Chap 1 线性方程组的直接法

§1.1 三角形方程组的解法

(一) 下三角阵 (前代法)

$$\begin{cases} l_{11}y_1 = b_1 \\ l_{21}y_1 + l_{22}y_2 = b_2 \\ \vdots \qquad \vdots \\ l_{n1}y_1 + l_{n2}y_2 + \cdots + l_{nn}y_n = b_n \end{cases}$$

$Ly = b$ $\quad L = \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix}$

$y_1 = b_1 / l_{11}$

$y_2 = (b_2 - l_{21} \cdot y_1) / l_{22}$

$\vdots \qquad \vdots$

$y_i = (b_i - \sum_{j=1}^{i-1} l_{ij} y_j) / l_{ii}$

算法1.1.1 for j=1 : n-1

b(j) = b(j)/L(j,j) # 得到的 y_i 存在 b_i

b(j+1 : n) = b(j+1 : n) - b(j)L(j+1:n, j)

end

b(n) = b(n)/L(n,n)

乘除法运算量：$\sum_{j=1}^{n-1}(1+n-j)+1 = \frac{n(n+1)}{2}$

总运算量(加减乘除)：$\sum_{j=1}^{n-1}[1+2(n-j)]+1 = n^2$

(二) 下三角阵

$$\begin{cases} u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n = b_1 \\ \qquad u_{22}x_2 + \cdots + u_{2n}x_n = b_2 \\ \qquad\qquad \ddots \qquad \vdots \\ \qquad\qquad\qquad u_{nn}x_n = b_n \end{cases}$$

$x_n = b_n/u_{nn}$

$\vdots$

$x_i = (b_i - \sum_{j=i+1}^{n} u_{ij} x_j)/u_{ii}$

算法1.1.2 for $j=n:-1:2$

$y(j) = y(j)/U(j,j)$

$y(1:j-1) = y(1:j-1) - y(j)U(1:j-1,j)$

end

$y(1)=y(1)U(1,1)$

运算量 n^2

§1.2 Gauss消元法(一般线性方程组化为三角形方程组)

对于一般线性方程组的矩阵形式 $AX=b$，记 (A,b) 为增广矩阵

对 (A,b) 作一系列行的初等变换，得 $(\tilde{A},\tilde{b})$，则 $\tilde{A}x=\tilde{b}$ 为 $AX=b$ 的同解方程组.

把 (A,b) 作一系列行的初等变换，使之化为三角形方程组，且与原方程同解的求解方法称为消元法.

(一) 消元过程 $(AX=b)$

$A^{(1)}x = b^{(1)}$. 把方程组写为 $\begin{cases} a_{11}^{(1)} x_1^{(1)} + a_{12}^{(1)} x_2^{(1)} + \cdots + a_{1n}^{(1)} x_n^{(1)} = b_1^{(1)} \\ a_{21}^{(1)} x_1^{(1)} + a_{22}^{(1)} x_2^{(1)} + \cdots + a_{2n}^{(1)} x_n^{(1)} = b_2^{(1)} \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ a_{n1}^{(1)} x_1^{(1)} + a_{n2}^{(1)} x_2^{(1)} + \cdots + a_{nn}^{(1)} x_n^{(1)} = b_n^{(1)} \end{cases}$

增广矩阵 $(A^{(1)}, b^{(1)}) = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{pmatrix}$

若 $a_{11} \neq 0$，分别以 $-a_{i1}^{(1)}/a_{11}^{(1)}$ 乘第1行，加到第i行，得到

$$(A^{(2)}, b^{(2)}) = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{pmatrix}$$

$A^{(2)}x=b^{(2)}$ 和 $A^{(1)}x=b^{(1)}$ 同解. 且 $a_{ij}^{(2)}=-a_{i1}^{(1)}/a_{11}^{(1)}\cdot a_{1j}^{(1)}+a_{ij}^{(1)}$

$$b_i^{(2)}=-a_{i1}^{(1)}/a_{11}^{(1)}\cdot b_1^{(1)}+b_i^{(1)}$$

$(A^{(1)},b^{(1)})\longrightarrow(A^{(2)},b^{(2)})$ 经过n-1次行的初等变换

$(A^{(2)},b^{(2)})$ 由 $(A^{(1)},b^{(1)})$ 左乘n-1个初等矩阵得到的

记 $l_{i1}=\dfrac{a_{i1}^{(1)}}{a_{11}^{(1)}}$ $\begin{pmatrix}1&&&\\-l_{21}&\ddots&&\\&&\ddots&\\&&&1\end{pmatrix}\begin{pmatrix}1&&&\\0&\ddots&&\\-l_{31}&&\ddots&\\&&&1\end{pmatrix}\cdots\begin{pmatrix}1&&&\\0&\ddots&&\\\vdots&&\ddots&\\-l_{n1}&&&1\end{pmatrix}$

乘积为 $\begin{pmatrix}1&&&\\-l_{21}&\ddots&&\\-l_{31}&&\ddots&\\\vdots&&&\\-l_{n1}&&&1\end{pmatrix}\triangleq \underline{L_1}$ $A^{(2)}=L_1A^{(1)}$ $b^{(2)}=L_1b^{(1)}$

↑这种类型的初等下三角阵为Gauss变换

$L_1=I-l_1e_1^T$, 其中 I 为 $n\times n$ 单位阵, $l_1=(0,l_{21},l_{31}\cdots l_{n1})^T$ → Gauss向量

$$e_1=(1,0\cdots\cdots 0)^T$$

$L_1^{-1}=\begin{pmatrix}1&&&\\l_{21}&\ddots&&\\\vdots&&\ddots&\\l_{n1}&&&1\end{pmatrix}=I+l_1e_1^T$

类似的消元过程进行至k-1步

得到 $(A^{(k)},b^{(k)})=\begin{pmatrix}a_{11}^{(1)}&a_{12}^{(1)}&a_{13}^{(1)}&\cdots\cdots&a_{1n}^{(1)}&b_1^{(1)}\\0&a_{22}^{(2)}&a_{23}^{(2)}&\cdots&a_{2n}^{(2)}&b_2^{(2)}\\\vdots&0&a_{33}^{(3)}&\cdots\cdots&a_{3n}^{(3)}&b_3^{(3)}\\\vdots&\vdots&\ddots\ a_{kk}^{(k)}&\cdots&a_{kn}^{(k)}&b_k^{(k)}\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0\cdots a_{nk}^{(k)}&\cdots&a_{nn}^{(k)}&b_n^{(k)}\end{pmatrix}$

若 $a_{kk}^{(k)}\neq0$, 分别以 $-a_{ik}^{(k)}/a_{kk}^{(k)}$ 乘第k行, 加到第i行, 得到 $(A^{(k+1)},b^{(k+1)})$

$A^{(k+1)}x=b^{(k+1)}$ 与 $A^{(k)}x=b^{(k)}$ 同解

$(A^{(k+1)},b^{(k+1)})$ 与 $(A^{(k)},b^{(k)})$ 前k行元素相同, 且

$$\begin{cases}a_{ij}^{(k+1)}=a_{ij}^{(k)}-a_{ik}^{(k)}/a_{kk}^{(k)}\cdot a_{kj}^{(k)}\\b_i^{(k+1)}=b_i^{(k)}-a_{ik}^{(k)}/a_{kk}^{(k)}\cdot b_k^{(k)}\\a_{ik}^{(k+1)}=0\quad i=k+1\cdots n\end{cases}$$

类似于第一步: $L_k=\begin{pmatrix}1&&&\\&\ddots&&\\&&1&&\\&&-l_{k+1,k}&\ddots&\\&&\vdots&&\\&&-l_{n,k}&&1\end{pmatrix}$ $A^{(k+1)}=L_kA^{(k)}$, $b^{(k+1)}=L_kb^{(k)}$

$L_k=I-l_ke_k^T$ $l_k=(0,\cdots0,l_{k+1,k}\cdots l_{n,k})^T$

$e_k^T=(0,\cdots0,1,0\cdots0)^T$ (I的第k列)

以上消元过程进行n-1步后化为上三角阵 $(A^{(n)}, b^{(n)})$

$A^{(n)}x=b^{(n)}$ 与 $Ax=b$ 同解

且 $U=A^{(n)}=\underbrace{L_{n-1}L_{n-2}\cdots L_1}_{L^{-1}} \underset{A}{A^{(1)}}$

则 $A=LU$ $Ax=b$ 求解 $\to LUx=b \to Ly=b,\ Ux=y$ LU分解

$L=L_1^{-1}L_2^{-1}\cdots L_{n-1}^{-1}=(I+l_1e_1^T)(I+l_2e_2^T)\cdots(I+l_{n-1}e_{n-1}^T)$

$=I+l_1e_1^T+l_2e_2^T+\cdots+l_{n-1}e_{n-1}^T$

$=\begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & l_{32} & \ddots & \\ l_{n1} & l_{n2} & \cdots\ l_{n,n-1} & 1 \end{pmatrix}$ 单位下三角阵

算法1.1.3 for $k=1:n-1$

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

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

end

#输出 $A=\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ l_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & u_{nn} \end{pmatrix}$

算法运算量: $\sum_{k=1}^{n-1}\left((n-k)+2(n-k)^2\right)=\frac{2}{3}n^3+O(n^2)$

求解 $Ax=b$ 的步骤:

① Gauss消元: $A=LU$,算法1.1.3

② 用算法1.1.1求 $Ly=b$,得到 y

③ 用算法1.1.2求 $Ux=y$,得到 x

(二)Gauss消元法的可行性条件

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix}$$

第一步：$a_{11}^{(1)} \neq 0$. 第二步：$a_{22}^{(2)} \neq 0$ $\Longleftrightarrow$ A的一阶,二阶顺序主子式不为0.

Thm 1.1.1 主元 $a_{ii}^{(i)}(i=1,2,3\cdots n)$ 均不为0的充分必要条件是

A的i阶顺序主子阵 A_i 都是非奇异的.

Pf：数学归纳法

Thm 1.1.2 若A的各阶顺序主子式不为0,则A可分解为一个单位下三角阵和一个上三角阵的乘积. 即：$A=LU$

Pf：Gauss消元过程

Thm 1.1.3 线性方程组 $Ax=b$ 可用Gauss消元法求解的充分必要条件是其系数矩阵A的各阶顺序主子式不为0.

Pf："$\Rightarrow$"：用定理1.1.1

"$\Leftarrow$"：用定理1.1.2

思考：① $\det A \neq 0 \overset{?}{\Rightarrow}$ A的顺序主子式不为0

② Gauss消元法是否改变行列式的值？

§1.3 选主元消元法

(一)全主元消元法

(A,b) 经过k-1步消元得到 $(A^{(k)}, b^{(k)})$

$$(A^{(k)}, b^{(k)}) = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & \cdots & & a_{2n}^{(2)} & b_2^{(2)} \\ & & \ddots & & \vdots & \vdots \\ & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\ & & \vdots & & \vdots & \vdots \\ & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & a_{nn}^{(k)} \end{pmatrix}$$

选主元 $|a_{pq}^{(k)}| = \max\limits_{k\leq i,j\leq n} |a_{ij}^{(k)}|$, 交换 $(A^{(k)}, b^{(k)})$ 的第k行与第p行；第k列与第q列.

记交换以后的矩阵为 $(\tilde{A}^{(k)}, \tilde{b}^{(k)}) \xrightarrow{\text{Gauss 消元法}} (A^{(k+1)}, b^{(k+1)})$

初等置换矩阵（单位阵的列与列交换所得）

k.p行交换　　k.q列交换

$P_k = I_{kp}$.　$Q_k = I_{kq}$　　$(P_k^2 = Q_k^2 = I)$

则 $\tilde{A}^{(k)} = P_k A^{(k)} Q_k$　　$A^{(k+1)} = L_k \tilde{A}^{(k)} = L_k P_k A^{(k)} Q_k$

$U = A^{(n)} = L_{n-1} P_{n-1} A^{(n-1)} Q_{n-1}$

$$= \underbrace{L_{n-1} P_{n-1} L_{n-1} P_{n-2} L_{n-3} P_{n-3} \cdots L_1 P_1} A \underbrace{Q_1 Q_2 \cdots Q_{n-1}}_{Q}$$

引入 $Q = Q_1 Q_2 \cdots Q_{n-1}$

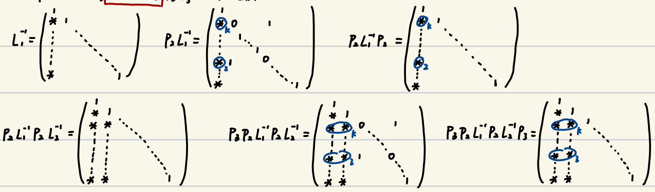
$P = P_{n-1} P_{n-2} \cdots P_1$　　$L = P(L_{n-1} P_{n-1} L_{n-2} P_{n-2} \cdots L_1 P_1)^{-1}$

则 $PAQ = LU$（全主元消元法）

问题：L为下三角阵？

$L = P_{n-1} P_{n-2} \cdots P_1 P_1 L_1^{-1} P_2 L_2^{-1} P_3 L_3^{-1} \cdots P_{n-1} L_{n-1}^{-1}$

$= P_{n-1} P_{n-2} \cdots P_3 \boxed{P_2 L_1^{-1} P_2 L_2^{-1}} P_3 L_3^{-1} \cdots P_{n-1} L_{n-1}^{-1}$



…… L应为下三角阵.（可归纳证明）

Thm: 1.2.1 $A \in \mathbb{R}^{n\times n}$，则存在排列矩阵 $P, Q \in \mathbb{R}^{n\times n}$ 以及单位下三角阵 L 和上三角阵 U，使得 $PAQ = LU$

且 L 的所有元素满足 $|L_{ij}| \le 1$，U 的非零对角元个数 $= \operatorname{rank}(A)$　　$|l_{ik}| = \left|\frac{\tilde{a}_{ik}^{(k)}}{\tilde{a}_{kk}^{(k)}}\right| \le 1$

算法1.2.1

选主元：要进行 $\sum_{k=1}^{n-1}(n-k+1)^2 \sim \frac{1}{3}n^3$ 次比较.

用全主元Gauss消元法求解AX=b的步骤

① 用算法1.2.1计算A的全主元三角分解. $PAQ=LU$.

② 用算法1.1.1计算 $Ly=Pb$ 得 y　　$\underbrace{(PAQ)}_{LU}\cdot Q^T x=Pb$

③ 用算法1.1.2计算 $Uz=y$ 得 z

④ $x=Qz$.

(二) 列主元消元法 (为了减少比较次数)

(A,b) 经过 $k-1$ 步消元后得到 $(A^{(k)},b^{(k)})=$

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & \cdots & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ & & \ddots & & \vdots & \vdots \\ & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\ & & \vdots & & \vdots & \vdots \\ & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & a_{nn}^{(k)} \end{pmatrix}$$

选主元: $|a_{pk}^{(k)}|=\max\limits_{k\le i\le n}|a_{ik}^{(k)}|$

交换 $(A^{(k)},b^{(k)})$ 的第 p 行和第 k 行, 记交换后的矩阵为 $(\tilde{A}^{(k)},\tilde{b}^{(k)})\xrightarrow{\text{Gauss消元}}(A^{(k+1)},b^{(k+1)})$

记初等置换矩阵 $P_k=I_{kp}$, $\tilde{A}^{(k)}=P_kA^{(k)}$

$$A^{(k+1)}=L_k\tilde{A}^{(k)}=L_kP_kA_k$$

上三角阵 $U=A^{(n)}=L_{n-1}P_{n-1}A^{(n-1)}=L_{n-1}P_{n-1}L_{n-2}P_{n-2}\cdots\cdots L_1P_1A$

记排列矩阵 $P=P_{n-1}P_{n-2}\cdots\cdots P_1$, $L=P(L_{n-1}P_{n-1}L_{n-2}P_{n-2}\cdots L_1P_1)^{-1}$ 为单位下三角阵、 $PA=LU$

用列主元消元法求解 $Ax=b$ 的步骤

① 用算法1.2.2计算A的列主元分解 $PA=LU$

② 用算法1.1.1计算 $Ly=Pb$ 得 y

③ 用算法1.1.2计算 $Ux=y$ 得 x

选主元要进行 $\sum\limits_{k=1}^{n-1}(n-k+1)\sim O(n^2)$

总结: 列主元Gauss消元法与全主元Gauss消元法在数值稳定性方面完全可以媲美, 但计算量却大为减少.

应用: 中小型稠密线性方程组.

§1.4 直接分解法

(一) 正定矩阵的平方根法

$A\in\mathbb{R}^{n\times n}$ 是对称正定的，即A满足 $A=A^T$ 且 $x^TAx>0$，$\forall x\in\mathbb{R}^n$，$x\neq 0$

(I) Thm 1.3.1 (Cholesky分解定理) 若 $A\in\mathbb{R}^{n\times n}$ 对称正定，则存在一个对角元均为正数的下三角阵 $L\in\mathbb{R}^{n\times n}$. s.t.

$$A=\underbrace{L}_{\text{Cholesky因子}}L^T.$$

Pf: A正定．则A的全部顺序主子阵均正定.

由Gauss消元法的LU分解定理．∃一个单位下三角阵 $\tilde{L}$ 和一个上三角阵 U. s.t. $A=\tilde{L}U$.

$\because$ A正定．$\Rightarrow \det(A)\neq 0 \Rightarrow \det(\tilde{L})\neq 0,\ \det(U)\neq 0$ $\Rightarrow u_{ii}\neq 0,\ i=1,2,\cdots n$

令 $D=\mathrm{diag}(u_{11},u_{22},\cdots,u_{nn}) \Rightarrow D$ 可逆

记 $\tilde{U}=D^{-1}U$．则 $\tilde{U}$ 为单位上三角阵.

且有 $A=\tilde{L}D\tilde{U}$，$A^T=\tilde{U}^TD\tilde{L}^T$

$\because A=A^T \Rightarrow \tilde{L}D\tilde{U}=\tilde{U}^TD\tilde{L}^T \Rightarrow \underbrace{D^{-1}\tilde{U}^{-T}\tilde{L}D}_{\text{单位下三角}}=\underbrace{\tilde{L}^T\tilde{U}^{-1}}_{\text{单位下三角}}$

$\Rightarrow D^{-1}\tilde{U}^{-1}\tilde{L}D=\tilde{L}^T\tilde{U}^{-1}=I$

$\therefore \tilde{U}=\tilde{L}^T \Rightarrow A=\tilde{L}D\tilde{L}^T$. ($\forall x\neq 0$, $x^TAx>0$ 即 $x^T\tilde{L}D\tilde{L}^Tx>0$, $(\tilde{L}^Tx)^TD(\tilde{L}^Tx)>0$. $\therefore$ D对角元为正)

令 $L=\tilde{L}\,\mathrm{diag}(\sqrt{u_{11}},\sqrt{u_{22}}\cdots\sqrt{u_{nn}}) \Rightarrow A=LL^T$. L的对角元为正

(II) 如何求L?

