin{bmatrix} \lambda & 1 \\ 0 & -\lambda \end{bmatrix}$$

取初始向量: $u_0 = (1, 1)^T$, 则迭代得到

$$y_1 = Bu_0 = (\lambda + 1, -\lambda)^T, \quad m_1 = \max(y_1) = \lambda + 1, \quad u_1 = \left(1, -\frac{\lambda}{\lambda + 1}\right)^T;$$

$$y_2 = Bu_1 = \left(\lambda - \frac{\lambda}{\lambda + 1}, \frac{\lambda^2}{\lambda + 1}\right)^T, \quad m_2 = \max(y_2) = \frac{\lambda^2}{\lambda + 1}, \quad u_2 = (1, 1)^T;$$

由此我们可以看出，迭代数列： $\{m_k\}$ ，和向量序列： $\{u_k\}$ 均不收敛，但它们都各对应地存在两个收敛子列。
 即当脚标为奇数时，

$$m_k \equiv \lambda + 1, \quad u_k \equiv \left(1, -\frac{\lambda}{\lambda+1}\right)^T, \quad k=1,3,5,\dots;$$

当脚标为偶数时，

$$m_k \equiv \frac{\lambda^2}{\lambda+1}, \quad u_k \equiv (1, 1)^T, \quad k=2,4,6,\dots$$

8. 在幂法中，取

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad u_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

得到一个精确到 5 位数字的特征向量需要多少次迭代？

[解] 分析幂法可知，由算法得到的向量序列 $\{u_k\}$ 满足：

$$y_k = A^k u_0, \quad u_k = y_k / \max(y_k), \quad k=1,2,\dots$$

注意到：

$$A^k = \begin{bmatrix} 1 & C_k^1 & C_k^2 \\ & 1 & C_k^1 \\ & & 1 \end{bmatrix}, \quad k > 1.$$

于是得知

$$y_k = A^k u_0 = \begin{bmatrix} C_k^2 \\ C_k^1 \\ 1 \end{bmatrix}, \quad u_k = \frac{y_k}{\max(y_k)} = \begin{bmatrix} C_k^2 \\ C_k^1 \\ 1 \end{bmatrix} \bigg/ C_k^2 = \begin{bmatrix} 1 \\ \frac{2}{k-1} \\ \frac{2}{k(k-1)} \end{bmatrix}, \quad k > 1,$$

欲使计算结果精确到 5 位数字，只需

$$\frac{2}{k-1} < 10^{-5}$$

解之得： $k > 2 \times 10^5 + 1$.

9. 设 $A \in C^{n \times n}$ 有实特征值并满足 $\lambda_1 > \lambda_2 \geq \dots \geq \lambda_{n-1} > \lambda_n$ ，现应用幂法于矩阵

$A - \mu I$ ，试证：选择 $\mu = \frac{1}{2}(\lambda_2 + \lambda_n)$ 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。

[证明] 由于特征值满足 $\lambda_1 > \lambda_2 \geq \dots \geq \lambda_{n-1} > \lambda_n$, 因此不论 μ 如何, $A - \mu I$ 的按模最大的特征值为 $\lambda_1 - \mu$ 或 $\lambda_n - \mu$, 当我们希望计算 λ_1 和 u_1 时, 首先就选择 μ 使

$$|\lambda_1 - \mu| > |\lambda_n - \mu|$$

且使收敛速度的比值

$$\omega = \max \left\{ \frac{|\lambda_2 - \mu|}{|\lambda_1 - \mu|}, \frac{|\lambda_n - \mu|}{|\lambda_1 - \mu|} \right\} = \min$$

显然, 当 $\lambda_1 - \mu = -(\lambda_n - \mu)$, 即 $\mu = \frac{1}{2}(\lambda_1 + \lambda_n)$ 时, 收敛速度的比值最小, 即收敛速度最快.

10. 应用幂法给出求多项式

$$p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0$$

之模最大根的一种算法。

[解] 根据拉普拉斯展开公式得知

$$\begin{aligned} p(z) &= \begin{vmatrix} 1 & a_1 & a_2 & \dots & a_n \\ -1 & z & & & \\ & -1 & z & & \\ & & \ddots & \ddots & \\ & & & -1 & z \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & z+a_1 & a_2 & \dots & a_n \\ & -1 & z & & \\ & & \ddots & \ddots & \\ & & & -1 & z \end{vmatrix} \\ &= \begin{vmatrix} z+a_1 & a_2 & \dots & a_n \\ -1 & z & & \\ & \ddots & \ddots & \\ & & -1 & z \end{vmatrix} \end{aligned}$$

若令

$$A = \begin{vmatrix} -a_1 & -a_2 & \dots & -a_n \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{vmatrix}$$

则有 $p(z) = \det(zI - A)$ 。这样, 我们就可以针对矩阵 A 利用幂法求得多项式的一个模最大特征值。

1 1. 利用反幂法计算矩阵

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

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

[解] 取初始向量: $v_0 = (1,1,1)^T$, 位移取为 $\tilde{\lambda} = 1.2679$, 用反幂法迭代一次, 即先求解方程组

$$\begin{bmatrix} 0.7321 & 1 & 0 \\ 1 & 1.7321 & 1 \\ 0 & 1 & 2.7321 \end{bmatrix} \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

解得: $\zeta_1 = 6776.39885$, $\zeta_2 = -4960.001597$, $\zeta_3 = 1815.81992$;

单位化得

$$\zeta_1 = 1.0000, \quad \zeta_2 = -0.7320, \quad \zeta_3 = 0.2680.$$

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

[证明] 根据题设可知, v 和 u 都是非零向量, 因此存在 Householder 变换 H 使

$$H \left(\frac{v}{\|v\|_2} \right) = \frac{u}{\|u\|_2}.$$

令

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

由于 H 是正交的, 故 $\|H\|_F = 1$, 因此

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

同时有

$$(A + E)v = Av + Ev = Av + u = Av + \lambda v - Av = \lambda v.$$

即向量 $v = (\lambda I - A)^{-1}u$ 是 $A + E$ 关于其特征值 λ 的一个特征向量.

1 3. 设 $A, E \in C^{n \times n}$, 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 $u, v \in C^n$ 使得

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

[证明] 设非零向量 $v \in C^n$ 为 $A+E$ 关于 λ 的特征向量, 即

$$(A+E)v = \lambda v,$$

再取: $u = \lambda v - Av$, 则 $Ev = u$, 从而 $\|u\|_2 = \|Ev\|_2 \leq \|E\|_2 \|v\|_2$, 即

$$\frac{\|u\|_2}{\|v\|_2} \leq \|E\|_2.$$

1.4. 应用基本的 Q R 迭代于矩阵

$$A = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix},$$

并考察所得的矩阵序列的特点, 并判断该矩阵序列是否收敛?

[解] 用 Givens 正交变换实现 Q R 迭代:

第 1 步: $r = \sqrt{2}, \quad c = \frac{\sqrt{2}}{2}, \quad s = \frac{\sqrt{2}}{2},$

$$A_1 = G_1 A G_1^T \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{3}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix},$$

第 2 步: $r = \frac{\sqrt{2}}{2}, \quad c = \frac{\sqrt{2}}{2}, \quad s = -\frac{\sqrt{2}}{2},$

$$A_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{3}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}.$$

由上面的两步迭代, 即已说明对该矩阵进行 Q R 迭代, 其矩阵序列由两个矩阵构成其奇数项和偶数项, 因此该矩阵序列不收敛.

1.5. 设 $A \in R^{n \times n}$. 证明: 存在初等置换矩阵 P 和初等下三角矩阵 M 使得

$(MP)A(MP)^{-1}$ 有如下形式

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

[证明] 若 A 的第一列除第一个非分量外, 其余各分量为零, 则结论显然成立. 不妨假设 A 的第一列除第一个分量外, 至少还有一个分量非零, 并设 $a_{i1} \neq 0 (i > 1)$, 那么我们取

$$P = I_{2,i},$$

从而有

$$PAP^{-1} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ 0 & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

其中: $a_{21} \neq 0$. 继而构造 Gauss 变换

$$M = I - l_2 e_2^T,$$

其中: $l_2 = (0, 0, a_{31}/a_{21}, \dots, a_{n1}/a_{21})^T$. 则不难验证 $(MP)A(MP)^{-1}$ 具有题中的形式.

1.6. 依照 1.5 即可完成.

1.7. 设 $A \in C^{n \times n}$, $x \in C^n$, $X = [x, Ax, \dots, A^{n-1}x]$. 证明: 如果 X 是非奇异的, 则 $X^{-1}AX$ 是上 Hessenberg 矩阵.