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & & a_{nn}\end{pmatrix}=A=LL^T=\begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn}\end{pmatrix}\begin{pmatrix} l_{11} & l_{21} & \cdots & l_{n1}\\ & l_{22} & \cdots & l_{n2}\\ & & \ddots & \vdots\\ & & & l_{nn}\end{pmatrix}$$

比较 $A=LL^T$ 两边元素有

第一列：$a_{11}=l_{11}^2$

$\Rightarrow l_{11}=\sqrt{a_{11}}$　$l_{i1}=a_{i1}/a_{11}$

$a_{i1}=l_{i1}l_{11}$

若已计算出L的前k-1列元素，比较第k列有：

$$a_{kk}=\sum_{p=1}^{k}l_{kp}^2=\sum_{p=1}^{k-1}l_{kp}^2+l_{kk}^2$$

$$a_{ik}=\sum_{p=1}^{k}l_{ip}l_{kp}=\sum_{p=1}^{k-1}l_{ip}l_{kp}+l_{ik}l_{kk}\quad i=k+1,\cdots n$$

$$\Rightarrow\begin{cases}l_{kk}=\sqrt{a_{kk}-\sum_{p=1}^{k-1}l_{kp}^2}\\ l_{ik}=(a_{ik}-\sum_{p=1}^{k-1}l_{ip}l_{kp})/l_{kk}\quad i=k+1,\cdots\cdots,n\end{cases}$$

$$l_{32}=(a_{32}-\ l_{31}\cdot l_{21})/l_{22}$$

算法1.3.1 运算量为Gauss消元法一半. $O(n^3)$

若$Ax=b$的系数矩阵$A>0$, 则求解$Ax=b$的步骤：

①用算法1.3.1计算A的平方根分解 $A=LL^T$.

②用算法1.1.1计算 $Ly=b$ 得y.

③用算法1.1.2计算 $L^Tx=y$ 得x.

(二)改进的平方根法(避免开方)

同(1)的定理证明：存在单位下三角阵L和一个对角阵D, 使得 $A=LDL^T$.

$$\begin{pmatrix}a_{11}&a_{12}&\cdots\cdots&a_{1n}\\a_{21}&a_{22}&\cdots\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&&a_{nn}\end{pmatrix}=A=LDL^T=\begin{pmatrix}1&&&\\l_{21}&1&&\\\vdots&\vdots&\ddots&\\l_{n1}&l_{n2}&\cdots\cdots&1\end{pmatrix}\begin{pmatrix}d_1&&&\\&d_2&&\\&&\ddots&\\&&&d_n\end{pmatrix}\begin{pmatrix}1&l_{21}&\cdots\cdots&l_{n1}\\&1&\cdots\cdots&l_{n2}\\&&\ddots&\vdots\\&&&1\end{pmatrix}$$

$$=\begin{pmatrix}1&&&\\l_{21}&1&&\\\vdots&\vdots&\ddots&\\l_{n1}&l_{n2}&\cdots\cdots&1\end{pmatrix}\begin{pmatrix}d_1&d_1l_{21}&\cdots\cdots&d_1l_{n1}\\&d_2&\cdots\cdots&d_2l_{n2}\\&&\ddots&\vdots\\&&&d_n\end{pmatrix}$$

对比两边元素：

第一列：$a_{11}=d_1\quad a_{i1}=l_{i1}d_1\ \Rightarrow d_1=a_{11}\quad l_{i1}=a_{i1}/d_1$

假设已知$d_1\cdots d_{k-1}$, L的前k-1列元素

对比第k列：$a_{ik}=\sum_{p=1}^{k}l_{ip}d_pl_{kp}=\sum_{p=1}^{k-1}l_{ip}d_pl_{kp}+l_{ik}d_kl_{kk}$

$$\begin{cases}d_k=a_{kk}-\sum_{p=1}^{k-1}d_pl_{kp}^2\\ l_{ik}=(a_{ik}-\sum_{p=1}^{k-1}l_{ip}d_pl_{kp})/d_k.\end{cases}$$

算法1.3.2 运算量 $\frac{1}{3}n^3$.

若 $Ax=b$ 的系数矩阵 $A>0$, 则求解 $Ax=b$ 的步骤:

①用算法1.3.2计算 $A=LDL^T$

②用算法1.1.1计算 $Ly=b$ 得 y.

③用算法1.1.2计算 $Dz=y$ 得 z.

④用算法1.1.2计算 $L^Tx=z$ 得 x.

(三) 带状矩阵分解

(I) 非零元素分布在对角线元素的两侧.

若上侧有 m_1 条次对角线, 下侧有 m_2 条次对角线有非零元素.

称为带宽为 m_1+m_2+1 的带状矩阵.

若其各阶主子式不为0. 则可分解为带宽 m_1+1 和 m_2+1 的上三角阵 U 与下三角阵 L 的乘积.

(II) 三对角线矩阵

$$A=\begin{pmatrix} a_1 & b_1 & & & \\ c_2 & a_2 & b_2 & & \\ & c_3 & a_3 & \ddots & \\ & & \ddots & \ddots & b_{n-1} \\ & & & c_n & a_n \end{pmatrix}=\begin{pmatrix} \alpha_1 & & & \\ \gamma_2 & \alpha_2 & & \\ & \gamma_3 & \alpha_3 & \\ & & \ddots & \ddots \\ & & \gamma_n & \alpha_n \end{pmatrix}\begin{pmatrix} 1 & \beta_1 & & & \\ & 1 & \beta_2 & & \\ & & 1 & \ddots & \\ & & & \ddots & \beta_{n-1} \\ & & & & 1 \end{pmatrix}$$

$$\left.\begin{array}{ll} \Rightarrow c_i=\gamma_i \quad i=2,3\cdots n & \text{定义 } c_1=\gamma_1=0 \\ a_i=\alpha_i+\gamma_i\beta_{i-1} & b_i=\alpha_i\beta_i \end{array}\right\} \Rightarrow \begin{array}{l} \gamma_i=c_i \\ \boxed{\begin{array}{l} \alpha_i=a_i-c_i\beta_{i-1} \quad \text{定义 } c_1=0 \\ \beta_i=b_i/\alpha_i. \quad \text{定义 } b_n=0 \end{array}} \end{array}$$

(III) 求 $Ax=f \rightarrow Ly=f \quad Ux=y$

$$\Rightarrow \boxed{\begin{array}{ll} y_i=(f_i-c_iy_{i-1})/\alpha_i & i=1,2,\cdots n \\ x_i=y_i-\beta_i x_{i+1} & i=n,n-1,\cdots,1 \end{array}} \quad \begin{array}{l} c_1=0 \\ \beta_n=0 \end{array}$$

$$\begin{cases} \left.\begin{array}{ll} \alpha_i=a_i-c_i\beta_{i-1} & c_1=0 \\ \beta_i=b_i/\alpha_i & \end{array}\right\} i=1\cdots n \\ y_i=(f_i-c_iy_{i-1})/\alpha_i \\ x_k=y_k-\beta_k x_{k+1} \quad k=n\cdots 1 \end{cases}$$

追赶法

通过对三对角矩阵的直接分解法. 求解具有三对角线系数矩阵的线性方程组的方法.

(四) 一般矩阵. $A=LU$

(I) Doolittle分解 $A=\underset{\text{单位下}}{L}\,\underset{\text{上三角}}{U}$

(II) Courant 分解 $A=\underset{\text{下三角}}{L}\,\underset{\text{单位上}}{U}$

} 对比行列

扩展：(1) 还有其他方法：参数法、线性插值法等.

(2) 求逆A. $AB=I$ $B=(b_1\cdots b_n)$ $Ab_i=e_i$.

Gauss-Jordan消元：$(A,I)\to(I,B)$

$$\begin{pmatrix} a_{11} & * & 1 & 0 \\ * & A_{22} & 0 & I \end{pmatrix} \xrightarrow{\text{Gauss消元}} \begin{pmatrix} a_{11} & * & 1 & 0 \\ 0 & \tilde{A}_{22} & * & I \end{pmatrix} \xrightarrow{\text{第一行}\times\frac{1}{a_{11}}} \begin{pmatrix} 1 & * & \frac{1}{a_{11}} & 0 \\ 0 & \tilde{A}_{22} & * & I \end{pmatrix}$$

$$\xrightarrow[\text{加到}1,3\cdots n\text{行}]{\text{第2行}\times\text{倍}} \begin{pmatrix} 1 & 0 & * & \frac{1}{a_{11}} & * & * \\ 0 & \tilde{a}_{22} & * & * \\ 0 & 0 & \tilde{A}_{33} & * \end{pmatrix}$$

- 算法
  - 思想
  - 格式
  - 分析
    - 敏度分析：稳定性
    - 误差分析
      - 精度
      - 机器误差

Chap 2 敏度分析和误差分析

§2.1 向量范数和矩阵范数.

(一) 向量范数.

(1) 定义：函数 $\|\cdot\|:\mathbb{R}^n\to\mathbb{R}$ 满足

1° 正定性：$\forall x\in\mathbb{R}^n$，有 $\|x\|\geq 0$，而且 $\|x\|=0$ 当且仅当 $x=0$

2° 齐次性：$\forall x\in\mathbb{R}^n$，$\alpha\in\mathbb{R}$. 有 $\|\alpha x\|=|\alpha|\cdot\|x\|$

3° 三角不等式：$\forall x,y\in\mathbb{R}^n$，有 $\|x+y\|\leq\|x\|+\|y\|$

称 $\|\cdot\|$ 为 $\mathbb{R}^n$ 上的向量范数.

最常用的向量范数为 p 范数 (Hölder 范数)

$x=(x_1\cdots x_n)^T$. $\|x\|_p=(|x_1|^p+|x_2|^p+\cdots+|x_n|^p)^{\frac{1}{p}}$. $p\geq 1$ 最重要: $p=1,2,\infty$. $\left(\|x\|_\infty=\max\limits_{1\leq i\leq n}|x_i|\right)$

(II) 性质

1° 连续性: 定理: $\|\cdot\|$ 作为 $\mathbb{R}^n$ 上的向量范数, 是其 n 个分量的连续函数.

Pf: $\left|\|x\|-\|y\|\right|\leq\|x-y\|=\|\sum\limits_{i=1}^n(x_i-y_i)e_i\|$ 单位阵的第 i 列

$\leq\sum\limits_{i=1}^n\|(x_i-y_i)e_i\|$

$\leq\sum\limits_{i=1}^n|x_i-y_i|\cdot\|e_i\|$

$\leq\max\limits_{1\leq i\leq n}|x_i-y_i|\left[\sum\limits_{i=1}^n\|e_i\|\right]^{N}$

$\forall\varepsilon>0$, $\exists\delta=\frac{\varepsilon}{N}>0$, 当 $|x_i-y_i|<\delta$ 时, $\forall i$ 有 $\left|\|x\|-\|y\|\right|\leq\frac{\varepsilon}{N}\cdot N=\varepsilon$.

2° 等价性: $\mathbb{R}^n$ 中任意范数都是等价的. 即若 $\|\cdot\|_\alpha$ 和 $\|\cdot\|_\beta$ 为 $\mathbb{R}^n$ 中任意两个范数

则存在 $m, M>0$, s.t. $\forall x\in\mathbb{R}^n$, 有 $m\|x\|_\alpha\leq\|x\|_\beta\leq M\|x\|_\alpha$

想证: $\|\cdot\|_\alpha$ 与 $\|\cdot\|_2$ 等价.

Pf: 记 $S=\{x\in\mathbb{R}^n \mid \|x\|_2=1\}$ 为有界闭集.

设 $\|\cdot\|_\alpha$ 为 $\mathbb{R}^n$ 上任意向量范数, 由连续性可知 $\|\cdot\|_\alpha$ 为 S 上的连续函数.

则 $\|\cdot\|_\alpha$ 在 S 上可取最大最小值, 记 $m=\min\limits_{x\in S}\|x\|_\alpha$, $M=\max\limits_{x\in S}\|x\|_\alpha$.

$\forall x\in\mathbb{R}^n$, $\frac{x}{\|x\|_2}\in S$ $\Rightarrow m\leq\left\|\frac{x}{\|x\|_2}\right\|_\alpha\leq M$. 由齐次性 $m\|x\|_2\leq\|x\|_\alpha\leq M\|x\|_2$. □

Ex: $\|x\|_2\leq\|x\|_1\leq\sqrt{n}\|x\|_2$

$\|x\|_\infty\leq\|x\|_2\leq\sqrt{n}\|x\|_\infty$

$\|x\|_\infty\leq\|x\|_1\leq n\|x\|_\infty$.

3° 收敛性:

记 $x^{(m)}=(x_1^{(m)}, x_2^{(m)}, x_3^{(m)}\cdots\cdots x_n^{(m)})^T$, $\{x^{(m)}\}_{m=1}^\infty$ 为 $\mathbb{R}^n$ 上的向量序列

若$\exists x=(x_1,\cdots x_n)^T$. s.t. $\lim\limits_{m\to\infty}\|x^{(m)}-x\|=0$

称$x^{(m)}$收敛于x. 记作$\lim\limits_{m\to\infty}x^{(m)}=x$ 或 $x^{(m)}\to x$. $(m\to\infty)$.

Thm: $\lim\limits_{m\to\infty}\|x^{(m)}-x\|=0 \Leftrightarrow \lim\limits_{m\to\infty}|x_i^{(m)}-x_i|=0$. $\forall i=1,2,\cdots n$

范数等价. 取1范数证明即可.

(二)矩阵范数

$m\times n$的矩阵全体$\mathbb{R}^{m\times n}$与$m\times n$维的向量空间$\mathbb{R}^{mn}$同构.

则$\mathbb{R}^{m\times n}$上的范数可由向量范数的概念直接推广.

$A\in\mathbb{R}^{m\times n}$: $A:\mathbb{R}^n\to\mathbb{R}^m$ $\quad \|Ax\|\le\|A\|\cdot\|x\|$

$\|AB\|\le\|A\|\cdot\|B\|$

相容性

为了保持矩阵作为线性算子的特征以及矩阵的乘法性质, 须引入相容性.

不是所有的向量范数都具有相容性.

(I) 定义$\mathbb{R}^{m\times n}\to\mathbb{R}$满足

1° 正定性: $\forall A\in\mathbb{R}^{m\times n}$, 有$\|A\|\ge 0$. 而且$\|A\|=0$当且仅当$A=0$

2° 齐次性: $\forall A\in\mathbb{R}^{m\times n}$, $\alpha\in\mathbb{R}$. 有$\|\alpha A\|=|\alpha|\cdot\|A\|$

3° 三角不等式: $\forall A,B\in\mathbb{R}^{m\times n}$, 有$\|A+B\|\le\|A\|+\|B\|$

4° $A\in\mathbb{R}^{m\times n}$, $B\in\mathbb{R}^{n\times p}$, $x\in\mathbb{R}^n$ 有$\|AB\|\le\|A\|\cdot\|B\|$, $\|Ax\|\le\|A\|\cdot\|x\|$

矩阵范数又称相容范数.

(II) 性质(具有向量范数的所有性质)

1° 连续性: 定理: $\|\cdot\|$作为$\mathbb{R}^{m\times n}$上的矩阵范数, 是其$m\times n$个分量的连续函数.

2° 等价性: $\mathbb{R}^{m\times n}$中任意范数都是等价的

(III) 诱导范数

$A\in\mathbb{R}^{n\times n}$. 定义 $|||A||| = \max\limits_{\substack{x\in\mathbb{R}^n\\ x\neq 0}} \frac{\|Ax\|}{\|x\|} = \max\limits_{\|x\|=1}\|Ax\|$

(a): $|||\cdot|||$ 为 $\mathbb{R}^{n\times n}$ 上矩阵范数

Pf: 1° $|||A|||\geq 0$ $|||A|||=0 \Leftrightarrow \forall x\in\mathbb{R}^n, \|x\|=1, \|Ax\|=0 \Leftrightarrow \forall x\in\mathbb{R}^n, \|x\|=1, Ax=0$

$\Leftrightarrow A=0.$

2° $|||\alpha A||| = \max\limits_{\|x\|=1}\|\alpha Ax\| = \alpha\max\limits_{\|x\|=1}\|Ax\| = \alpha|||A|||$

3° $|||A+B||| = \max\limits_{\|x\|=1}\|(A+B)x\| = \max\limits_{\|x\|=1}\|Ax+Bx\| \leq \max\limits_{\|x\|=1}(\|Ax\|+\|Bx\|)$

$\leq \max\limits_{\|x\|=1}\|Ax\| + \max\limits_{\|x\|=1}\|Bx\| = |||A|||+|||B|||$

4° $\forall x\in\mathbb{R}^n$ 有 $\|Ax\|\leq|||A|||\cdot\|x\|$. 即 $\|Ax\| = \frac{\|Ax\|}{\|x\|}\cdot\|x\| \leq |||A|||\cdot\|x\|.$

$|||AB||| = \max\limits_{\|x\|=1}\|ABx\| = \max\limits_{\|x\|=1}\frac{\|ABx\|}{\|Bx\|}\cdot\|Bx\|$

$\leq |||A|||\cdot\max\limits_{\|x\|=1}\|B\| = |||A|||\cdot|||B|||$

注: ① $|||I|||=1$

② $|||\cdot|||$ 称为向量范数 $\|\cdot\|$ 诱导出的矩阵范数. 后面简记为 $\|\cdot\|$.

(b) 由向量范数 $\|\cdot\|_p$ 诱导出的矩阵范数定义为

$\|A\|_p = \max\limits_{\|x\|=1}\|Ax\|_p$. 相应常用的有 $\|A\|_1$, $\|A\|_2$, $\|A\|_\infty$

Thm: $A=(a_{ij})\in\mathbb{R}^{n\times n}$, $\|A\|_1 = \max\sum\limits_{i=1}^n|a_{ij}|$ (列和范数)

Pf: $\forall x\in\mathbb{R}^n$, $\|x\|_1=1$.