[证明] 由假设可知, X 的 n 个列向量 $x, Ax, \dots, A^{n-1}x$ 是线性无关的. 我们令

$$M = [m_{ij}] = X^{-1}AX$$

则

$$AX = XM$$

现在我们来考察上式两边矩阵的前 $n-1$ 列

$$AXe_i = XMe_i, \quad i = 1, 2, \dots, n-1.$$

则有

$$A^i x = m_{1i} x + m_{2i} Ax + \dots + m_{ni} A^{n-1} x,$$

由于 $x, Ax, \dots, A^{n-1}x$ 是线性无关的, 因此

$$m_{1i} = \dots = m_{ii} = m_{i+2,i} = \dots = m_{ni} = 0, \quad m_{i+1,i} = 1.$$

从而得知

$$M = X^{-1}AX = \begin{bmatrix} 0 & 0 & \dots & 0 & m_{1n} \\ 1 & 0 & \dots & 0 & m_{2n} \\ & \ddots & \ddots & \vdots & \vdots \\ & & \ddots & 0 & m_{n-1,n} \\ & & & 1 & m_{n,n} \end{bmatrix}.$$

1.8. 证明: 若 H 是一个非亏损的上 Hessenberg 矩阵. 则 H 没有重特征值.

[证明] 设 λ_i 是 H 的一个特征值, 则由于 H 是不可约的, 即次对角线上的各元素 $h_{k+1,k} \neq 0$, 故 $A - \lambda_i I$ 的秩为 $n-1$, 因此方程组

$$(A - \lambda_i I)x = 0,$$

的基础解系的秩为 1, 即 H 对应于 λ_i 的特征向量都是线性相关的. 又 H 是非亏损的, 因此, H 将有 n 个线性无关的特征向量. 而每个特征值只对应存在一个线性无关的特征向量, 因此 H 必具有 n 个互不相等的特征值. 即 H 没有重特征值.

1.9. 设 H 是一个为可约的上 Hessenberg 矩阵. 证明存在一个对角矩阵 D 使得 $D^{-1}HD$ 的次对角元素均为 1. $k_2(D) = \|D\|_2 \|D^{-1}\|_2$ 是多少?

[证明] 令 $D = \text{diag}(d_1, d_2, \dots, d_n)$, $D^{-1} = \text{diag}(d_1^{-1}, d_2^{-1}, \dots, d_n^{-1})$. 则由矩阵乘积的定义可知 $D^{-1}HD$ 的次对角元素依次为

$$\frac{d_1}{d_2} h_{21}, \quad \frac{d_2}{d_3} h_{32}, \quad \dots, \quad \frac{d_{n-1}}{d_n} h_{n,n-1}.$$

令其均等于 1, 则可得到

$$d_2 = h_{21}d_1, \quad d_3 = h_{32}h_{21}d_1, \quad \dots, \quad d_n = h_{n,n-1}h_{n-1,n-2} \dots h_{21}d_1.$$

令 $d_1 = 1$, 则 $d_2 = h_{21}$, $d_3 = h_{32}h_{21}$, \dots , $d_n = h_{n,n-1} \dots h_{21}$. 从而

$$D = \text{diag}(1, h_{21}, h_{32}h_{21}, \dots, h_{n,n-1} \dots h_{21}),$$

$$D^{-1} = \text{diag}(1, h_{21}^{-1}, h_{32}^{-1}h_{21}^{-1}, \dots, h_{n,n-1}^{-1} \dots h_{21}^{-1}).$$

记 $m = \min \{1, h_{21}, h_{32}h_{21}, \dots, h_{n,n-1} \dots h_{21}\}$, $M = \max \{1, h_{21}, h_{32}h_{21}, \dots, h_{n,n-1} \dots h_{21}\}$, 则

$$\kappa_2(D) = \frac{M}{m}.$$

2 0. 设 $H \in R^{n \times n}$ 是一个上 Hessenberg 矩阵, 并假定 $x \in R^n$ 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

$$Q^T H Q = \begin{bmatrix} \lambda & w^T \\ 0 & H_1 \end{bmatrix},$$

其中 H_1 是 $n-1$ 阶上 Hessenberg 矩阵.

[解] 不妨假定 $\|x\|_2 = 1$, 以 x 作为第一列构造正交矩阵

$$Q_0 = [x, \tilde{Q}_0], \quad \tilde{Q}_0 \in R^{n \times (n-1)},$$

显然 $x^T \tilde{Q}_0 = 0$, $\tilde{Q}_0^T x = 0$, 从而便有

$$\begin{aligned} Q_0^T H Q_0 &= \begin{bmatrix} x^T \\ \tilde{Q}_0^T \end{bmatrix} H \begin{bmatrix} x & \tilde{Q}_0 \end{bmatrix} = \begin{bmatrix} x^T H \\ \tilde{Q}_0^T H \end{bmatrix} \begin{bmatrix} x & \tilde{Q}_0 \end{bmatrix} \\ &= \begin{bmatrix} \lambda & x^T H \tilde{Q}_0 \\ \tilde{Q}_0^T H x & \tilde{Q}_0^T H \tilde{Q}_0 \end{bmatrix} = \begin{bmatrix} \lambda & x^T H \tilde{Q}_0 \\ 0 & \tilde{Q}_0^T H \tilde{Q}_0 \end{bmatrix} \end{aligned}$$

其中 $\tilde{Q}_0^T H \tilde{Q}_0 \in R^{(n-1) \times (n-1)}$.

将算法 6.4.1 用于 $n-1$ 阶矩阵 $\tilde{Q}_0^T H \tilde{Q}_0$, 得

$$H_1 = \tilde{Q}^T \tilde{Q}_0^T H \tilde{Q}_0 \tilde{Q},$$

再令

$$Q = Q_0 \begin{bmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{bmatrix},$$

为本题所求的正交矩阵, 即

$$Q^T H Q = \begin{bmatrix} \lambda & w^T \\ 0 & H_1 \end{bmatrix}.$$

2.1. 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.

[证明] 由于 H 是奇异的不可约上 Hessenberg 矩阵, 所以它的秩为 $n-1$. 且由定理 3.3.1 可知, H 的 QR 分解如下

$$H = QR = \begin{bmatrix} R_1 \\ 0 \end{bmatrix},$$

其中 $R_1 \in R^{(n-1) \times n}$ 是上三角矩阵, 其对角元素均为正数. 于是

$$\tilde{H} = RQ = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} Q = \begin{bmatrix} R_1 Q \\ 0 \end{bmatrix}.$$

由 QR 算法可知, \tilde{H} 仍是上 Hessenberg 矩阵, 这样零特征值便出现了 \tilde{H} 的右下角.

2.2. 证明: 若给定 $H = H_0$, 并由

$$H_k - \mu_k I = U_k R_k \quad \text{和} \quad H_{k+1} = R_k U_k + \mu_k I$$

产生矩阵序列 $\{H_k\}$, 则

$$(U_0 \dots U_j)(R_j \dots R_0) = (H - \mu_0 I) \dots (H - \mu_j I).$$

[证明] 先来证明下面的等式关系:

$$R_p(H_p - \mu_q I) = (H_{p+1} - \mu_q I)R_p, \quad q > p.$$

事实上,

$$\begin{aligned} R_p(H_p - \mu_q I) &= R_p(U_p R_p + \mu_p I - \mu_q I) \\ &= (R_p U_p + \mu_p I - \mu_q I)R_p \\ &= (H_{p+1} - \mu_q I)R_p \end{aligned}$$

特别地, 当 $q = p+1$ 时,

$$R_p(H_p - \mu_{p+1} I) = U_{p+1} R_{p+1} R_p.$$

下面我们用归纳法证明本题结论.

当 $j=0$ 时, 结论显然成立. 现归纳假设 $j=k$ 时, 结论仍然成立, 即有

$$(U_0 \dots U_k)(R_k \dots R_0) = (H - \mu_0 I) \dots (H - \mu_k I).$$

现在来证明对于 $j = k+1$ 时, 结论仍然成立. 事实上,

$$\begin{aligned} & (H - \mu_0 I) \dots (H - \mu_k I) (H - \mu_{k+1} I) \\ &= (U_0 \dots U_k) (R_k \dots R_0) (H - \mu_{k+1} I) \\ &= (U_0 \dots U_k) (R_k \dots R_1) (H_1 - \mu_{k+1} I) R_0 \\ &= \dots \\ &= (U_0 \dots U_k) R_k (H_k - \mu_{k+1} I) R_{k-1} \dots R_0 \\ &= (U_0 \dots U_k U_{k+1}) (R_{k+1} R_k \dots R_0) \end{aligned}$$

2.3. 设 $A \in R^{n \times n}$ 是一具有互不相同对角元素的上三角矩阵, 给出计算 A 的全部特征向量的详细算法.