$\|Ax\|_1 = \sum\limits_{i=1}^n|(Ax)_i| = \sum\limits_{i=1}^n\left|\sum\limits_{j=1}^n a_{ij}x_j\right| \leq \sum\limits_{i=1}^n\sum\limits_{j=1}^n|a_{ij}||x_j| = \sum\limits_{j=1}^n\left(\sum\limits_{i=1}^n a_{ij}\right)|x_j|$

$\leq \max\limits_{1\leq j\leq n}\sum\limits_{i=1}^n|a_{ij}|\cdot\underbrace{\sum\limits_{j=1}^n|x_j|}_{=\|x\|_1} = \max\limits_{1\leq j\leq n}\sum\limits_{i=1}^n|a_{ij}|$

即 $\|A\|_1 \leq \max\limits_{1\leq j\leq n}\sum\limits_{i=1}^n|a_{ij}|$

另一方面若列和的最大值在第k列取到, 即 $\max\limits_{1\leq j\leq n}\sum\limits_{i=1}^n|a_{ij}| = \sum\limits_{i=1}^n|a_{ik}|$

取 $x=e_k$ 则 $\|x\|_1=1$ 则 $\|Ax\|_1=\sum_{i=1}^{n}|a_{ik}|$，即 $\|A\|_1 \geq \|Ae_k\|_1=\sum_{i=1}^{n}|a_{ik}|=\max_{1\leq j\leq n}\sum_{i=1}^{n}|a_{ij}|$

$\therefore \|A\|_1=\max_{1\leq j\leq n}\sum_{i=1}^{n}|a_{ij}|$.

Thm. $\|A\|_\infty=\max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}|$ (行和范数)

Pf: $\forall x\in\mathbb{R}^n$, $\|x\|_\infty=1$

$\|Ax\|_\infty=\max_{1\leq i\leq n}\left|\sum_{j=1}^{n}a_{ij}x_j\right| \leq \max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}||x_j| \leq \max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}|$

即 $\|Ax\|_\infty \leq \max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}|$

若 $\max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}|=\sum_{j=1}^{n}|a_{kj}|$，令 $\tilde{x}=(\mathrm{sgn}(a_{k1}),\cdots,\mathrm{sgn}(a_{kn}))^T$.

$\|A\tilde{x}\|_\infty=\sum_{j=1}^{n}|a_{kj}|=\max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}|$

若A有特征值 $\lambda_1,\lambda_2,\cdots\lambda_n$，记 $\rho(A)=\max_{1\leq i\leq n}|\lambda_i|$ 为A的谱半径.

Thm: $\|A\|_2=\sqrt{\rho(A^TA)}$

Pf. A^TA 是对称半正定的，则存在正交阵 U，使得 $U^T(A^TA)U$ 为对角阵

U 的列向量为其特征向量

设 A^TA 的特征值 $0\leq\lambda_n\leq\lambda_{n-1}\leq\cdots\leq\lambda_1$；$v_1\cdots v_n$ 为对应的单位正交特征向量，构成 $\mathbb{R}^n$ 的一组基　$A^TAv_i=\lambda_iv_i$

$\forall x\in\mathbb{R}^n$, $\|x\|_2=1$，$x=\sum_{i=1}^{n}\alpha_iv_i$ 且 $\sum_{i=1}^{n}\alpha_i^2=1$ $\Leftarrow \underbrace{\left(\sum_{i=1}^{n}\alpha_iv_i^T\right)}_{x^T}\underbrace{\left(\sum_{i=1}^{n}\alpha_iv_i\right)}_{x}=\|x\|_2^2=1$

$\|Ax\|_2^2=(Ax)^TAx=x^TA^TAx$

$=\left(\sum_{i=1}^{n}\alpha_iv_i^T\right)A^TA\left(\sum_{j=1}^{n}\alpha_jv_j\right)$

$=\left(\sum_{i=1}^{n}\alpha_iv_i^T\right)\left(\sum_{j=1}^{n}\alpha_j\lambda_jv_j\right)$

$=\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_j\alpha_i\alpha_jv_i^Tv_j=\sum_{i=1}^{n}\lambda_i\alpha_i^2v_i^Tv_i=\sum_{i=1}^{n}\lambda_i\alpha_i^2\leq\lambda_1$　$\Rightarrow \|A\|_2\leq\sqrt{\rho(A^TA)}$

另一方面：$\|Av_1\|_2^2=\lambda_1=\rho(A^TA)$　$\therefore \|A\|_2\geq\sqrt{\rho(A^TA)}$

推论：若A对称，$\|A\|_2=\rho(A)$

(IV) 矩阵范数和谱半径的关系

1° $\rho(A) \le \|A\|$

若 $\rho(A)=|\lambda|$，对应的特征向量为 v，$Av=\lambda v$

则 $\rho(A)=|\lambda|=\frac{\|Av\|}{\|v\|}\le\|A\|$ 相容性

2° $\forall \varepsilon>0$，存在诱导范数 $\|\cdot\|$，s.t. $\|A\|\le\rho(A)+\varepsilon$

(V) 收敛矩阵

1° 定义：记 $A^{(m)}=(a_{ij}^{(m)})$，对于矩阵序列 $\{A^{(m)}\}_{m=1}^{\infty}$，若存在矩阵 $A=(a_{ij})$，使得 $\lim\limits_{m\to+\infty}\|A^{(m)}-A\|=0$

称 $\{A^{(m)}\}$ 收敛于 A，记作 $\lim\limits_{m\to+\infty}A^{(m)}=A$

注：$\{A^{(m)}\}$ 的收敛性与范数定义无关.

Thm: $\lim\limits_{n\to+\infty}A^{(m)}=A \Leftrightarrow \lim\limits_{m\to+\infty}a_{ij}^{(m)}=a_{ij}\quad i,j=1,2,\cdots n$

2° 定义：对于 $A=(a_{ij})$，若矩阵序列 $\{A^k\}$ 收敛于 0 矩阵，则称 A 为收敛矩阵.

Thm：A 为收敛矩阵 $\Leftrightarrow \rho(A)<1$.

Pf: "$\Rightarrow$" 若 A 为收敛矩阵. 则 $\lim\limits_{k\to+\infty}\|A^k\|=0$

记 $\lambda(A)$ 为 A 的特征值全体构成的集合. 设 $\lambda_0\in\lambda(A)$ 且 $|\lambda_0|=\rho(A)$

λ_0^k 为 A^k 的特征值. $0\le|\lambda_0|^k=|\lambda_0^k|\le\rho(A^k)\le\|A^k\|\to0 \quad\Rightarrow \lim\limits_{k\to+\infty}|\lambda_0|^k=0$

$\therefore \rho(A)=|\lambda_0|<1$.

"$\Leftarrow$" $\rho(A)<1$. 取 $\varepsilon=\frac{1-\rho(A)}{2}>0$

$\exists$ 矩阵范数 $\|\cdot\|_\alpha$，s.t. $\|A\|_\alpha\le\rho(A)+\varepsilon=\frac{1+\rho(A)}{2}<1$

$0\le\|A^k\|_\alpha\le\|A\|_\alpha^k\to0,\ k\to+\infty \quad\Rightarrow \lim\limits_{k\to+\infty}\|A^k\|_\alpha\to0.$

推论：$\|A\|<1$，则 A 为收敛矩阵.

Thm①: $\sum_{k=0}^{\infty}A^k$ 收敛 $\Leftrightarrow \rho(A)<1$.

Pf: $B^{(m)}=\sum_{k=0}^{m}A^k \quad B=\sum_{k=0}^{\infty}A^k$

"$\Rightarrow$": $0\le\|A^{(m)}\|=\|B^{(m)}-B^{(m-1)}\|\le\|B^{(m)}-B\|+\|B-B^{(m-1)}\|\to 0,\ m\to\infty$

即 $\|A^m\|\to 0,\ m\to+\infty$. A为收敛矩阵. 即 $\rho(A)<1$.

"$\Leftarrow$". $\rho(A)<1$. 取 $\varepsilon=\frac{1-\rho(A)}{2}>0$

存在矩阵范数 $\|\cdot\|_\alpha$ s.t. $\|A\|_\alpha\le\rho(A)+\varepsilon=\frac{1+\rho(A)}{2}<1$

$\|B^{(m)}-B\|_\alpha=\|\sum_{k=m+1}^{\infty}A^k\|_\alpha\le\sum_{k=m+1}^{\infty}\|A^k\|_\alpha\le\sum_{k=m+1}^{\infty}\|A\|_\alpha^k=\frac{\|A\|_\alpha^{m+1}}{1-\|A\|_\alpha}\to 0.\ m\to+\infty.$

Thm②: 当 $\sum_{k=1}^{\infty}A^k$ 收敛时, $I-A$ 可逆且 $(I-A)^{-1}=\sum_{k=0}^{+\infty}A^k$.

并存在诱导范数 $\|\cdot\|_\alpha$, $\|(I-A)^{-1}-\sum_{k=0}^{m}A^k\|_\alpha\le\frac{\|A\|_\alpha^{m+1}}{1-\|A\|_\alpha}$

Pf: 假设 $I-A$ 不可逆, 即 $(I-A)x=0$ 有非零解 x.

即 $x\ne0$. $x=Ax$

$\because \sum_{k=0}^{\infty}A^k$ 收敛 $\Rightarrow\rho(A)<1$, 取 $\varepsilon=\frac{1-\rho(A)}{2}$, $\exists$ 诱导范数 $\|\cdot\|_\alpha$. s.t. $\|A\|_\alpha<1$

$\|x\|_\alpha=\|Ax\|_\alpha\le\|A\|_\alpha\|x\|_\alpha<\|x\|_\alpha$ 矛盾.

$(I-A)\sum_{k=0}^{m}A^k=(I-A)(I+A+\cdots+A^m)=I-A^{m+1}\to I,\ m\to+\infty$

$(I-A)\sum_{k=0}^{\infty}A^k=I \Rightarrow (I-A)^{-1}=\sum_{k=0}^{\infty}A_k.\quad \|(I-A)^{-1}-\sum_{k=0}^{m}A^k\|_\alpha\le\frac{\|A\|_\alpha^{m+1}}{1-\|A\|_\alpha}$

推论: $\|A\|<1$, $\|\cdot\|$ 满足 $\|I\|<1$, 则 $I-A$ 可逆, 且 $\|(I-A)^{-1}\|\le\frac{1}{1-\|A\|}$

Pf: $\rho(A)\le\|A\|<1$.

§2.2 敏度分析和误差估计

(一) 敏度分析

① 若 $Ax=b$, 系数矩阵A非奇异. 有小扰动 $\underset{}{\delta A}$ (满足 $\|A^{-1}\delta A\|<1$), 即 $A\to A+\delta A$.

相应方程组的解也有扰动 δx. 即 $(A+\delta A)(x+\delta x)=b$.

$\Rightarrow (A+\delta A)\delta x+\delta A\cdot x=0 \Rightarrow \delta x=-(A+\delta A)^{-1}\cdot\delta A\cdot x$

$\|\delta x\|\le\|(A+\delta A)^{-1}\|\cdot\|\delta A\|\cdot\|x\|\le\dfrac{\|A^{-1}\|\cdot\|\delta A\|\cdot\|x\|}{1-\|A^{-1}\|\|\delta A\|}$

$\dfrac{\|\delta x\|}{\|x\|}\le\dfrac{\|A^{-1}\|\cdot\|A\|\dfrac{\|\delta A\|}{\|A\|}}{1-\|A^{-1}\|\cdot\|A\|\dfrac{\|\delta A\|}{\|A\|}}\approx\boxed{\|A^{-1}\|\cdot\|A\|}\dfrac{\|\delta A\|}{\|A\|}$ 相对误差放大率.

② 若 $Ax=b$, 常数项b有误差 δb, 也会引起方程解的扰动 δx.

即 $A(x+\delta x)=b+\delta b \Rightarrow A\delta x=\delta b \Rightarrow \delta x=A^{-1}\delta b$

$\|\delta x\|\le\|A^{-1}\|\cdot\|\delta b\|$ $\quad\|b\|=\|Ax\|\le\|A\|\cdot\|x\| \Rightarrow \dfrac{1}{\|x\|}\le\dfrac{\|A\|}{\|b\|}$

$\dfrac{\|\delta x\|}{\|x\|}\le\|A^{-1}\|\cdot\|A\|\dfrac{\|\delta b\|}{\|b\|}$.

定义(矩阵条件数): 若A非奇异, $\mathcal{K}(A)=Cond(A)=:\|A\|\cdot\|A^{-1}\|$ 为A的条件数.

注: ① 条件数大, 则A为病态的, 反之条件数小, 则A为良态的.

② 任意两个范数下的条件数是等价的.

③ 几何意义: $\dfrac{1}{\mathcal{K}_2(A)}=\dfrac{1}{\|A\|_2\|A^{-1}\|_2}=\min\left\{\dfrac{\|\delta A\|_2}{\|A\|_2}: A+\delta A\text{奇异}\right\}$.

即在谱范数下, 一个矩阵的条件数的倒数恰等于该矩阵与全体奇异矩阵所成集合的相对距离.

④ $\mathcal{K}(A)\ge1$ $\quad\|A\|\cdot\|A^{-1}\|\ge\|A\cdot A^{-1}\|=\|I\|=1$.

Thm: 若A非奇异. $b\in\mathbb{R}^n$ 非零. 假设 δA 满足 $\|A^{-1}\|\cdot\|\delta A\|<1$

若 x 和 $x+\delta x$ 分别是线性方程组 $Ax=b$ 和 $(A+\delta A)(x+\delta A)=(b+\delta A)$ 的解.

则 $\dfrac{\|\delta x\|}{\|x\|}\le\dfrac{\mathcal{K}(A)}{1-\mathcal{K}(A)\dfrac{\|\delta A\|}{\|A\|}}\left(\dfrac{\|\delta A\|}{\|A\|}+\dfrac{\|\delta b\|}{\|b\|}\right)$. 当 $\dfrac{\|\delta A\|}{\|A\|}$ 较小时, $\dfrac{\|\delta x\|}{\|x\|}\approx\mathcal{K}(A)\left(\dfrac{\|\delta A\|}{\|A\|}+\dfrac{\|\delta b\|}{\|b\|}\right)$

(二)精度估计

(I) 设用某种计算方法求解线性方程组 $Ax=b$ 得到的计算解为 $\hat{x}$.

令 $r=b-A\hat{x}=Ax-A\hat{x}$. $\|x-\hat{x}\|=\|A^{-1}\cdot r\|\le\|A^{-1}\|\cdot\|r\|$

即 $\frac{\|x-\hat{x}\|}{\|x\|}\le\|A^{-1}\|\cdot\|A\|\frac{\|r\|}{\|b\|}$. 取无穷范数时. $\frac{\|x-\hat{x}\|_\infty}{\|x\|_\infty}\le\kappa_\infty(A)\frac{\|r\|_\infty}{\|b\|_\infty}$. 关键在于如何计算 $\|A^{-1}\|_\infty$

(II) 盲人爬山法: $B\in\mathbb{R}^{n\times n}$. 估计 $\|B\|_1$

定义函数 $f(x)=\|Bx\|_1=\sum_{i=1}^{n}\left|\sum_{j=1}^{n}b_{ij}x_j\right|$ $D=\{x\in\mathbb{R}^n:\|x\|_1=1\}$

易证: $f(x)$ 是凸函数, D 为凸集. 则求 $\|B\|_1\Leftrightarrow$ 求 $f(x)$ 在 D 上最大值.

令 $\|x_0\|_1=1$. $\xi_i=\mathrm{sgn}\left(\sum_{j=1}^{n}b_{ij}x_j^{(0)}\right)\Rightarrow f(x)=\sum_{i=1}^{n}\xi_i\sum_{j=1}^{n}b_{ij}x_j^{(0)}$.

$\nabla f=V^TB=(B^TV)^T$. $V=(\xi_1,\xi_2,\cdots\cdots\xi_n)^T$.

$w=Bx_0$, $z=B^Tv$.

Thm 2.5.1 判断 $\|z\|_\infty\le z^Tx_0$ 则 $\|w\|_1=\|B\|_1$

若 $\|z\|_\infty>z^Tx_0$ 则 $\|Be_j\|_1>\|Bx_0\|_1$. 其中 j 满足 $|z_j|=\|z\|_\infty$. 更新 $x_0=e_j$

(III) 估计一个计算解 $\hat{x}$ 的精度.

① 利用算法 2.5.1 计算 $\|A^{-1}\|_\infty$, 得到估计值 $\tilde{v}$.

$\|A^{-1}\|_\infty=\|A^{-T}\|_1$, 令 $B=A^{-T}$ 在算法 2.5.1 中. 计算 $w=Bx$ 和 $z=B^Tv$ 相当于求解 $A^Tw=x$, $Az=v$.

② 计算 $\tilde{r}=\|r\|_\infty$, $\tilde{\beta}=\|b\|_\infty$ 和 $\|A\|_\infty\triangleq\tilde{\alpha}$

③ $\tilde{\rho}=\frac{\tilde{v}\cdot\tilde{\alpha}\cdot\tilde{r}}{\tilde{\beta}}$ 可作为计算解的相对误差的一个估计.

(IV) 若 $\hat{x}$ 的精度太低, 可将 $\hat{x}$ 作为初值, 进行迭代于 $f(x)=Ax-b$ 上.

步骤: ① 计算 $r=b-A\hat{x}$

② 求解 $Az=r$

③ 计算 $x=\hat{x}+z$.

④ 若 $\frac{\|x-\hat{x}\|_\infty}{\|x\|_\infty}\le\varepsilon$ 则结束. 否则 $\hat{x}=x$. go to ①.

§2.3 舍入误差分析

(1) 浮点数：计算机中的浮点数可表示为 $x=\pm \underset{\text{尾数}}{w}\times \underset{\text{基底}}{\beta}^{\overset{\text{阶}}{J}}$ $L\le J\le U$

$w=d_0.d_1d_2\cdots d_{t-1}$，$t$尾数位数. $0\le d_i<\beta$

若 $d_0\neq 0$. 则称为规范化浮点数.

记浮点数全体构成的集合为 $\mathcal{F}=\{0\}\cup\{x: x=\pm d_0.d_1d_2\cdots d_{t-1}\times\beta^J,\ 0\le d_i<\beta,\ d_0\neq 0,\ L\le J\le U\}$

· $\mathcal{F}$可用四元数组来刻画(β,t,L,U). $\mathcal{F}$有限集. 有 $2(\beta-1)\beta^{t-1}(U-L+1)+1$ 个元素.