[解] 根据题设, A 的对角元素恰是 n 个互不相等的特征值. 下面给出计算 A 关于特征值 r_{kk} 的特征向量方法. 记 A 关于特征值 r_{kk} 的特征向量为

$x_k = (\xi_{1k}, \xi_{2k}, \dots, \xi_{nk})^T$, 它是下列方程组的非零解向量

$$(A - r_{kk} I)x = 0.$$

为讨论其算法, 我们记

$$A = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \dots & r_{1n} \\ & r_{22} & r_{23} & \dots & r_{2n} \\ & & r_{33} & \dots & r_{3n} \\ & & & \ddots & \vdots \\ & & & & r_{nn} \end{bmatrix},$$

我们将 $(A - r_{kk} I)$ 和 x 做如下分块

$$A - r_{kk} I = \begin{bmatrix} R_{11} & v & R_{12} \\ & 0 & w^T \\ & & R_{22} \end{bmatrix} \begin{matrix} k-1 \\ 1 \\ n-k-1 \end{matrix}, \quad x = \begin{bmatrix} x_1 \\ \xi_{kk} \\ x_2 \end{bmatrix} \begin{matrix} k-1 \\ 1 \\ n-k-1 \end{matrix}$$

这里

$$R_{11} = \begin{bmatrix} r_{11} - r_{kk} & \dots & r_{1,k-1} \\ & \ddots & \vdots \\ & & r_{k-1,k-1} - r_{kk} \end{bmatrix}$$

是一个非奇异的 $k-1$ 阶上三角矩阵;

$$v = [r_{1k}, \dots, r_{k-1,k}]^T \in R^{k-1}; \quad w^T = [r_{k,k+1}, \dots, r_{k,n}];$$

$$R_{22} = \begin{bmatrix} r_{k+1,k+1} - r_{kk} & \cdots & r_{k+1,n} \\ & \ddots & \vdots \\ & & r_{n,n} - r_{kk} \end{bmatrix} \in R^{(n-k-1) \times (n-k-1)}$$

是一个非奇异矩阵;

$$R_{12} = \begin{bmatrix} r_{1,k+1} & \cdots & r_{1,n} \\ \vdots & & \vdots \\ r_{k-1,k+1} & \cdots & r_{k-1,n} \end{bmatrix} \in R^{(k-1) \times (n-k-1)};$$

$$x_1 = (\xi_{1k}, \dots, \xi_{k-1,k})^T, \quad x_2 = (\xi_{k+1,k}, \dots, \xi_{n,k})^T.$$

从而, 方程 $(A - r_{kk}I)x = 0$ 等价于

$$\begin{cases} R_{11}x_1 + \xi_{kk}v + R_{12}x_2 = 0 \\ w^T x_2 = 0 \\ R_{22}x_2 = 0 \end{cases}$$

显然, $x_2 = 0$, $x_1 = \xi_{kk}R_{11}^{-1}v$.

根据以上的分析, 我们得到计算 A 关于特征值 r_{kk} 的特征向量

$x_k = (\xi_{1k}, \xi_{2k}, \dots, \xi_{nk})^T$ 的算法步骤如下:

(1) 任取 $\xi_{kk} \neq 0$;

(2) 对于 $i = k-1, \dots, 1$ 依次计算:

$$\xi_{ik} = \left(\xi_{kk}r_{ik} - \sum_{p=i+1}^{k-1} r_{ip}\xi_{pk} \right) / (r_{ii} - r_{kk});$$

(3) 取: $\xi_{k+1,k} = \dots = \xi_{n,k} = 0$.

2.4. 证明: 对任意的 $A \in C^{n \times n}$ 有 QL 分解: $A = QL$, 其中 Q 是酉矩阵, L 是下三角矩阵. 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.

[证明] 对照定理 3.3.1 及定理 6.4.1 的证明.

2.5. 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 是排列矩阵, L 是下三角矩阵, U 是上三角矩阵. 证明:

$\tilde{H} = U(P^T L)$ 仍是上 Hessenberg 矩阵, 并且相似于 H .

[证明] 由于 H 是上 Hessenberg 矩阵, 故 H 的三角分解中的 L 具有如下形式:

$$L = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ & \ddots & \ddots & \\ & & l_{n,n-1} & 1 \end{bmatrix}$$

根据 P 的形成, 可知 $P^T L$ 呈三对角形, 即

$$P^T L = \begin{bmatrix} \times & \times & & \\ \times & \times & \times & \\ & \times & \times & \ddots \\ & & \ddots & \ddots & \times \\ & & & \times & \times \end{bmatrix}$$

从而对于 $j = 1, \dots, n-2$, 及 $i \geq j+2$, 我们有

$$\begin{aligned} \tilde{H}(i, j) &= e_i^T \tilde{H} e_j = e_i^T U (P^T L) e_j \\ &= \begin{bmatrix} \underbrace{0, \dots, 0}_{i-1}, u_{ij}, \dots, u_{in} \end{bmatrix} \begin{bmatrix} \underbrace{0}_{j-2} \\ \vdots \\ 0 \\ \times \\ \times \\ \times \\ \underbrace{0}_{n-j-1} \\ \vdots \\ 0 \end{bmatrix} = 0. \end{aligned}$$

这说明 \tilde{H} 是上 Hessenberg 矩阵.

另外, \tilde{H} 和 H 是相似的. 事实上

$$\tilde{H} = U(P^T L) = L^{-1} P A P^T L = (P^T L)^{-1} A (P^T L).$$

2.6. 设

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

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

$$Q^T A Q = \begin{bmatrix} \tilde{T} & u \\ 0 & \alpha_{11} \end{bmatrix},$$

其中 $\lambda(\tilde{T}) = \lambda(T)$.

[解] 为了给出构造正交矩阵的算法, 我们先看下面一个结论: 设 x 是 $A \in R^{n \times n}$ 对应于 λ 的一个左特征向量, 并假定 x 已规范化, 从而 $\|x\|_2 = 1$. 于是, 可以找到一个 $U \in R^{n \times n}$, 使 $[U, x]$ 是酉阵. 因为 $x^T A = \lambda x^T$ 所以

$$[U, x]^T A [U, x] = \begin{bmatrix} U^T \\ x^T \end{bmatrix} \begin{bmatrix} AU & Ax \end{bmatrix} = \begin{bmatrix} U^T AU & U^T Ax \\ \lambda x^T U & \lambda x^T x \end{bmatrix},$$

由于 $x^T x = 1, x^T U = [0, 0]$, 因此置

$$C = U^T AU, \quad h^T = U^T Ax,$$

那么

$$[U, x]^T A [U, x] = \begin{bmatrix} C & h^T \\ 0 & \lambda \end{bmatrix}.$$

据此结论, 我们设计的算法第一步当然就是计算矩阵 A 的一个左特征向量 x . 这可由下列方程组解出

$$\begin{bmatrix} 0 & 0 \\ v & T^T - \alpha_{11} I \end{bmatrix} x,$$

我们记 $x = (\xi_1, \xi_2, \xi_3)^T$, 并设定 $\xi_1 = 1$, 则 $\tilde{x} = (\xi_2, \xi_3)^T$ 可由下列二阶方程组解定

$$(\alpha_{11} I - T^T) \tilde{x} = v,$$

由于 α_{11} 不是 T 的特征值, 故该方程存在唯一的解. 计算出 A 关于 α_{11} 的左特征向量后, 再构造正交矩阵 $Q = [U, x]$, 便实现了本题需要的算法.

2.7. 设 $A \in R^{n \times n}$ 是一个如下形式的拟上三角矩阵

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ 0 & A_{22} & \cdots & A_{2k} \\ & & \ddots & \vdots \\ & & & A_{kk} \end{bmatrix},$$

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

[解] 若 A_{ii} 是 1×1 的矩阵, 即是 A 的一个实特征值, 此时可用反幂法求得相应的特征向量 x_i .

若 A_{ii} 是 2×2 的矩阵, 则存在一对共轭特征值 $\xi \pm i\eta$, 相应的特征向量为 x 和 \bar{x} . 现在我们较详细地讨论不用复运算计算相应特征向量的反幂法. 设相应的近似特征值为 $\tilde{\xi} \pm i\tilde{\eta}$, 复形式下的反幂法为

$$\begin{cases} (A - \tilde{\xi}I - i\tilde{\eta}I)y_{k+1} = u_k \\ u_{k+1} = y_{k+1} / \max(y_{k+1}) \end{cases} \quad (2.7-1)$$

令

$$y_{k+1} = x_{k+1} + iz_{k+1}, \quad u_k = p_k + iq_k,$$

代入 (2.7-1) 第一式, 令实部和虚部相等, 得到

$$(A - \tilde{\xi}I)x_{k+1} + \tilde{\eta}z_{k+1} = p_k \quad (2.7-2)$$

$$-\tilde{\eta}x_{k+1} + (H - \tilde{\xi}I)z_{k+1} = q_k \quad (2.7-3)$$

由上二式消去 z_{k+1} , 可以得到

$$[(A - \tilde{\xi}I)^2 + \tilde{\eta}^2 I]x_{k+1} = (A - \tilde{\xi}I)p_k - \tilde{\eta}q_k \quad (2.7-4)$$