· 对称分布在$[-M,-m]$和$[m,M]$中 (不等距)

$m=\beta^L$ $M=(\beta-1+\frac{\beta-1}{\beta}+\frac{\beta-1}{\beta^2}+\cdots+\frac{\beta-1}{\beta^{t-1}})\cdot\beta^U=\beta^{U+1}(1-\beta^{-t})$

IEEE standard

单精度$(2,24,-126,127)$

双精度$(2,53,-1022,1023)$ $m\approx 10^{-308}$, $M\approx 10^{308}$

(II) 实数表示成浮点数的相对误差.

记实数x在浮点数中表示为$fl(x)$

若$x=0$. $fl(x)=0$

若$m\le|x|\le M$. $fl(x)$为$\mathcal{F}$中最接近于x的浮点数. (舍入法)

若$|x|>M$, $fl(x)=NAN$ (上溢)

若$|x|<m$, $fl(x)=0$ (下溢)

引理: $x_1,x_2\in\mathcal{F}$ 且相邻, 则 $\beta^{-t}<\frac{|x_1-x_2|}{|x_1|}<\beta^{1-t}$

Pf: 不妨设 $0<x_1<x_2$.

$x_1=d_0.d_1\cdots d_{t-1}\times\beta^J \longrightarrow (d_0+\frac{d_1}{\beta}+\cdots+\frac{d_{t-1}}{\beta^{t-1}})\times\beta^J$ 最大值 $\beta(1-\beta^{-t})$

若 $1 \le x_1 < \beta(1-\beta^{-t})$. $x_2 - x_1 = \beta^{1-t} \times \beta^J$

若 $x_1 = \beta(1-\beta^{-t})$. $x_2 - x_1 = \beta^{1-t} \times \beta^J$.

$\beta^{-t} < \frac{\beta^{-t}}{1-\beta^{-t}} \le \frac{|x_1 - x_2|}{|x_1|} \le \beta^{1-t}$ □.

定义：机器精度 $u = \frac{1}{2} \times (1$和接下来最大的浮点数距离$) = \frac{1}{2}\beta^{1-t}$ 双精度：$u \approx 2^{-53} \approx 10^{-16}$

Thm：设 $m \le |x| \le M$，则 $fl(x) = x(1+\delta)$ $|\delta| < u$

Pf：不妨设 $x>0$，存在 x_L, x_R s.t. $x_L \le x \le x_R$

$\frac{|fl(x) - x|}{|x|} \le \frac{|fl(x) - x|}{|x_L|} \le \frac{1}{2} \frac{|x_R - x_L|}{|x_L|} \le \frac{1}{2}\beta^{1-t} = u$.

(III) 基本运算的舍入误差

∘表示 $+, -, \times, /$ 中任意一种运算

设 $a, b \in \mathcal{F}$，在不发生溢出的情况下，$fl(a \circ b) = (a \circ b)(1+\delta)$ $|\delta| \le u$

(IV) 向量点积的舍入误差

引理：若 $|\delta_i| \le u$ 且 $n \cdot u \le 0.01$ 则 $1 - nu \le \prod_{i=1}^{n}(1+\delta_i) \le 1 + 1.01nu$.

或 $\prod_{i=1}^{n}(1+\delta_i) = 1+\delta$. $|\delta| \le 1.01nu$

Pf：$\because |\delta_i| < u$ 有 $(1-u)^n \le \prod_{i=1}^{n}(1+\delta_i) \le (1+u)^n$

由Taylor展开：$(1-x)^n = 1 - nx + \frac{n(n-1)}{2}(1-\theta x)^{n-2}x^2$ $0<\theta<1$.

$\Rightarrow 1 - nx \le (1-x)^n \Rightarrow 1 - 1.01nu \le 1 - nu \le (1-u)^n$

利用 $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$

$= 1 + x + \frac{x}{2} \cdot x\left(1 + \frac{x}{3} + \frac{2x^2}{4!} \cdots\cdots\right)$

当 $0 \le x \le 0.01$ 时有 $1 + x \le e^x \le 1 + x + \frac{0.01}{2}xe^x \le 1 + 1.01x$

$\Rightarrow (1+u)^n \le e^{nu} \le 1 + 1.01nu$ □.

设$x\in\mathcal{F}^n, y\in\mathcal{F}^n$. 估计$|fl(x^Ty)-x^Ty|$的上界

令 $S_k = fl(\sum_{i=1}^{k} x_i y_i)$

$s_1 = x_1 y_1 (1+\gamma_1) \quad |\gamma_1| \le u$

$s_k = fl(s_{k-1} + fl(x_k y_k)) = [s_{k-1} + x_k y_k (1+\gamma_k)](1+\delta_k) \quad |\delta_k|, |\gamma_k| \le u$

于是有 $fl(x^Ty) = S_n = \sum_{i=1}^{n} x_i y_i (1+\gamma_i) \prod_{j=i}^{n} (1+\delta_j)$

$= \sum_{i=1}^{n} (1+\varepsilon_i) x_i y_i$

其中 $1+\varepsilon_i = (1+\gamma_i)\prod_{j=i}^{n}(1+\delta_j)$

定义：$\delta_1 = 0$. $|fl(x^Ty) - x^Ty| \le \sum_{i=1}^{n} |\varepsilon_i| \cdot |x_i y_i| \le 1.01\, nu \sum_{i=1}^{n} |x_i y_i|$

(V) 矩阵基本运算的舍入误差

引入记号：$|E| = (|e_{ij}|)$ 且 $|E| \le |F| \Leftrightarrow |e_{ij}| \le |f_{ij}| \quad i,j = 1,2,\cdots n$

设$A, B \in \mathcal{F}^{n\times n}$，且$\alpha \in \mathcal{F}$

$fl(\alpha A) = \alpha A + E \quad |E| \le u|\alpha A|$.

$fl(A+B) = A+B+E \quad |E| \le u|A+B|$

$fl(AB) = AB + E \quad |E| \le 1.01\, nu\, |A|\cdot|B|$.

向前误差分析：通过估计计算解与精确解之间的误差得到. 舍入误差的界与精确解有关.

向后误差分析：把计算过程产生的误差归结为具有误差的原始数据的精确运算.

(VI) 列主元消元法的舍入误差分析

思想：利用浮点数的基本运算的舍入误差理论对列主元消去法进行向后误差分析.

即证明用列主元消去法求解$Ax=b$时，它的计算解$\hat{x}$满足$(A+E)\hat{x}=b$并给出误差矩阵E的上界估计.

1. A的三角分解的舍入误差.

定理: $A\in\mathcal{F}^{n\times n}$,有三角分解且 $1.01nu\le 0.01$,则高斯消元法得到的单位下三角阵 $\tilde L$ 和上三角阵 $\tilde U$ 满足:

$$\tilde L\cdot\tilde U = A+E. \text{ 其中 } |E|\le 2.05nu|\tilde L|\cdot|\tilde U|$$

Pf: 记 $\tilde L=(\tilde l_{ij})$ $\tilde U=(\tilde u_{ij})$

由高斯消元法的具体实现可知:

$$a_{ij}^{(1)}=a_{ij} \quad i=1 \quad ; \quad a_{ij}^{(k+1)} = fl(a_{ij}^{(k)} - fl(\tilde l_{ik}\cdot\tilde u_{kj})) \quad i,j=k+1,\cdots,n$$

$$\tilde u_{ij} = a_{ij}^{(i)} = fl(a_{ij}^{(i-1)} - fl(\tilde l_{i,i-1}\cdot \tilde u_{i-1,j}))$$

$$\Rightarrow a_{ij}^{(k+1)} = (a_{ij}^{(k)} - (\tilde l_{ik}\cdot\tilde u_{kj})(1+\gamma_k))(1+\varepsilon_k) \quad |\varepsilon_k|\le u, |\gamma_k|\le u$$

$$\tilde u_{ij} = a_{ij}(1+\delta_1) - \sum_{k=1}^{i-1}(\tilde l_{ik}\tilde u_{kj})(1+\delta_k) \quad |\delta_k|\le u$$

$$\Rightarrow a_{ij} = \frac{\tilde u_{ij}}{1+\delta_i} + \sum_{k=1}^{i-1}(\tilde l_{ik}\tilde u_{kj})\frac{1+\delta_k}{1+\delta_i}$$

$$=\sum_{k=1}^{i}\tilde l_{ik}\tilde u_{kj} - e_{ij}$$

$$e_{ij} = (\tilde l_{ii}\cdot u_{ij})\frac{\delta_i}{1+\delta_i} + \sum_{k=1}^{i-1}\frac{\delta_i-\delta_k}{1+\delta_i}$$

$\because |\delta_k|\le 1.01nu<0.01$ 有 $|e_{ij}|\le \frac{2.02nu}{1-0.01}\sum_{k=1}^{i}|\tilde l_{ik}||\tilde u_{kj}| \le 2.05nu\sum_{k=1}^{i}|\tilde l_{ik}|\cdot|\tilde u_{kj}|$

推论: 设 $A\in\mathcal{F}^{n\times n}$ 可逆, 且 $1.01nu\le 0.01$, 则用列主元消元法计算得到的单位下三角阵 $\tilde L$, 上三角阵 $\tilde U$ 以及排列方阵 $\tilde P$ 满足

$\tilde L\cdot\tilde U=\tilde PA+E$, 其中 $E\le 2.05nu|\tilde L|\cdot|\tilde U|$ 交换矩阵的行或列不引入舍入误差

2. 求解三角形方程组的舍入误差

$Sx=b$ 为浮点三角形方程组, 可逆, 且假设 $1.01nu\le 0.01$

则 $Sx=b$ 的计算解满足 $(S+\tilde H)\tilde x=b$ 且 $|\tilde H|\le 1.01nu|S|$

3. 列主元消去法的舍入误差:

列主元分解: $PA+E=\tilde L\cdot\tilde U$ $\quad |E|\le 2.05nu|\tilde L||\tilde U|$

求解三角形方程组 $\tilde{L}y=\tilde{P}b$　$\tilde{U}x=y$

即得到的计算解 $\tilde{x}$ 应满足：$(\tilde{L}+F)(\tilde{U}+G)\tilde{x}=Pb$

即 $(\tilde{L}\tilde{U}+F\tilde{U}+\tilde{L}G+FG)\tilde{x}=Pb$. 其中 $|F|\le 1.01nu|\tilde{L}|$，$|G|\le 1.01nu|\tilde{U}|$

$\delta A=P^T(E+F\tilde{U}+\tilde{L}G+FG)$

$|\delta A|\le(2.05nu+1.01nu+1.01nu+1.01nu\times 1.01nu)|\tilde{L}||\tilde{U}|$

$\le 4.09|\tilde{L}|\cdot|\tilde{U}|$

$\tilde{L}$ 元素的绝对值 $\le 1\Rightarrow\|\tilde{L}\|_\infty\le n$

列主元消元法的增长因子 $\rho=\max\limits_{i,j}|\tilde{u}_{ij}|/\max\limits_{i,j}|a_{ij}|$

$\|\tilde{U}\|_\infty\le n\max\limits_{i,j}|\tilde{u}_{ij}|=\rho n\max\limits_{i,j}|a_{ij}|\le n\cdot\rho\|A\|_\infty$

Thm. $A\in\mathcal{F}^{n\times n}$ 可逆且 $1.01nu\le 0.01$

则用列主元消元法解线性方程组 $Ax=b$ 得到的计算解 $\tilde{x}$ 满足

$(A+\delta A)\tilde{x}=b$，其中 $\|\delta A\|_\infty\le 4.09n^3\rho u\|A\|_\infty$

Chap3 最小二乘问题的解法 $A\in\mathbb{R}^{m\times n}$

1° 问题是什么 2° 解的存在唯一性 3° 数值算法

4° 敏度分析和误差分析

Ex1: θ. 作 n 次测量 $x_1, x_2 \cdots x_n$ $\theta \approx \frac{x_1+x_2+\cdots+x_n}{n}$

$H=\sum_{i=1}^{n}(x_i-\theta)^2$ (线性回归)

Ex2. $\{(x_i, y_i)\}_{i=1}^{n}$

§3.1 最小二乘问题的定义. 存在唯一性

(一) 定义: 给定矩阵 $A\in\mathbb{R}^{m\times n}$ 及向量 $b\in\mathbb{R}^m$, 求解 $x\in\mathbb{R}^n$ 使得 $\|b-Ax\|=\min\limits_{y\in\mathbb{R}^n}\|b-Ay\|_2$

即为所谓的最小二乘问题, 简称 Least Square (LS) 问题

$r(x)=b-Ax$ 为残向量

最小二乘问题的解为线性方程组 $Ax=b$ 的最小二乘解.

当 $m>n$. 称为超定方程组; 当 $m<n$. 称为欠定方程组

根据 m.n 及 rank(A) 的不同, 分为以下情形:

1° $m=n$

(i) $\mathrm{rank}(A)=m=n$

(ii) $\mathrm{rank}(A)=k<m=n$

2° $m>n$

(i) $\mathrm{rank}(A)=n<m$ 只考虑这一情形

(ii) $\mathrm{rank}(A)=k<n<m$

3° $m<n$

(i) $\mathrm{rank}(A)=m<n$

(ii) $\mathrm{rank}(A)=k<m<n$

(二) 存在性和唯一性

$A\in\mathbb{R}^{m\times n}$. 记 $R(A)=\{y\in\mathbb{R}^m: y=Ax, x\in\mathbb{R}^n\}$

易证：$R(A)=\mathrm{span}(a_1,a_2\cdots a_n)$，其中 a_i 为 A 的列向量.

A 的零空间 $\mathcal{N}(A)=\{x\in\mathbb{R}^n: Ax=0\}$

$S\subset\mathbb{R}^n$ 为一个子空间，正交补 $S^\perp=\{y\in\mathbb{R}^n: y^Tx=0, \forall x\in S\}$

(I) 线性方程组 $Ax=b$ 解的存在唯一性

定理：方程组 $Ax=b$ 解存在的充要条件为 $\mathrm{rank}(A)=\mathrm{rank}([A,b])$

Pf："$\Rightarrow$" 设存在 x. s.t. $Ax=b \Rightarrow b\in R(A)$

而 $R(A)=\mathrm{span}(a_1,a_2\cdots a_n)$ 则 $R(A)=R([A,b])$

$\therefore \mathrm{rank}(A)=\dim R(A)=\dim R([A,b])=\mathrm{rank}([A,b])$

"$\Leftarrow$" 若 $\mathrm{rank}(A)=\mathrm{rank}([A,b]) \Rightarrow b\in R(A)$

$\exists\{x_i\}_{i=1}^n$，即 $b=\sum_{i=1}^n x_ia_i$. 令 $x=(x_1,\cdots x_n)^T$ 则 $Ax=b$.

定理：设方程组 $Ax=b$ 的解存在且假定 x 为其任一给定的解，

则方程组 $Ax=b$ 全部解的集合为 $x+\mathcal{N}(A)$

Pf：1° 若 y 满足 $Ay=b \Rightarrow A(y-x)=0 \Rightarrow y-x\in\mathcal{N}(A)$，$y\in x+\mathcal{N}(A)$

2° 若 $y\in x+\mathcal{N}(A)$，则 $\exists z\in\mathcal{N}(A)$，$y=x+z$

则 $Ay=A(x+z)=b$

推论：线性方程组 $Ax=b$ 解唯一的充分必要条件为 $\mathcal{N}(A)=\{0\}$.

(II) 最小二乘解的存在唯一性

$$\|b-Ax\|_2=\min_{z\in\mathbb{R}^n}\|b-Az\|_2$$

向量 b 可表示为 $b=b_1+b_2$，其中 $b_1\in R(A)$，$b_2\in R(A)^\perp$

当 $b-y\perp R(A)$ 时，$\|b-y\|_2$ 极小. 此时 $y_{min}=b_1$，b_2 为取极小时的残向量.

求解 $Ax=b_1$

定理：线性最小二乘问题的解总是存在的且其解唯一的充分必要条件是 $N(A)=\{0\}$.

Pf：$\mathbb{R}^m=R(A)\oplus R(A)^\perp \Rightarrow b=b_1+b_2$ (唯一表示) 其中 $b_1\in R(A)$, $b_2\in R(A)^\perp$

$\forall x\in\mathbb{R}^n$, $b_1-Ax\in R(A)$ 且与 b_2 正交.

$\|r(x)\|_2^2=\|b-Ax\|_2^2=\|(b_1-Ax)+b_2\|_2^2=\|b_1-Ax\|_2^2+\|b_2\|_2^2$

$\therefore$ $\|r(x)\|_2$ 达到极小 $\Leftrightarrow$ $\|b_1-Ax\|$ 达到极小

$b_1\in R(A)\Rightarrow$ 解存在

$N(A)=\{0\}\Leftrightarrow$ 解唯一.

记最小问题解集为 $\mathcal{X}_{LS}$ 则 ① $\mathcal{X}_{LS}\neq\Phi$ ② $\mathcal{X}_{LS}$ 仅有一个元素 $\Leftrightarrow N(A)=0\Leftrightarrow A$ 列满秩.

③ $A^Tb_2=0$ $<b_2\in R(A)^\perp$, $\forall z\in\mathbb{R}^n$, $Az\in R(A)$, $(Az)^Tb_2=0$ $\therefore z^T(A^Tb_2)=0\Rightarrow A^Tb_2=0>$

§3.2 正则化方法

(I) 正则化方程

定理：$x\in\mathcal{X}_{LS}\Longleftrightarrow A^TAx=A^Tb$

Pf："$\Rightarrow$" $x\in\mathcal{X}_{LS}\Rightarrow Ax=b_1$，$A^TAx=A^Tb_1=A^T(b_1+b_2)=A^Tb$

"$\Leftarrow$" $\forall y\in\mathbb{R}^n$ $\|b-A(x+y)\|_2^2=(b-A(x+y))^T(b-A(x+y))$

$=\|b-Ax\|_2^2-2y^TA^T(b-Ax)+\|Ay\|_2^2$

$=\|b-Ax\|_2^2+\|Ay\|_2^2\geq\|b-Ax\|_2^2$

(四)正则化方法步骤

①计算 $C=A^TA \quad d=A^Tb$

②用平方根法计算 $C=LL^T$

③用前代法求解 $Ly=d$，得 y

④用后代法求解 $L^Tx=y$，得 x.

(五)敏度分析

1° Moore-Penrose 广义逆

$X=(A^TA^{-1})A^Tb \triangleq A^+b$，其中 A^+ 为 A 的 Moore-Penrose 广义逆

定义：$X\in\mathbb{R}^{n\times m}$ 满足 $AXA=A$，$XAX=X$，$(AX)^T=AX$，$(XA)^T=XA$，则 X 为 A 的 M-P 广义逆.

2° b 扰动的影响.

假设 b 有扰动 δb，且 x 与 $x+\delta x$ 分别是最小二乘问题 $\min\limits_{y\in\mathbb{R}^n}\|b-Ay\|_2$ 和 $\min\limits_{y\in\mathbb{R}^n}\|(b+\delta b)-Ay\|_2$ 的解

$x=A^+b$，$x+\delta x=A^+(b+\delta b)=A^+\tilde{b}$

设 $b=b_1+b_2$，$\tilde{b}=\tilde{b}_1+\tilde{b}_2$，其中 $b_1,\tilde{b}_1\in R(A)$，$b_2,\tilde{b}_2\in R(A)^\perp$

$A^+b=A^+b_1$，$A^+\tilde{b}=A^+\tilde{b}_1$

$\|\delta x\|_2=\|A^+\tilde{b}-A^+b\|_2=\|A^+(\tilde{b}_1-b_1)\|_2\le\|A^+\|_2\cdot\|\tilde{b}_1-b_1\|_2$

$Ax=b_1 \Rightarrow \|b_1\|_2\le\|A\|_2\|x\|_2 \Rightarrow \dfrac{\|\delta x\|_2}{\|x\|_2}\le\|A^+\|_2\|A\|_2\dfrac{\|\tilde{b}_1-b_1\|_2}{\|b_1\|_2}$

注：① $\mathcal{K}_2(A)=\|A^+\|_2\cdot\|A\|$ 为最小二乘问题的条件数

② b 的扰动只有在 $R(A)$ 上的投影有影响.

③误差估计 $O(u)$

§3.3 正交化方法 (更实用的方法)

正交阵：$Q\in\mathbb{R}^{n\times n}$，$QQ^T=Q^TQ=I$.

$$\|b-Ax\|_2^2=(b-Ax)^T(b-Ax)=(b-Ax)^TQ^TQ(b-Ax)=\|Q(b-Ax)\|_2^2=\|Qb-QAx\|_2^2$$

Q：如何找到一个具有⊥消元性质的正交矩阵？

(一) 初等正交变换

(I) Householder 变换.

1° 定义：设 $w\in\mathbb{R}^n$ 且 $\|w\|_2=1$. 定义 $H\in\mathbb{R}^{n\times n}$ 为 $H=I-2ww^T$

称 H 为 Householder 变换/矩阵.

2° 性质：① 对称性：$H=H^T$

② 正交性：$HH^T=I=H^TH$

③ 对合性：$H^2=I$

④ 反射性：$\forall x\in\mathbb{R}^n$，Hx 是 x 关于 w 的垂直超平面 $\text{span}\{w\}^\perp$ 的镜像反射.

Pf：$\mathbb{R}^n=\text{span}\{w\}\oplus\text{span}\{w\}^\perp$

$\forall x\in\mathbb{R}^n$，$x=\alpha w+y$ 其中 $\alpha\in\mathbb{R}$，$y^Tw=0$.

$Hx=(I-2ww^T)(\alpha w+y)$

$=\alpha w+y-2\alpha ww^Tw-2ww^Ty$

$=-\alpha w+y$.

Householder 变换也称为镜像变换或初等反射矩阵.

3° 定理：设 $x\in\mathbb{R}^n$，$x\neq 0$. 则存在 $w\in\mathbb{R}^n$，$\|w\|_2=1$，使得 Householder 变换 $H=I-2ww^T$ 满足 $Hx=\alpha e_1$，其中 $\alpha=\pm\|x\|_2$

Pf：$Hx=\alpha e_1$ 与 x 关于 $\text{span}\{w\}^\perp$ 镜像反射

$$w=\frac{x-Hx}{\|x-Hx\|_2}=\frac{x-\alpha e_1}{\|x-\alpha e_1\|_2}$$

又 $\because Hx=(I-2ww^T)x=x-2ww^Tx=\alpha e_1 \Rightarrow x-\alpha e_1=2ww^Tx=2w^Txw$

$\Rightarrow 2w^Txw=\|x-\alpha e_1\|_2 w \quad w\neq 0$

$\Rightarrow 2w^Tx=\|x-\alpha e_1\|_2 \Rightarrow 2(x-\alpha e_1)^Tx=\|x-\alpha e_1\|_2^2$

设 $x=(x_1,x_2,\cdots x_n)^T \quad x-\alpha e_1=(x_1-\alpha, x_2 \cdots x_n)$

$\therefore 2((x_1-\alpha)x_1+x_2^2+x_3^2+\cdots+x_n^2)=(x_1-\alpha)^2+x_2^2+\cdots+x_n^2$

$\Rightarrow 2(\|x\|_2^2-\alpha x_1)=\|x\|_2^2-2\alpha x_1+\alpha^2$

$\Rightarrow \alpha^2=\|x\|_2^2 \Rightarrow \alpha=\pm\|x\|_2$

4° 构造步骤

① $v=x-\alpha e_1$, $\alpha=\|x\|_2$

$v_1=x_1-\alpha=x_1-\|x\|_2$ (*1)

等价变形: $v_1=x_1-\|x\|_2=\dfrac{x_1^2-\|x\|_2^2}{x_1+\|x\|_2}=-\dfrac{x_2^2+\cdots+x_n^2}{x_1+\|x\|_2}$ (*2)

若 $x_1>0$. 用 *2 算 v_1.

若 $x_1<0$. 用 *1 算 v_1.

② $w=\dfrac{v}{\|v\|_2}$

③ $H=I-2ww^T=I-\dfrac{2vv^T}{v^Tv}=I-\beta vv^T$, $\beta=\dfrac{2}{v^Tv}$.

注: ① 不用求 w. 只需求 v 和 β.

② x 分量大, 平方时可能上溢
下溢发生时, v^Tv 可能为0 $\Bigg\} \rightarrow \dfrac{x}{\|x\|_\infty}$ 代替 x

③ 实际计算中. 把 v 规范化为 $v_1=1$. 存 $v(2:n)$ 即可.

算法3.2.1

①可将向量中任意若干相邻元素化为0.

②实际计算中不需要H的具体形式.

③数值性态良好.(舍入误差分析 $O(u)$)

(II) Givens变换.(选择性化一些元素为0)

$$
\begin{array}{c|c}
\begin{pmatrix}
1 & & \vdots & & \vdots & & \\
 & \ddots & \vdots & & \vdots & & \\
\cdots & \cdots & c & \cdots & s & \cdots & \cdots \\
 & & \vdots & & \vdots & & \\
\cdots & \cdots & -s & \cdots & c & \cdots & \cdots \\
 & & \vdots & & \vdots & \ddots & \\
 & & \vdots & & \vdots & & 1
\end{pmatrix} & \begin{matrix} \\ \\ i \\ \\ k \\ \\ \\ \end{matrix} \\
\begin{matrix} \qquad\qquad i \qquad\quad k \qquad\qquad \end{matrix} &
\end{array}
= G(i,k,\theta) = I + s(e_i e_k^T - e_k e_i^T) + (c-1)(e_i e_i^T + e_k e_k^T)
$$

其中 $s=\sin\theta$, $c=\cos\theta$. 易证 $G(i,k,\theta)$ 为正交阵.

$\forall x \in \mathbb{R}^n$, $y=G(i,k,\theta)x$. 则有

$$
\begin{cases}
y_i = cx_i + sx_k \\
y_k = -sx_i + cx_k \\
y_j = x_j \quad j \neq i,k.
\end{cases}
$$

若取θ满足

$c=\cos\theta=\dfrac{x_i}{\sqrt{x_i^2+x_k^2}}$, $s=\sin\theta=\dfrac{x_k}{\sqrt{x_i^2+x_k^2}}$

$\Rightarrow y_i=\sqrt{x_i^2+x_k^2}$. $y_k=0$.

几何意义: $G(i,k,\theta)$ 是在 (i,k) 平面内将 x 按逆时针旋转θ度, 所以Givens变换又称平面旋转变换

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$$

$$\Rightarrow c=\frac{a}{\sqrt{a^2+b^2}} \qquad s=\frac{b}{\sqrt{a^2+b^2}}$$

算法3.2.2 为避免溢出 $b\neq 0$.

① $|b|>|a|$ $\tau=a/b$, $s=\dfrac{1}{\sqrt{1+\tau^2}}$ $c=s\tau$.

② $|a|\geq|b|$ $\tau=b/a$, $c=\dfrac{1}{\sqrt{1+\tau^2}}$ $s=c\tau$.

注: ① r 不一定为正

② 数值性态良好.

(二)正交化方法

基本思想:选取适当的正交阵Q. s.t.原来的LS问题化为易求解的LS问题.

(I) QR分解定理.

设$A\in\mathbb{R}^{m\times n}$，则$A$有QR分解$A=\overset{m\times m}{Q}\begin{pmatrix}\overset{n\times n}{R}\\0\end{pmatrix}$，其中$Q\in\mathbb{R}^{m\times m}$为正交阵，$R\in\mathbb{R}^{n\times n}$为具有非负对角元的上三角阵

且当$m=n$且A非奇异时，上述分解是唯一的.

Pf: 对n用数学归纳法. 当$n=1$时，存在Householder变换H，s.t. $HA=\alpha e$，即$A=H^T\begin{pmatrix}\alpha\\0\\\vdots\\0\end{pmatrix}$ $\alpha=\|A\|_2$

假设对$n-1$成立.

即$A=[A_1,\ v]$，$A_1\in\mathbb{R}^{m\times(n-1)}$. 根据假设$A_1$有QR分解$A_1=Q_1\begin{pmatrix}R_1\\0\end{pmatrix}$

其中$Q_1\in\mathbb{R}^{m\times m}$正交阵. $R_1\in\mathbb{R}^{(n-1)\times(n-1)}$为具有非负对角元的上三角阵

$Q_1^TA=[Q_1^TA_1,\ Q_1^Tv]=\left[\begin{matrix}R_1\\0\end{matrix},\ Q_1^Tv\right]$

$y\in Q_1^Tv\in\mathbb{R}^m$. 设$y=\begin{pmatrix}y_1\\\vdots\\y_{n-1}\\y_n\\\vdots\\y_m\end{pmatrix}$ $\begin{pmatrix}y_n\\\vdots\\y_m\end{pmatrix}$构造$\tilde{H}_1\in\mathbb{R}^{(m-n+1)\times(m-n+1)}$, s.t. $\tilde{H}_1\begin{pmatrix}y_n\\\vdots\\y_m\end{pmatrix}=\begin{pmatrix}\alpha_1\\0\\\vdots\\0\end{pmatrix}$ $\alpha_1>0$.

令$H_1=\begin{pmatrix}I_{n-1}&\\&\tilde{H}_1\end{pmatrix}\in\mathbb{R}^{m\times m}$为正交阵.

$H_1Q_1^TA=\begin{pmatrix}I_{n-1}&\\&\tilde{H}_1\end{pmatrix}\begin{pmatrix}R_1&Q_1^Tv\\0&\end{pmatrix}=\begin{pmatrix}R_1&\begin{matrix}y_1\\\vdots\\y_{n-1}\end{matrix}\\0&\begin{matrix}\alpha_1\\0\end{matrix}\end{pmatrix}=\begin{pmatrix}R\\0\end{pmatrix}$

$Q=Q_1H_1$为正交阵. $A=Q\begin{pmatrix}R\\0\end{pmatrix}$

唯一性: $m=n$. A非奇异.

若QR分解不唯一，则$A=QR=\tilde{Q}\tilde{R}$，其中$Q.\tilde{Q}$为正交阵. $R,\tilde{R}$为具有正对角元的上三角阵

$\underbrace{\tilde{Q}^TQ}_{正交}=\underbrace{\tilde{R}R^{-1}}_{上三角}\triangleq B$ $\therefore B$单位阵 $\therefore\tilde{Q}=Q.$ $\tilde{R}=R$

(II) 正交化方法求解最小二乘问题

$A\in\mathbb{R}^{m\times n}$，$\overset{\text{解的唯一性}}{\mathrm{rank}(A)=n}$，$b\in\mathbb{R}^m$

存在正交矩阵Q. 使得$A=Q\begin{pmatrix}R\\0\end{pmatrix}$，$Q\in\mathbb{R}^{m\times m}$为正交阵.

$\|Ax-b\|_2^2=\|Q^TAx-Q^Tb\|_2^2=\left\|\begin{pmatrix}R\\0\end{pmatrix}x-Q^Tb\right\|_2^2$

$Q=[\underset{m\times n}{Q_1},\ \underset{m\times(m-n)}{Q_2}]$ $Q^Tb=\begin{pmatrix}Q_1^T\\Q_2^T\end{pmatrix}b=\begin{pmatrix}c_1\\c_2\end{pmatrix}\begin{matrix}\to\mathbb{R}^n\\\to\mathbb{R}^{m-n}\end{matrix}$

$\therefore \|Ax-b\|_2^2 = \left\|\begin{pmatrix} Rx \\ 0 \end{pmatrix} - \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}\right\|_2^2 = \|Rx-c_1\|_2^2 + \|c_2\|_2^2$

$\therefore$ x是LS问题的解当且仅当x为$Rx=c_1$的解.

基本步骤:

① 计算A的QR分解(关键)

② 计算$c_1 = Q_1^T b$.

③ 求解上三角形方程组$Rx = c_1 \Rightarrow x$.

(III) 利用Householder变换实现A的QR分解 $\alpha_1 > 0$

$A \xrightarrow{H_n \cdots H_2 H_1} \begin{pmatrix} R \\ 0 \end{pmatrix}$ $\quad H_k = \begin{pmatrix} I & 0 \\ 0 & \tilde{H}_k \end{pmatrix}$

则$A = Q\begin{pmatrix} R \\ 0 \end{pmatrix}$. $Q = H_1 H_2 \cdots H_n$.

注: ① 数值性态良好

Chap 4 线性方程组的古典迭代法

直接法：经过有限次运算得到解

迭代法：按照某种规则构造一个向量序列 $\{x^{(k)}\}$, s.t. $\lim\limits_{k\to\infty} x^{(k)} = x^*$. 且 x^* 满足 $Ax^*=b$.

§4.1 迭代法

(I) 定义：线性方程组 $Ax=b$, A 可分解为 $A=A_1-A_2$, 其中 A_1 为非奇异矩阵.

可得到同解方程组 $A_1x=A_2x+b \Rightarrow x=\underbrace{A_1^{-1}A_2}_{M}x+\underbrace{A_1^{-1}b}_{g}$

可简写为 $x=Mx+g$.

任给初始向量 $x^{(0)}$. 作迭代 $x^{(k+1)}=Mx^{(k)}+g$. (*1)

(*1) 为迭代公式. M 迭代矩阵, g 为常数项 单步线性定常迭代

由(*1)生成的向量序列 $\{x^{(k)}\}$ 对于 $\lim\limits_{k\to\infty} x^{(k)}=x^*$. 则 $x^*=Mx^*+g$ 由于同解性 $Ax^*=b$

定义：若 $x=Mx+g$ 是 $Ax=b$ 的同解方程组. $\forall x^{(0)}\in\mathbb{R}^n$. 作迭代 $x^{(k+1)}=Mx^{(k)}+g$

所得到的向量序列 $\{x^{(k)}\}$ 恒有极限, 则称迭代是收敛的.

(II) 收敛性理论

定理：解 $Ax=b$ 的迭代法 $x^{(k+1)}=Mx^{(k)}+g$ 收敛的充分必要条件是 $\rho(M)<1$.

Pf: $\forall x^{(0)}\in\mathbb{R}^n$. 由迭代法得到 $\{x^{(k)}\}$. 若 x^* 是 $Ax=b$ 的解, 则 x^* 也满足 $x^*=Mx^*+g$

$\|x^*-x^{(k+1)}\|=\|Mx^*+g-(Mx^{(k)}+g)\|=\|M(x^*-x^{(k)})\| = \cdots = \|M^{k+1}(x^*-x^{(0)})\|$

由 $x^{(0)}$ 的任意性, $\lim\limits_{k\to\infty} x^{(k+1)}=x^* \Leftrightarrow \lim\limits_{k\to\infty}\|M^{k+1}\|=0 \Leftrightarrow M$ 为收敛矩阵 $\Leftrightarrow \rho(M)<1$.

推论：若迭代矩阵 M 满足 $\|M\|<1$. 则迭代法 $x^{(k+1)}=Mx^{(k)}+g$ 收敛

注：通常选迭代矩阵 1 范数和 ∞ 范数.

(III) 误差估计

Thm 1. 解 $Ax=b$ 的迭代 $x^{(k+1)}=Mx^{(k)}+g$，若 $q\triangleq\|M\|<1$，x^* 为 $Ax=b$ 的解

则迭代法(*)所产生的近似解 $x^{(k)}$ 与准确解 x^* 的误差有如下估计式：

$$\|x^{(k)}-x^*\|\le\frac{q^k}{1-q}\|x^{(1)}-x^{(0)}\|.$$

Pf: $\|x^{(k)}-x^*\|=\|Mx^{(k-1)}+g-(Mx^*+g)\|$

$=\|M(x^{(k-1)}-x^*)\|$

$\le q\|x^{(k-1)}-x^*\|\le q^k\|x^{(0)}-x^*\|$

$\|x^{(0)}-x^*\|\le\|x^{(0)}-x^{(1)}\|+\|x^{(1)}-x^*\|$

$\le\|x^{(0)}-x^{(1)}\|+q\|x^{(0)}-x^*\|$ $\therefore\ \|x^{(k)}-x^*\|\le\frac{q^k}{1-q}\|x^{(1)}-x^{(0)}\|.$

Thm 2. $\|x^{(k)}-x^*\|\le\frac{q}{1-q}\|x^{(k-1)}-x^{(k)}\|$

注：① 只要 $\|M\|$ 不接近于1，则用 $\|x^{(k-1)}-x^{(k)}\|$ 来控制误差 $\|x^{(k)}-x^*\|$.

② 迭代法通常不能得到方程组的精确解.