联合 (2.7-4) 和 (2.7-2) 即可由 (p_k, q_k) 求得 (x_{k+1}, z_{k+1}) . 由 (2.7-4)

可知, 迭代时要求解的方程组系数矩阵是 $(A - \tilde{\xi}I)^2 + \tilde{\eta}^2 I$, 因为 A 是拟上三角矩阵, 这个矩阵只是比上三角阵多一条次对角线元素, 所以用列主元 Gauss 消去法进行三角分解时, 计算量还是比较小的.

2.8. 借助幂法设计一种计算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] 设 $A \in \mathbb{C}^{m \times n}$ 的奇异值就是矩阵 A^*A 的特征值的算术平方根. 所以可以通过下列算法求得一个给定矩阵的最奇异值, 其收敛性与幂法的收敛性的讨论类同. 算法步骤:

- 计算矩阵: $B = A^*A$;
- 利用幂法于矩阵 B , 求得矩阵 B 的最大特征值 λ ;
- 计算得到奇异值: $\sigma = \sqrt{\lambda}$.

2.9. 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] 由奇异值和奇异向量的定义可知: 若 $\sigma = \sqrt{\lambda}$ 是矩阵 $A \in C^{m \times n}$ 的奇异值, 则 A 关于奇异值 σ 的右奇异向量 u 和左奇异向量 v 分别满足

$$A^*Au = \lambda u, \quad AA^*v = \lambda v.$$

由此, 我们可以分别按以下两种算法计算得到最大奇异值的左奇异向量和右奇异向量:

算法 1 (计算右奇异向量的反幂法)

- 计算: $B = A^*A$;
- 利用反幂法计算 B 的按模最大的特征值的特征向量 (即 A 的右奇异向量)。

算法 2 (计算左奇异向量的反幂法)

- 计算: $B = AA^*$;
- 利用反幂法计算 B 的按模最大的特征值的特征向量 (即 A 的左奇异向量)。

30. 设 $A = X\Lambda X^{-1}$, 其中

$$X = [x_1, \dots, x_n], \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n).$$

并假定

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

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

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

这里 y_i^* 表示 X^{-1} 的第 i 行向量.

[证明] 对任意的初始向量

$$u_0 = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n,$$

注意到 $X = [x_1, \dots, x_n]$, $(X^{-1})^* = [y_1^*, \dots, y_n^*]$, 于是

$$y_i^* u_0 = \alpha_1 y_i^* x_1 + \alpha_2 y_i^* x_2 + \dots + \alpha_n y_i^* x_n,$$

再因 $A = X\Lambda X^{-1}$, 从而

$$y_i^k u_0 = \alpha_i, \quad i = 1, 2, \dots, n.$$

这样便有

$$u_0 = (y_1^* u_0) x_1 + (y_2^* u_0) x_2 + \dots + (y_n^* u_0) x_n$$

于是, 由幂法可以产生向量序列 $\{u_m\}$, 其中

$$u_m = \frac{A^m u_0}{\lambda_2^m}, \quad m = 1, 2, \dots$$

注意: 上式中分母原应为 λ_1^m , 只因 $|\lambda_1| = |\lambda_2|$, 故亦可用 λ_2^m 代替.
因为

$$A^m u_0 = (y_1^* u_0) \lambda_1^m x_1 + (y_2^* u_0) \lambda_2^m x_2 + \dots + (y_n^* u_0) \lambda_n^m x_n,$$

所以

$$u_m = (y_1^* u_0) e^{\frac{i 2ms}{t} \pi} x_1 + (y_2^* u_0) x_2 + (y_3^* u_0) \left(\frac{\lambda_3}{\lambda_2}\right)^m x_2 + \dots + (y_n^* u_0) \left(\frac{\lambda_n}{\lambda_2}\right)^m x_n,$$

由题设可知, 对于 $i \geq 3$ 有

$$\left(\frac{\lambda_i}{\lambda_2}\right)^m \rightarrow 0, \quad m \rightarrow \infty,$$