对于预先给定的精度控制小量 $\varepsilon>0$，只要 $\|x^{(k)}-x^{(k-1)}\|<\varepsilon$

可视 $x^{(k)}$ 为方程组的近似，否则迭代继续进行，直到满足精度要求.

③ 如果迭代收敛，由于与初始选取无关，因此迭代过程中出现运算错误

也不影响迭代序列的收敛.

(IV) 收敛速度

判断迭代"好"与"不好"的标准为 $\|x^{(k)}-x^*\|$ 收敛于0的快慢.

给定 $\delta>0$，满足 $\|x^{(k)}-x^*\|<\delta\|x^{(0)}-x^*\|$ 所需迭代次数的大小.

$\|x^{(k)}-x^*\|\le\|M^k\|\cdot\|x^{(0)}-x^*\|.$

引理：$\lim\limits_{k\to\infty}\|M^k\|^{\frac{1}{k}}=\rho(M)$

Pf：λ为M的特征值，λ^k为M^k的特征值

$(\rho(M))^k\le\rho(M^k)\le\|M^k\|\Rightarrow\rho(M)\le\|M^k\|^{\frac{1}{k}}$

另一方面．$\forall\varepsilon>0$，$B_\varepsilon=\frac{1}{\rho(M)+\varepsilon}M\Rightarrow\rho(B_\varepsilon)<1\Rightarrow B_\varepsilon$为收敛矩阵．

即$\lim\limits_{k\to\infty}B_\varepsilon^k=0$．则存在自然数$K$．当$k>K$时$\|B_\varepsilon^k\|<1$．即$\|M^k\|\le(\rho(M)+\varepsilon)^k$

$\forall\varepsilon>0$．$\exists K$ 当$k>K$时 $\rho(M)\le\|M^k\|^{\frac{1}{k}}\le\rho(M)+\varepsilon$．

由引理：对充分大k．用$\|M^k\|<\delta$ 刻画迭代速度时，可近似地按$(\rho(M)^k)<\delta$来决定k．

$k\ge\frac{\ln\delta}{\ln\rho(M)}=\frac{-\ln\delta}{-\ln\rho(M)}$ 与k与$(-\ln\rho(M))$成反比

$\forall\delta>0$，只要$k\ge\frac{-\ln\delta}{-\ln\rho(M)}$ 有$\|x^{(k)}-x^*\|\le\delta\|x^{(0)}-x^*\|$．

记$R_\infty(M)=-\ln\rho(M)$为迭代的渐近收敛速度．

$R_k(M)=\frac{-\ln\|M^k\|}{k}$为$k$次迭代平均收敛速度．

注：① $\rho(M)$越小．$R_\infty(M)$越大．收敛速度越高．

② 对迭代矩阵分别为M_1．M_2的两种迭代，值$R_\infty(M_1)/R_\infty(M_2)$称为两种迭代的速率比

例如：$\rho(M_1)=\rho^2(M_2)<1$．则$R_\infty(M_1)/R_\infty(M_2)=2$．

(V) 迭代构造

$A=(a_{ij})$ $D=\begin{pmatrix}a_{11}&&\\&\ddots&\\&&a_{nn}\end{pmatrix}$ L,U分别为上下三角部分．

$A=D-L-U$

① Jacobi迭代 $A=D-(L+U)$

迭代 $x^{(k+1)}=D^{-1}(L+U)x^{(k)}+D^{-1}b\triangleq Bx^{(k)}+g$．

② Gauss-Seidel迭代：$A=(D-L)+U$

$$x^{(k+1)}=(D-L)^{-1}Ux^{(k)}+(D-L)^{-1}b \triangleq Mx^{(k)}+g$$

③松弛迭代：$A=\frac{1}{w}D-\frac{1-w}{w}D-L-U$

$$=(\frac{1}{w}D-L)-(\frac{1-w}{w}D+U)$$

$$x^{(k+1)}=(\frac{1}{w}D-L)^{-1}(\frac{1-w}{w}D+U)x^{(k)}+(\frac{1}{w}D-L)^{-1}b$$

w：松弛因子 → 最佳松弛因子.

准备知识：一些矩阵的概念

定义1：对$A=(a_{ij})$. 若有$|a_{ii}|\geq\sum_{j\neq i}|a_{ij}|$. $i=1,2,\cdots n$. 至少有一个i有严格不等号成立.

则称A为弱严格对角占优的. 若有$|a_{ii}|>\sum_{j\neq i}|a_{ij}|$ 则称A为严格对角占优的.

命题1：若A为严格对角占优的，则A必非奇异.

Pf：若A奇异，则$\exists x\in\mathbb{R}^n$. $x\neq 0$ 使得$Ax=0$.

不妨设$|x_k|=\|x\|_\infty=\max_{1\leq i\leq n}|x_i|\neq 0$

$Ax=0$ 的第k行：$|\sum_{j=1}^{n}a_{kj}x_j|=|a_{kk}x_k+\sum_{j\neq k}a_{kj}x_j|$

$$\geq|a_{kk}|\cdot|x_k|-\sum_{j\neq k}|a_{kj}|\cdot|x_j|$$

$$\geq|a_{kk}|\cdot|x_k|-\sum_{j\neq k}|a_{kj}|\cdot|x_k|$$

>0. 矛盾.

定义2：设$A=(a_{ij})$, 若存在排列矩阵P. 使得$PAP^T=\begin{pmatrix}A_{11}&0\\A_{21}&A_{22}\end{pmatrix}$ 则称A为可约矩阵，否则称A为不可约矩阵.

等价说法：$w=\{1,2,3,\cdots,n\}$有两个非空子集合满足$S\cup T=w$, $S\cap T=\Phi$. s.t. $a_{ij}=0$ $(i\in S, j\in T)$ 则称A为可约的.

Thm：A不可约⟺A对应的有向图是连通的.

命题2：若A为弱严格对角占优的不可约阵，则A必非奇异.

pf：假设A奇异，则$\exists x\neq 0$. s.t. $Ax=0$，不失一般性，假设$\|x\|_\infty=1$

①可证不可能所有分量都满足$|x_i|=1$. 否则若A的k行满足$|a_{kk}|>\sum_{j\neq k}|a_{kj}|$

$Ax=0$的第k行：$a_{kk}x_k+\sum_{j\neq k}a_{kj}x_j=0$ 而$|a_{kk}x_k+\sum_{j\neq k}a_{kj}x_j|>|a_{kk}|-\sum_{j\neq k}|a_{kj}|>0$ 矛盾.

②记$W_1=\{i:|x_i|=1\}$ $W_2=\{i:|x_i|<1\}$

则知$W_1\cup W_2=\{1,2,3\cdots n\}$. $W_1\cap W_2=\varnothing$.

若$i\in W_1$，则$Ax=0$的第i行

$$|a_{ii}x_i+\sum_{j\neq i}a_{ij}x_j|\geq|a_{ii}|\cdot|x_i|-\sum_{\substack{j\in W_1\\ j\neq i}}|a_{ij}|\cdot|x_j|-\sum_{j\in W_2}|a_{ij}|\cdot|x_j|$$

假设$m\in W_2$，$a_{im}\neq 0$则

$$|a_{ii}x_i+\sum_{j\neq i}a_{ij}x_j|\geq|a_{ii}|-\sum_{\substack{j\in W_1\\ j\neq i}}|a_{ij}|-\sum_{j\in W_2}|a_{ij}|\cdot|x_j|$$

$$>|a_{ii}|-\sum_{\substack{j\in W_1\\ j\neq i}}|a_{ij}|-\sum_{j\in W_2}|a_{ij}|=|a_{ii}|-\sum_{j\neq i}|a_{ij}|\text{ 矛盾}.$$

§4.2 Jacobi迭代

(一)迭代公式 $A=D-(L+U)$ $x^{(k+1)}=D^{-1}(L+U)x^{(k)}+D^{-1}b$

计算公式：$x_1^{(k+1)}=-\frac{1}{a_{11}}(a_{12}x_2^{(k)}+a_{13}x_3^{(k)}+\cdots+a_{1n}x_n^{(k)}-b_1)$

$$x_2^{(k+1)}=-\frac{1}{a_{22}}(a_{21}x_1^{(k)}+a_{23}x_3^{(k)}+\cdots+a_{2n}x_n^{(k)}-b_2)$$
$$\vdots$$
$$x_n^{(k+1)}=-\frac{1}{a_{nn}}(a_{n1}x_1^{(k)}+a_{n2}x_2^{(k)}+\cdots+a_{nn-1}x_{n-1}^{(k)}-b_n)$$

(二)收敛性

(1)定理：解$Ax=b$与$A^Tx=b$的Jacobi迭代有相同的收敛性.

Pf: 记 $Ax=b$ 的迭代矩阵为 B　$B=D^{-1}(L+U)=D^{-1}(D-A)=I-D^{-1}A$.

记 $A^Tx=b$ 的迭代矩阵为 B_1　$B_1=I-D^{-1}A^T$

相似. 特征值相同.

$\rho(B_1)=\rho(I-D^{-1}A^T)\overset{trans}{=\!=\!=}\rho(I-AD^{-1})=\rho(D^{-1}(I-AD^{-1})D)=\rho(I-D^{-1}A)=\rho(B)$

(II) 定理: 若 A 满足下列条件之一, 则其 Jacobi 迭代收敛.

① A 是严格对角占优的

② A 为弱严格对角占优的不可约阵.

Pf: $\forall i$, $a_{ii}\neq 0$

① 迭代矩阵 $B=(b_{ij})$. $B=D^{-1}(L+U)$　$b_{ij}=-\frac{a_{ij}}{a_{ii}}\ (i\neq j)$　$b_{ii}=0$

$\|B\|_\infty=\max\limits_i\sum\limits_{j=1}^n|b_{ij}|=\max\limits_i\sum\limits_{j=1}^n\frac{|a_{ij}|}{|a_{ii}|}<1$ ⇒ 迭代法收敛.

法二: 设 λ 为 B 的特征值. $|\lambda|\geq 1$　$\det(\lambda I-B)=0$

$\lambda I-B=\lambda I-D^{-1}(L+U)=D^{-1}(\lambda D-L-U)$

$\lambda D-L-U$ 仍严格对角占优 ⇒ $\lambda D-L-U$ 非奇异 ⇒ $\det(\lambda I-B)\neq 0$ 矛盾.

$\Rightarrow \rho(B)<1$

② A 是弱严格对角占优的不可约阵, λ 为 B 特征值. $|\lambda|\geq 1$

$\lambda D-L-U$ 仍是弱严格对角占优的不可约阵 ⇒ $\lambda D-L-U$ 为非奇异. ⇒ $\det(\lambda I-B)\neq 0$. 矛盾. ⇒ $\rho(B)<1$.

(III) 定理: 若 $Ax=b$ 的系数矩阵 A 对称且 $a_{ii}>0$, $i=1,2,\cdots n$

则 $Ax=b$ 的 Jacobi 迭代收敛的充要条件是 $A>0$, $2D-A>0$

Pf: 迭代矩阵 $B=D^{-1}(L+U)=I-D^{-1}A=D^{-\frac{1}{2}}(I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}$

A 对称 ⇒ $I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 对称. 其特征值为实数.

又 $B\overset{S}{\sim}I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ ⇒ B 的特征值为实数.

Jacobi 迭代收敛 ⟺ $\rho(B)<1$ ⟺ $I-B$ 与 $I+B$ 特征值均为正实数.

$I-B=D^{-\frac{1}{2}}(D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}$, $I+B=D^{-\frac{1}{2}}(2I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}$

$\therefore D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$, $2I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 正定. $\Leftrightarrow A$, $2D-A$ 正定

§4.3 Gauss-Seidel迭代

(一)迭代公式: $A=(D-L)-U$ $\quad x^{(k+1)}=(D-L)^{-1}Ux^{(k)}+(D-L)^{-1}b$

实际计算: 改写: $(D-L)x^{(k+1)}=Ux^{(k)}+b$

$\Rightarrow Dx^{(k+1)}=Lx^{(k+1)}+Ux^{(k)}+b$

$\Rightarrow x^{(k+1)}=D^{-1}Lx^{(k+1)}+D^{-1}Ux^{(k)}+D^{-1}b$

计算公式: $x_1^{(k+1)}=-\frac{1}{a_{11}}(a_{12}x_2^{(k)}+a_{13}x_3^{(k)}+\cdots+a_{1n}x_n^{(k)}-b_1)$

$x_2^{(k+1)}=-\frac{1}{a_{22}}(a_{21}x_1^{(k+1)}+a_{23}x_3^{(k)}+\cdots+a_{2n}x_n^{(k)}-b_2)$

$\vdots$

$x_n^{(k+1)}=-\frac{1}{a_{nn}}(a_{n1}x_1^{(k+1)}+a_{n2}x_2^{(k+1)}+\cdots+a_{nn-1}x_{n-1}^{(k+1)}-b_n)$

(二)收敛性

(I)定理: $Ax=b$的系数矩阵A若为严格对角占优或弱严格对角占优的不可约阵.

则解$Ax=b$的Gauss-Seidel迭代收敛.

(II)定理: 若$Ax=b$的系数矩阵A对称正定, 则其Gauss-Seidel迭代收敛.

注: 松弛迭代对应定理的特殊情形.

§4.4 松弛迭代

G-S迭代: $x^{(k+1)}=D^{-1}Lx^{(k+1)}+D^{-1}Ux^{(k)}+D^{-1}b$

$=x^{(k)}+\underline{D^{-1}(Lx^{(k+1)}+(U-D)x^{(k)}+b)}$ 修正项.

$x^{(k+1)}$可看作$x^{(k)}$加上修正项而得，修正因子为1.

松弛迭代：$x^{(k+1)} = x^{(k)} + w(Lx^{(k+1)} + (U-D)x^{(k)} + b)$

$$= (1-w)x^{(k)} + w(D^{-1}Lx^{(k+1)} + D^{-1}Ux^{(k)} + D^{-1}b)$$

计算式：$x_1^{(k+1)} = (1-w)x_1^{(k)} - \frac{w}{a_{11}}(a_{12}x_2^{(k)} + \cdots + a_{1n}x_n^{(k)} - b_1)$

$$x_2^{(k+1)} = (1-w)x_2^{(k)} - \frac{w}{a_{22}}(a_{21}x_1^{(k+1)} + \cdots + a_{2n}x_n^{(k)} - b_2)$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$x_n^{(k+1)} = (1-w)x_n^{(k)} - \frac{w}{a_{nn}}(a_{n1}x_1^{(k+1)} + \cdots + a_{n,n-1}x_{n-1}^{(k)} - b_n)$$

迭代矩阵：$x^{(k+1)} = ((1-w)I + wD^{-1}U)x^{(k)} + wD^{-1}Lx^{(k+1)} + wD^{-1}b.$

$$x^{(k+1)} = (I - wD^{-1}L)^{-1}((1-w)I + wD^{-1}U)x^{(k)} + w(I - wD^{-1}L)^{-1}D^{-1}b$$

迭代矩阵 $L_w = (I - wD^{-1}L)^{-1}((1-w)I - wD^{-1}U)$

$$= (D^{-1}(D-wL))^{-1}[D^{-1}((1-w)D - wU)]$$

$$= (D-wL)^{-1}((1-w)D - wU)$$

称为松弛迭代，w为松弛因子

当$w>1$时，超松弛迭代 (over-relaxation)

当$0<w<1$时，低松弛迭代 (under-relaxation)

当$w=1$时，G-S迭代

统称为逐次的超松弛迭代
successive over-relaxation (SOR)

(二) 收敛性.

(1) 定理：松弛迭代收敛的必要条件是 $0<w<2$

Pf: 迭代法收敛 $\Rightarrow \rho(L_w) < 1$

设$\lambda_1, \lambda_2, \cdots \lambda_n$为$L_w$的特征值. $|\det(L_w)| = |\lambda_1 \ldots \lambda_n| < 1$

$\det(L_w) = \det((I - wD^{-1}L)^{-1})\det((1-w)I - wD^{-1}U) = (1-w)^n.$

$\therefore 0 < w < 2.$

(II)定理：$Ax=b$ 的系数矩阵 A 若为严格对角占优或弱严格对角占优的不可约阵，

则解 $Ax=b$ 的松弛迭代当 $0<w\le 1$ 时收敛。

注：两个严格对角占优矩阵加起来不一定是严格对角占优。

(III)定理：若 $Ax=b$ 的系数矩阵 A 对称正定，则解 $Ax=b$ 的松弛迭代对 $0<w<2$ 均收敛。

Pf：设 λ 为迭代矩阵 L_w 的任一特征值，x 为其对应的特征向量，即

$(D-wL)^{-1}((1-w)D+wU)x=\lambda x$. $x\neq 0$

即 $((1-w)D+wU)x=\lambda(D-wL)x$

共轭转置

$\Rightarrow x^*((1-w)D+wU)x=\lambda x^*(D-wL)x$ $\quad \lambda=\dfrac{x^*((1-w)D+wU)x}{x^*(D-wL)x}$

设 $x^*Dx=\delta$. $x^*Lx=\alpha+i\beta$. $x^*Ux=\alpha-i\beta$. $(L^T=U)$

$\lambda=\dfrac{(1-w)\delta+w(\alpha-i\beta)}{\delta-w(\alpha+i\beta)}$ $\quad |\lambda|^2=\dfrac{((1-w)\delta+w\alpha)^2+(w\beta)^2}{(\delta-w\alpha)^2+(w\beta)^2}$ $\quad$ 分母－分子 $=(\delta-w\alpha)^2-((1-w)\delta+w\alpha)^2$

A 正定. $a_{ii}>0$. $\Rightarrow \delta=x^*Dx>0$. $\qquad\qquad =(2-w)\delta w(\delta-2\alpha)$.

$x^*Ax=x^*(D-L-U)x=\delta-2\alpha>0$.

$\Rightarrow 0<w<2$ 时 $(2-w)\delta w(\delta-2\alpha)>0$. $\Rightarrow |\lambda|^2<1 \Rightarrow \rho(L_w)<1$.

(三)最佳松弛因子

当 w 为何值时，迭代收敛速度最快。收敛速度：$k\ge -\dfrac{\ln\varepsilon}{-\ln\rho(L_w)}$ $\quad$ w 取何值时，$\rho(L_w)$ 小。

①实际计算中，通常选用不同的 w 值，用相同的初始向量进行计算：

迭代相同次数，比较残向量最小的 w 作为松弛因子。

②根据 $\rho(L_w)$ 随 w 的变化曲线

在不能得到准确的最佳松弛因子时，宁可取得稍大一些。

Chap 5 共轭梯度法（A正定）

(一) $Ax=b$. $A\in\mathbb{R}^{n\times n}$，对称正定，$x,b\in\mathbb{R}^n$

令 $\varphi(x)=x^TAx-2b^Tx=\sum_{i,j}x_ia_{ij}x_j-2\sum_i b_ix_i$

定理：若 $Ax=b$ 的系数矩阵A对称正定，则 $Ax=b$ 的求解等价于求 $\varphi(x)$ 的极小值点.

Pf：1° 若 x^* 为 $\varphi(x)$ 的极小值点，则 x^* 为 $\varphi(x)$ 的驻点.

即 x^* 满足 $\frac{\partial\varphi}{\partial x_i}=2\sum_{j=1}^n a_{ij}x_j-2b_i=0 \quad i=1,\cdots n$

$\Rightarrow \sum_{j=1}^n a_{ij}x_j^*-b_i=0$，即 $Ax^*=b$.

2° 若 x^* 为 $Ax=b$ 的解

$\forall y\in\mathbb{R}^n$. $\varphi(x^*+y)=(x^*+y)^TA(x^*+y)-2b^T(x^*+y)$

$=(x^*)^TAx^*+y^TAx^*+(x^*)^TAy+y^TAy-2b^Tx^*-2b^Ty$

$=\varphi(x^*)+2y^TAx^*+y^TAy-2y^Tb$

$=\varphi(x^*)+y^TAy\geq\varphi(x^*)$ $\quad\therefore$ x^* 是 φ 的极小值点.

法二：$\mathrm{grad}\,\varphi(x)=2Ax-2b=0$

$\left(\frac{\partial^2\varphi}{\partial x_i\partial x_j}\right)=2(a_{ij})=2A>0$

思路：给定 $x^{(0)}$，逐次求 $x^{(1)}$，$x^{(2)}$，… 使得 $\varphi(x^{(k+1)})<\varphi(x^{(k)})$

给定 $x^{(0)}$，确定方向 $p^{(0)}$，在 $x^{(0)}+\alpha p^{(0)}$ 上求 φ 的极小值点 $x^{(1)}$.

…

两个问题：1° 如何确定方向序列 $\{p^{(k)}\}$

2° 由 $x^{(k)}$ 及方向 $p^{(k)}$，如何求 $\varphi(x)$ 在 $x^{(k)}+\alpha p^{(k)}$ 上的极小值点 $x^{(k+1)}$

(二) 若有 $x\in\mathbb{R}^n$，给定向量 p，$p\neq0$，求方向 $x+\alpha p$ 上 $\varphi(x)$ 的极小值点 x^*.

令 $f(\alpha)=\varphi(x+\alpha p)=(x+\alpha p)^TA(x+\alpha p)-2b^T(x+\alpha p)=\varphi(x)+2\alpha(Ax-b)^Tp+\alpha^2p^TAp$.

令 $r=b-Ax$. 则 $f(\alpha)=\alpha^2p^TAp-2\alpha r^Tp+\varphi(x)$

驻点：$f'(\alpha)=2\alpha p^TAp-2r^Tp=0 \Rightarrow \alpha=\frac{r^Tp}{p^TAp}$ 为 $f(\alpha)$ 的极小值点.

$f''(\alpha)=2p^TAp>0$

即 $x^*=x+\frac{r^Tp}{p^TAp}p$ 　 $r^*=b-Ax^*$

注：$(r^*)^Tp=0$

总结：给定 $x^{(k)}$ 和方向 $p^{(k)}$, 则 $\varphi(x)$ 在 $x^{(k)}+\alpha p^{(k)}$ 上的极小值点为

$$x^{(k+1)}=x^{(k)}+\frac{(r^{(k)})^Tp^{(k)}}{(p^{(k)})^TAp^{(k)}}p^{(k)}$$

(三) 最速下降法

$\varphi(x)$ 增加最快的方向是梯度方向. $\text{grad}\,\varphi=2(Ax-b)=-2r$

取 $p^{(k)}=r^{(k)}$. 负梯度方向.

1° 取 $x^{(0)}\in\mathbb{R}^n$, $r^{(0)}=b-Ax^{(0)}$

2° 迭代. $\begin{cases}\alpha_k=\frac{\langle r^{(k)},r^{(k)}\rangle}{\langle r^{(k)},Ar^{(k)}\rangle}\\ x^{(k+1)}=x^{(k)}+\alpha_kr^{(k)}\\ r^{(k+1)}=b-Ax^{(k+1)}\end{cases}$ 　 $\langle r^{(k+1)},r^{(k)}\rangle=0$

注：① 此方法是收敛的 (定理5.1.2)

② 误差估计

③ 只是逐步最速，但整体不一定最好.

(四) 共轭梯度法

(1) 过程

给定初始向量 $x^{(0)}$, 选 $p^{(0)}=r^{(0)}=b-Ax^{(0)}$

有 $\alpha_0=\frac{\langle p^{(0)},r^{(0)}\rangle}{\langle p^{(0)},Ap^{(0)}\rangle}$, $x^{(1)}=x^{(0)}+\alpha_0p^{(0)}$, $r^{(1)}=b-Ax^{(1)}$.

下一步，在过点 $x^{(1)}$，由向量 $r^{(1)}$ 和 $p^{(0)}$ 所张成的平面

$\pi_2=\{x=x^{(1)}+\xi r^{(1)}+\eta p^{(0)}:\xi,\eta\in\mathbb{R}\}$ 上，找出使 φ 下降最快的方向作为新的方向 $p^{(1)}$.

令 $g(\xi,\eta)=\varphi(x^{(1)}+\xi r^{(1)}+\eta p^{(0)})$

$$=(x^{(1)}+\xi r^{(1)}+\eta p^{(0)})^TA(x^{(1)}+\xi r^{(1)}+\eta p^{(0)})-2b^T(x^{(1)}+\xi r^{(1)}+\eta p^{(0)})$$

直接计算：$\frac{\partial g}{\partial \xi}=2(\xi(r^{(1)})^TAr^{(1)}+\eta(r^{(1)})^TAp^{(0)}-(r^{(1)})^Tr^{(1)})$

$\frac{\partial g}{\partial \eta}=2(\xi(r^{(1)})^TAp^{(0)}+\eta(p^{(0)})^TAp^{(0)})$

令 $\frac{\partial g}{\partial \xi}=\frac{\partial g}{\partial \eta}=0$ 得 φ 在 π_2 内有唯一极小值点 (ξ_0,η_0)

即 $\frac{\partial g}{\partial \xi}\Big|_{(\xi_0,\eta_0)}=\frac{\partial g}{\partial \eta}\Big|_{(\xi_0,\eta_0)}=0$

令 $\beta_0=\frac{\eta_0}{\xi_0}=-\frac{\langle r^{(1)},Ap^{(0)}\rangle}{\langle p^{(0)},Ap^{(0)}\rangle}$ $\quad p^{(1)}=r^{(1)}+\beta_0p^{(0)}$.

$\alpha_1=\frac{\langle r^{(1)},p^{(1)}\rangle}{\langle p^{(1)},Ap^{(1)}\rangle}$ $\quad x^{(2)}=x^{(1)}+\alpha_1p^{(1)}$，$r^{(2)}=b-Ax^{(2)}$.

第k步：$\beta_k=-\frac{\langle r^{(k)},Ap^{(k-1)}\rangle}{\langle p^{(k-1)},Ap^{(k-1)}\rangle}$ $\quad p^{(k)}=r^{(k)}+\beta_{k-1}p^{(k-1)}$.

$\alpha_k=\frac{\langle r^{(k)},p^{(k)}\rangle}{\langle p^{(k)},Ap^{(k)}\rangle}$ $\quad x^{(k+1)}=x^{(k)}+\alpha_kp^{(k)}$，$r^{(k+1)}=b-Ax^{(k+1)}$.

(II)定理：由共轭梯度法得到的向量组 $\{r^{(i)}\}$ 和 $\{p^{(i)}\}$ 有如下性质：

① $\langle p^{(i)},r^{(j)}\rangle=0$，$0\le i<j\le k$.

② $\langle r^{(i)},r^{(j)}\rangle=0$，$0\le i,j\le k$ $\quad i\ne j$.

③ $\langle p^{(i)},Ap^{(j)}\rangle=0$，$0\le i,j\le k$ $\quad i\ne j$.

④ $\mathrm{span}\{r^{(0)},r^{(1)},\cdots,r^{(k)}\}=\mathrm{span}\{p^{(0)},p^{(1)},\cdots,p^{(k)}\}$

$=\mathrm{span}\{r^{(0)},Ar^{(0)},\cdots,A^kr^{(0)}\}\triangleq K(A,r^{(0)},k+1)\to$ Krylov子空间.

Pf：数学归纳法

$p^{(1)}=r^{(1)}+\beta_0p^{(0)}$

$\langle p^{(1)},Ap^{(0)}\rangle=\langle r^{(1)},Ap^{(0)}\rangle-\frac{\langle r^{(1)},Ap^{(0)}\rangle}{\langle p^{(0)},Ap^{(0)}\rangle}\langle p^{(0)},Ap^{(0)}\rangle=0$

$\langle r^{(1)},p^{(0)}\rangle=0$，$\langle r^{(2)},p^{(1)}\rangle=0$

$\langle r^{(2)}, p^{(0)}\rangle = \langle b-Ax^{(2)}, p^{(0)}\rangle = \langle b-Ax^{(1)}-\alpha_1 Ap^{(1)}, p^{(0)}\rangle$

$= \langle r^{(1)}, p^{(0)}\rangle - \alpha_1\langle Ap^{(1)}, p^{(0)}\rangle = 0$

$\langle r^{(1)}, p^{(0)}\rangle = \langle r^{(1)}, r^{(0)}\rangle = 0$, $\langle r^{(2)}, r^{(1)}\rangle = \langle r^{(2)}, p^{(1)}-\beta_0 p^{(0)}\rangle = \langle r^{(2)}, p^{(1)}\rangle - \beta_0\langle r^{(2)}, p^{(0)}\rangle = 0$.

定义：设A是对称正定阵. 若有$\mathbb{R}^n$上非零向量系$\{p^{(i)}\}_{i=0}^{m}$

满足$(p^{(i)}, Ap^{(j)}) = 0$, $\forall i\neq j$，则称$\{p^{(i)}\}_{i=0}^{m}$为A的共轭向量系.

性质：若$\{p^{(i)}\}_{i=0}^{m}$为A的共轭向量系，则

① $\{p^{(i)}\}_{i=0}^{m}$为$\mathbb{R}^n$上线性无关组 ② $m\le n-1$.

Pf：在$\mathbb{R}^n$上定义$[x,y]_A = x^T Ay$

满足 1° $[x,x]_A \ge 0$，当且仅当$x=0$时$[x,x]_A=0$

2° $[x,y]_A = [y,x]_A$

3° $\forall\lambda\in\mathbb{R}$. $[\lambda x,y]_A = [x,\lambda y]_A = \lambda[x,y]_A$

则$[x,y]_A$为x,y的内积.

即共轭向量系$\{p^{(i)}\}_{i=0}^{m}$为内积$[\cdot,\cdot]_A$下的正交系 $\Rightarrow$ ①②

注：① $\mathbb{R}^n$上的线性无关组在内积$[\cdot,\cdot]_A$下经Schmidt正交化可作出包含n个向量的共轭向量系.

② $\{r^{(i)}\}_{i=0}^{k}$和$\{p^{(i)}\}_{i=0}^{k}$分别是Krylov子空间$\mathcal{K}(A, r^{(0)}, k+1)$的正交基. 共轭正交基.

③ 由性质①，$\langle p^{(i)}, r^{(n)}\rangle = 0$, $0<i<n$.

最后一个向量的残量$r^{(n)}$正交于$p^{(0)}, p^{(1)}, \dots, p^{(n-1)}$

$\{p^{(i)}\}_{i=0}^{n-1}$为$\mathbb{R}^n$中的线性无关，构成$\mathbb{R}^n$中一组基

$\therefore \forall y\in\mathbb{R}^n$. $(y, r^{(n)}) = 0 \Rightarrow r^{(n)} = b-Ax^{(n)} = 0$ 即$x^{(n)}$为$Ax=b$的解.

(III) 运算过程

1° 给定初始向量 $x^{(0)}$，选 $p^{(0)}=r^{(0)}=b-Ax^{(0)}$

$$\alpha_0=\frac{\langle r^{(0)},p^{(0)}\rangle}{\langle p^{(0)},Ap^{(0)}\rangle},\quad x^{(1)}=x^{(0)}+\alpha_0 p^{(0)}.$$

2° 若已有 $x^{(0)},x^{(1)}\cdots x^{(k)}$，$p^{(0)},p^{(1)},\cdots p^{(k-1)}$，作 $p^{(k)}=r^{(k)}+\beta_{k-1}p^{(k-1)}$

选取 β_{k-1} s.t. $\langle p^{(k)},Ap^{(k-1)}\rangle=0\Rightarrow\beta_{k-1}=-\frac{\langle r^{(k)},Ap^{(k-1)}\rangle}{\langle p^{(k-1)},Ap^{(k-1)}\rangle}$

$$\alpha_k=\frac{\langle r^{(k)},p^{(k)}\rangle}{\langle p^{(k)},Ap^{(k)}\rangle},\ x^{(k+1)}=x^{(k)}+\alpha_k p^{(k)}.\quad r^{(k+1)}=b-Ax^{(k+1)}.$$

简化运算：

$$\alpha_k=\frac{\langle r^{(k)},p^{(k)}\rangle}{\langle p^{(k)},Ap^{(k)}\rangle}=\frac{\langle r^{(k)},r^{(k)}+\beta_{k-1}p^{(k-1)}\rangle}{\langle p^{(k)},Ap^{(k)}\rangle}=\frac{\langle r^{(k)},r^{(k)}\rangle}{\langle p^{(k)},Ap^{(k)}\rangle}$$

$$\beta_k=-\frac{\langle r^{(k+1)},Ap^{(k)}\rangle}{\langle p^{(k)},Ap^{(k)}\rangle}=\frac{\langle r^{(k+1)},r^{(k+1)}\rangle}{\langle r^{(k)},r^{(k)}\rangle}$$

算法 5.2.1

$x^{(0)}\in\mathbb{R}^n$. 计算 $p^{(0)}=r^{(0)}=b-Ax^{(0)}$，$\gamma_0=\langle r_0,r_0\rangle$，$S_0=\langle p^{(0)},Ap^{(0)}\rangle$.

对 $k=0,1,\cdots n-1$

① $\alpha_k=\frac{\gamma_k}{S_k}$　$x^{(k+1)}=x^{(k)}+\alpha_k p^{(k)}$

② $r^{(k+1)}=b-Ax^{(k+1)}$　$\gamma_{k+1}=\langle r^{k+1},r^{k+1}\rangle$

③ $\beta=\frac{\gamma_{k+1}}{\gamma_k}$，$p^{(k+1)}=r^{(k+1)}+\beta_k p^{(k)}$

④ $S_{k+1}=\langle p^{(k+1)},Ap^{(k+1)}\rangle$

注：① 理论上迭代至多 n 步就可得到 $Ax=b$ 的精确解.

② 实际上，终止迭代：$\|r^{(k)}\|$ 很小.及迭代次数达到最大.

③ 收敛性分析

定理：$A=I+B$. $\mathrm{rank}(B)=k$. 则共轭梯度法至多迭代 $k+1$ 步即可得到 $Ax=b$ 的解.

Pf. $\mathrm{span}\{r^{(0)},Ar^{(0)},\cdots,A^kr^{(0)}\}=\mathrm{span}\{r^{(0)},Br^{(0)},\cdots B^kr^{(0)}\}$.

④误差估计

$$\|x^{(k)}-x^*\|_A \le 2\left(\frac{\sqrt{K_2}-1}{\sqrt{K_2}+1}\right)^k \|x_0-x^*\|_A ,\quad K_2=\|A\|_2\|A^{-1}\|_2 .$$

Chap 6 矩阵特征值问题的计算方法

§6.1 矩阵特征值问题的相关知识

(一) 基本概念和性质

(I) $A\in\mathbb{R}^{n\times n}$, $\lambda\in\mathbb{C}$, s.t. $Ax=\lambda x$, $x\ne 0$, $x\in\mathbb{C}^n$

特征多项式 $P_A(\lambda)=\det(\lambda I-A)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_r)^{n_r}$

$n_1+n_2+\cdots+n_r=n$. $\lambda_1\cdots\lambda_r$ 互异. n_i 为 λ_i 的代数重数.

特征向量为 $(\lambda_iI-A)X=0$ 的解. 构成特征子空间 $\mathrm{Ker}(\lambda_iI-A)$. 其维数为 λ_i 的几何重数.

$m_i=n-\mathrm{rank}(\lambda_iI-A)$. 若 $n_i=m_i$. λ_i 为 A 的半单特征值. A的所有特征值为半单 $\Leftrightarrow$ A可对角化

A有n个线性无关的特征向量.

(II) 若存在非奇异矩阵 P s.t. $B=PAP^{-1}$, 则A与B相似.

相似的矩阵有相同的特征值. x 为A的一个特征向量 $\Longleftrightarrow$ $y=Px$ 为B的一个特征向量

(III) 三种常见的标准型:

1° Jordan 标准型

$A\in\mathbb{C}^{n\times n}$ 有 r 个互不相同的特征值 $\lambda_1\cdots\lambda_r$. 其重数分别为 $n(\lambda_1)\cdots n(\lambda_r)$

则必存在一个非奇异矩阵 $P\in\mathbb{C}^{n\times n}$. s.t.

$$P^{-1}AP=\begin{pmatrix} J(\lambda_1) & & \\ & \ddots & \\ & & J(\lambda_r)\end{pmatrix}\qquad J(\lambda_i)=\begin{pmatrix} J_1(\lambda_i) & & \\ & \ddots & \\ & & J_{k_i}(\lambda_i)\end{pmatrix}$$

$$J_j(\lambda_i)=\begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{pmatrix}\in\mathbb{C}^{n_j(\lambda_i)\times n_j(\lambda_i)}$$

其中 $n_1(\lambda_i)+\cdots+n_{k_i}(\lambda_i)=n(\lambda_i)$

2° Shur 分解定理：设 $A\in\mathbb{C}^{n\times n}$，则存在酉矩阵 $U\in\mathbb{C}^{n\times n}$ 使得 $U^HAU=T$

其中 T 为上三角阵，且可适当选取 U，使 T 的对角元按指定顺序排列。

3° 实数域上的 Shur 分解定理：$A\in\mathbb{R}^{n\times n}$，则存在正交阵 $Q\in\mathbb{R}^{n\times n}$，s.t. $Q^TAQ=\begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ & R_{22} & \cdots & R_{2n} \\ & & \ddots & \vdots \\ & & & R_{nn}\end{pmatrix}$ 拟上三角阵

其中 R_{ii} 是一个实数或者是一个有一对复共轭特征值的 2 阶方阵。

(二) 特征值的简单定位

最简单的结论：$|\lambda|\le\rho(A)\le\|A\|$

1° (Gerschgorin 第一圆盘定理)

矩阵 $A=(a_{ij})$ 的任意一个特征值至少落在复平面 n 个圆盘 $S_i=\{\lambda:|\lambda-a_{ii}|\le\sum\limits_{j\ne i}|a_{ij}|\}$ $i=1,2,\cdots n$ 中的某一个上。

(三) 特征值问题的敏度分析

Ex. $\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0\end{pmatrix}$ 特征值为 0 $\xrightarrow[\varepsilon>0]{\text{左下扰动}}$ $\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \varepsilon & & & 0\end{pmatrix}$ 特征值 $\varepsilon^{\frac{1}{n}}$.