再由 $e^{\frac{i 2ms}{t} \pi}$ 的周期性可知, 对任意的非负整数 l , 及 $k = 1, 2, \dots, t$ 当 $m = lt + k$ 时

$$e^{\frac{i 2ms}{t} \pi} = e^{\frac{i 2ks}{t} \pi}.$$

从而 $\{u_m\}$ 存在 k 个收敛子序列, 它们是

$$\{u_{lt+k}\}_{l=0}^{\infty}, \quad k = 1, 2, \dots, t$$

并且

$$u_{lt+k} \rightarrow e^{\frac{i 2ks}{t} \pi} (y_1^* u_0) x_1 + (y_2^* u_0) x_2, \quad l \rightarrow \infty.$$

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

$$q_0 = \frac{u}{\|u\|_2}, \quad q_k = \frac{A q_{k-1}}{\|A q_{k-1}\|_2} \quad (k \geq 1),$$

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

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

[证明] 由于 A 是非亏损矩阵, 故与对角矩阵相似, 即存在

$$X = [x_1, \dots, x_n], \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n) \quad \text{使} \quad A = X \Lambda X^{-1}.$$

假设

$$u = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n, \quad \alpha_1 \neq 0.$$

于是

$$\begin{aligned} A^k u &= \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \dots + \alpha_n \lambda_n^k x_n \\ &= \lambda_1^k \left(\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right) \\ &= \lambda_1^k \left(\alpha_1 x_1 + O\left(\left|\lambda_2 / \lambda_1\right|^k\right) \right) \end{aligned} \quad (k \geq 1)$$

再由 q_k 的定义不难得知

$$\begin{aligned} q_k &= \frac{A q_{k-1}}{\|A q_{k-1}\|_2} = \dots = \frac{A^k q_0}{\|A^k q_0\|_2} = \frac{A^k u}{\|A^k u\|_2} \\ A q_k &= \frac{A^{k+1} u}{\|A^k u\|_2}, \end{aligned}$$

由于

$$\begin{aligned} \|A^k u\|_2^2 &= u^* (A^k)^* A^k u \\ &= (\bar{\alpha}_1 \bar{\lambda}_1^k x_1^* + \bar{\alpha}_2 \bar{\lambda}_2^k x_2^* + \dots + \bar{\alpha}_n \bar{\lambda}_n^k x_n^*) (\alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \dots + \alpha_n \lambda_n^k x_n) \\ &= |\alpha_1|^2 |\lambda_1|^{2k} \|x_1\|_2^2 + \bar{\alpha}_1 \alpha_2 \bar{\lambda}_1^k \lambda_2^k x_1^* x_2 + \dots + \bar{\alpha}_1 \alpha_n \bar{\lambda}_1^k \lambda_n^k x_1^* x_n \\ &\quad + \dots \\ &\quad + \bar{\alpha}_n \alpha_1 \bar{\lambda}_n^k \lambda_1^k x_n^* x_1 + \bar{\alpha}_n \alpha_2 \bar{\lambda}_n^k \lambda_2^k x_n^* x_2 + \dots + |\alpha_n|^2 |\lambda_n|^{2k} \|x_n\|_2^2 \\ &= |\lambda_1|^{2k} \left(|\alpha_1|^2 \|x_1\|_2^2 + O\left(\left|\lambda_2 / \lambda_1\right|^k\right) \right) \end{aligned}$$

同理可得

$$(A^k u)^* A^{k+1} u = q_k^* A q_k \|A^k u\|_2^2 = \lambda_1 |\lambda_1|^{2k} \left(|\alpha_1|^2 \|x_1\|_2^2 + O\left(\left|\lambda_2 / \lambda_1\right|^k\right) \right)$$

从而

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

若假定 A 是 Hermite 矩阵，则前述中的非奇异矩阵 X 将是酉矩阵，从而

$$\begin{aligned}\|A^k u\|_2^2 &= |\alpha_1|^2 |\lambda_1|^{2k} \|x_1\|_2^2 + \dots + |\alpha_n|^2 |\lambda_n|^{2k} \|x_n\|_2^2 \\ &= |\lambda_1|^{2k} \left(|\alpha_1|^2 \|x_1\|_2^2 + O\left(|\lambda_2 / \lambda_1|^{2k}\right) \right),\end{aligned}$$

$$(A^k u) A^{k+1} u = \lambda_1 |\lambda_1|^{2k} \left(|\alpha_1|^2 \|x_1\|_2^2 + O\left(|\lambda_2 / \lambda_1|^{2k}\right) \right)$$

于是

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