定义：矩阵 A 的特征值整体条件数 $\nu(A)=\inf\limits_{Q\in D(A)}\|Q\|\|Q^{-1}\|$

$D(A)$ 包含所有使 A 对角化的相似变换矩阵。

定义：λ 为 A 的单根，x,y 分别为相应的单位右特征向量和单位左特征向量。$Ax=\lambda x$，$y^TA=\lambda y^T$

相应的特征值局部条件数 $w(\lambda;A)=\frac{1}{|y^*x|}$

定理：(λ,x) 为可对角化的矩阵 A 的近似特征信息，相应的残量 $r=Ax-\lambda x$.

则 $\exists$ 某个特征值 λ_i. s.t. $|\lambda-\lambda_i|\le\nu(A)\frac{\|r\|_2}{\|x\|_2}$

思路：① 全部特征值：对 A 作一系列变换 $P_k^{-1}AP_k$ 收敛于一个对角阵或一个(拟)上三角阵 → A 的所有特征值的近似值。

② 部分特征值：(针对性的方法)

③ 特征向量是 $Ax=\lambda x$ 的非零解，在求特征值过程中稍作处理可得特征向量。

§6.2 幂法

目的：计算一个矩阵的按模最大特征值对应的特征向量

(一) 思想：

$A\in\mathbb{R}^{n\times n}$. 可对角化，即有n个线性无关的特征向量，令其特征值为$\lambda_1\cdots\lambda_n$

不妨设$|\lambda_1|\geq|\lambda_2|\geq\cdots\geq|\lambda_n|$. 对应的特征向量$v_1,v_2,\cdots v_n$ $\mathbb{R}^n$中一组基

$\forall x^{(0)}\in\mathbb{R}^n$, $x^{(0)}$可表示为$x^{(0)}=\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n$

作迭代$x^{(k+1)}=Ax^{(k)}$ 得向量序列$\{x^{(k)}\}$.

$$x^{(k)}=A^kx^{(0)}=\sum_{j=1}^{n}\alpha_jA^kv_j=\sum_{j=1}^{n}\alpha_j\lambda^kv_j=\lambda_1^k[\alpha_1v_1+\sum_{j=2}^{n}(\frac{\lambda_j}{\lambda_1})^kv_j]$$

当$k\to\infty$时，$x^{(k)}\approx\lambda_1^k\alpha_1v_1$

(二) 分析：

A可对角化. $x^{(0)}=\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n$

$$x^{(k)}=\alpha_1\lambda_1^kv_1+\alpha_2\lambda_2^kv_2+\cdots+\alpha_n\lambda_n^kv_n.$$

(1) $|\lambda_1|>|\lambda_2|\geq\cdots\geq|\lambda_n|$

$$x^{(k)}=\lambda^k[\alpha_1v_1+\alpha_2(\frac{\lambda_2}{\lambda_1})^kv_2+\cdots+\alpha_n(\frac{\lambda_n}{\lambda_1})^kv_n]$$

若$\alpha_1\neq0$. 由于$|\frac{\lambda_i}{\lambda_1}|<1$ $i=2,\cdots n$

对充分大的k：$x^{(k)}\approx\lambda_1^k\alpha_1v_1$ $x^{(k+1)}\approx\lambda_1^{k+1}\alpha_1v_1$ $\Rightarrow\lambda_1=\frac{x^{(k+1)}}{x^{(k)}}$. $v_1=x^{(k)}$.

(2) $\lambda_1=-\lambda_2$.

$$x^{(k)}=\lambda^k[\alpha_1v_1+(-1)^k\alpha_2v_2+\cdots+\alpha_n(\frac{\lambda_n}{\lambda_1})^kv_n]$$

对充分大的k：$x^{(k)}\approx\lambda_1^k(\alpha_1v_1+(-1)^k\alpha_2v_2)$ $x^{(k+1)}\approx\lambda_1^{k+1}(\alpha_1v_1-(-1)^k\alpha_2v_2)$ $x^{(k+2)}\approx\lambda_1^{k+2}(\alpha_1v_1+(-1)^k\alpha_2v_2)$

$$\Rightarrow\lambda_1^2\approx\frac{x^{(k+2)}}{x^{(k)}}$$

$$\left.\begin{aligned}x^{(k+1)}+\lambda_1x^{(k)}&\approx2\lambda_1^{k+1}\alpha_1v_1\\x^{(k+1)}-\lambda_1x^{(k)}&\approx2(-1)^{k+1}\lambda_1^{k+1}\alpha_1v_1\end{aligned}\right.\Rightarrow\begin{aligned}v_1&\approx x^{(k+1)}+\lambda_1x^{(k)}\\v_2&\approx x^{(k+1)}+\lambda_2x^{(k)}\end{aligned}$$

(3) $\lambda_2 = \bar{\lambda}_1$

A是实矩阵 $\Rightarrow v_1 = \bar{v}_2$ $\quad Av_1 = \lambda_1 v_1$

$x^{(0)} \in \mathbb{R}^n$, $x^{(0)} = \alpha_1 v_1 + \bar{\alpha}_1 \bar{v}_1 + \alpha_3 v_3 + \cdots + \alpha_n v_n$

令 $\lambda_1 = \rho e^{i\theta}$. $\lambda_2 = \rho e^{-i\theta}$.

$$x^{(k)} = \lambda_1^k \alpha_1 v_1 + \bar{\lambda}_1^k \bar{\alpha}_1 v_1 + \lambda_3^k \alpha_3 v_3 + \cdots + \lambda_n^k \alpha_n v_n$$

$$= \rho^k [\alpha_1 e^{ik\theta} v_1 + \bar{\alpha}_1 e^{-ik\theta} \bar{v}_1 + \alpha_3 (\tfrac{\lambda_3}{\rho})^k v_3 + \cdots + \alpha_n (\tfrac{\lambda_n}{\rho})^k v_n]$$

对充分大的k有 $x^{(k)} \approx \rho^k (\alpha_1 e^{ik\theta} v_1 + \bar{\alpha}_1 e^{-ik\theta} \bar{v}_1)$ $\quad$ 记 $(\alpha_1 v_1)_j = r_j e^{i\theta_j}$

$$x_j^{(k)} \approx 2\rho^k r_j \cos(k\theta + \theta_j) \quad x_j^{(k+1)} \approx 2\rho^{k+1} r_j \cos((k+1)\theta + \theta_j) \quad x_j^{(k+2)} \approx 2\rho^{k+2} r_j \cos((k+2)\theta + \theta_j)$$

$$\lambda_1 + \lambda_2 = 2\rho\cos\theta \triangleq -p \quad \lambda_1 \cdot \lambda_2 = \rho^2 \triangleq q.$$

$$\Rightarrow x_j^{(k+2)} - (\lambda_1 + \lambda_2) x_j^{(k+1)} + \lambda_1 \lambda_2 x_j^{(k)} = 0. \quad j = 1, 2, \cdots n$$

有 $\begin{cases} x_1^{(k+2)} - (\lambda_1 + \lambda_2) x_1^{(k+1)} + \lambda_1 \lambda_2 x_1^{(k)} = 0 \\ x_2^{(k+2)} - (\lambda_1 + \lambda_2) x_2^{(k+1)} + \lambda_1 \lambda_2 x_2^{(k)} = 0 \end{cases}$ 可求出 p, q $\Rightarrow \lambda_1 = -\frac{p}{2} + i\sqrt{q - (\frac{p}{2})^2}$ $\quad \lambda_2 = -\frac{p}{2} - i\sqrt{q - (\frac{p}{2})^2}$

$\because x^{(k+1)} - \lambda_2 x^{(k)} = \lambda_1^k (\lambda_1 - \lambda_2) \alpha_1 v_1$ 取为 v_1

$x^{(k+1)} - \lambda_1 x^{(k)} = \lambda_2^k (\lambda_2 - \lambda_1) \alpha_2 v_2$ 取为 v_2

注: 幂法的收敛速度取决于 $|\frac{\lambda_2}{\lambda_1}|$ (第一种). $|\frac{\lambda_3}{\lambda_1}|$ (第二.三种).

缺点: 若 λ_1 很大, 则易上溢; λ_1 很小, 则易下溢.

(三) 幂法的规范性运算

对 $x^{(0)} \in \mathbb{R}^n$. $x^{(0)} \neq 0$. 作 $y^{(0)} = \frac{x^{(0)}}{\|x^{(0)}\|_\infty}$

迭代 $\begin{cases} x^{(k+1)} = Ay^{(k)} \\ y^{(k+1)} = \frac{x^{(k+1)}}{\|x^{(k+1)}\|_\infty} \end{cases}$

①若$|\lambda_1|>|\lambda_2|\geq\cdots\geq|\lambda_n|$

易看出$\{y^{(k)}\}$有极限，或$\{y^{(2k)}\}$与$\{y^{(2k+1)}\}$分别有极限.仅差一个负号.

当$\{y^{(k)}\}$有极限说明$\lambda_1>0$ 且 $\lambda_1\approx\max\limits_i|x_i^{(k)}|$ $v_1\approx x^{(k)}$

当$\{y^{(2k)}\}$与$\{y^{(2k+1)}\}$分别有极限 $v_1 . -v_1$ 说明$\lambda_1<0$. $\lambda_1\approx-\max\limits_i|x_i^{(k)}|$. $v_1=x^{(k)}$

运算时分别按$\|y^{(k)}-y^{(k-1)}\|_\infty<\varepsilon$, $\|y^{(k)}-y^{(k-2)}\|_\infty<\varepsilon$ 来控制

②若$\lambda_1=-\lambda_2$, $\lambda_1>0$. $|\lambda_1|=|\lambda_2|>|\lambda_3|\geq\cdots\geq|\lambda_n|$

$\{y^{(2k)}\}$ $\{y^{(2k+1)}\}$分别有两个不同的极限(不仅差一个负号)

若有k s.t. $\|y^{(k)}-y^{(k-2)}\|_\infty<\varepsilon$.

再做两次迭代 $\left.\begin{aligned}x^{(k+1)}&=Ay^{(k)}\\x^{(k+2)}&=Ax^{(k+1)}\end{aligned}\right\}$ $\lambda_1=\sqrt{x_i^{(k+2)}/y_i^{(k)}}\approx\sqrt{\max\limits_i|x_i^{(k+2)}|}$ $\lambda_2=-\lambda_1$

③$\{x^{(k)}\}$或$\{y^{(k)}\}$无规律. 尝试是否为$\lambda_1=\overline{\lambda_2}$.

作$\begin{cases}x^{(k+1)}=Ay^{(k)}\\x^{(k+2)}=Ax^{(k+1)}\\x^{(k+3)}=Ax^{(k+2)}\end{cases}$ 联立$\begin{cases}x_1^{(k+3)}+px_1^{(k+2)}+qx_1^{(k+1)}=0\\x_2^{(k+3)}+px_2^{(k+2)}+qx_2^{(k+1)}=0\end{cases}$ 求出$p.q$

计算$\|x^{(k+3)}+px^{(k+2)}+qx^{(k+1)}\|_\infty$ 若其值小于ε. 属于$\lambda_1=\overline{\lambda_2}$的情况. $\lambda_1=-\frac{p}{2}+i\sqrt{q-(\frac{p}{2})^2}$ $\lambda_2=-\frac{p}{2}-i\sqrt{q-(\frac{p}{2})^2}$

否则重复以上步骤. 直至$\|x^{(k+3)}+px^{(k+2)}+qx^{(k+1)}\|_\infty<\varepsilon$.

若反复运算始终达不到要求. 幂法失败.

(四)反幂法(按模最小特征值)

A非奇异. 若A的特征值$|\lambda_1|\geq|\lambda_2|\geq\cdots\geq|\lambda_n|$特征向量为$v_1, v_2\cdots v_n$

则A^{-1}的特征值为$\lambda_1^{-1}, \lambda_2^{-1}\cdots\lambda_n^{-1}$特征向量为$v_1, v_2\cdots v_n$

因而只需求A^{-1}的按模最大特征值μ及特征向量v.

反幂法：

取 $x^{(0)}\neq 0$．作 $y^{(0)}=x^{(0)}/\|x^{(0)}\|_\infty$

迭代 $\begin{cases} x^{(k+1)}=A^{-1}y^{(k)} \\ y^{(k+1)}=x^{(k+1)}/\|x^{(k+1)}\|_\infty \end{cases}$ 注：为避免求逆，可以解方程.

(三) 移位幂法 (对近似特征值的精确化)

求与值 p 最近的特征值和特征向量.

1° 求 $A-pI$ 的按模最小特征值 μ 及特征向量 v (反幂法)

2° 则 $\mu+p$ 为所求特征值．特征向量 v

§6.3 QR方法

思想：利用相似的矩阵有相同的特征值，利用正交变换对A作一系列正交相似变换 → 上三角阵/拟上三角阵 → 求出全部特征值

(一) 基本迭代与收敛性

令 $A_0=A$，作分解 $A_0=Q_1R_1$，令 $A_1=R_1Q_1$ $A_1=Q_1^TQ_1R_1Q_1=Q_1^TA_0Q_1$．$A_0$ 与 A_1 正交相似.

作分解 $A_1=Q_2R_2$，令 $A_2=R_2Q_2$

⋮

$A_m=Q_{m+1}R_{m+1}$，令 $A_{m+1}=R_{m+1}Q_{m+1}=Q_{m+1}^TA_mQ_{m+1}$.

得到矩阵序列 $\{A_m\}$．且序列中的矩阵是相似的．(正交相似)

$$A_{m+1}=Q_{m+1}^TA_mQ_{m+1}=Q_{m+1}^TQ_m^T\cdots Q_1^TA_0Q_1\cdots Q_mQ_{m+1}$$

令 $\tilde{Q}_m=Q_1Q_2\cdots Q_m$ $\tilde{R}_m=R_m\cdots R_1$

$$A_{m+1}=\tilde{Q}_{m+1}^TA_0\tilde{Q}_{m+1}\Rightarrow R_{m+1}Q_{m+1}=\tilde{Q}_{m+1}^TA_0\tilde{Q}_{m+1}$$

$$\Rightarrow \tilde{Q}_{m+1}R_{m+1}=A_0\underbrace{\tilde{Q}_{m+1}Q_{m+1}^T}_{\tilde{Q}_m}$$

即 $\tilde{Q}_{m+1}R_{m+1}=A_0\tilde{Q}_m \Rightarrow \tilde{Q}_{m+1}\tilde{R}_{m+1}=A_0\tilde{Q}_m\tilde{R}_m=A_0^mQ_1R_1=A_0^{m+1}$

令 $\tilde{R}_m=(r_{ij})$ $\tilde{Q}_m$ 的第一列为 $q_1^{(m)} \Rightarrow A^m e_1 = r_{11} q_1^{(m)}$

即 $q_1^{(m)}$ 可看作A用 e_1 作为初始向量进行幂法所得到的向量.

若A的特征值 $\lambda_1 \cdots \lambda_n$ 满足 $|\lambda_1|>|\lambda_2|\geq|\lambda_3|\geq\cdots\geq|\lambda_n|$，则 $q_1^{(m)}$ 将收敛到A的一个属于 λ_1 的特征向量.

定理：设A的n个特征值满足 $|\lambda_1|>|\lambda_2|>\cdots>|\lambda_n|>0$. 若Y为A的n个线性无关特征向量构成的矩阵，即 $A=Y^{-1}\Lambda Y$.

若Y有LU分解. 则由QR方法迭代格式产生的矩阵序列 $\{A_m\}$ 是基本收敛的.

实矩阵在正交相似变换下的标准形问题：

定理（实Shur标准形）

$A\in\mathbb{R}^{n\times n}$，则存在正交阵 $Q\in\mathbb{R}^{n\times n}$，s.t. $Q^TAQ=\begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ & R_{22} & \cdots & R_{2m} \\ & & \ddots & \vdots \\ & & & R_{mm} \end{pmatrix}$

其中 R_{ii} 或是一个实数，或是一个具有一对复共轭特征值的2阶方阵.

注：①基本迭代有以下缺点 a) 每一次迭代运算量大　b) 收敛速度慢.

②实际计算中（为了减少每次迭代的计算量）

$A \xrightarrow{\text{相似变换}}$ 准上三角阵 $\xrightarrow{\text{QR迭代}}$ 求出全部特征值.

(二) 上Hessenberg化

(I) 定义：$H=(h_{ij})$ 满足 $h_{ij}=0$ $i>j+1$. 即 $H=\begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ & h_{32} & \cdots & h_{3n} \\ & & \ddots & \vdots \\ & & h_{n,n-1} & h_{nn} \end{pmatrix}$ 称形如H的矩阵为Hessenberg矩阵.

若 $h_{i+1,i}\neq 0$, $i=1,\cdots n-1$ 则H为不可约的.

注：若 $h_{k+1,k}=0$. 可把H分块化 $H=\begin{pmatrix} H_1 & \\ & H_2 \end{pmatrix}$

(II) 定理：$A\in\mathbb{R}^{n\times n}$，则恒有与A正交相似的Hessenberg矩阵H.

Pf：记 $A^{(1)}=A$，并写成分块形式 $A^{(1)}=\begin{pmatrix} a_{11} & a_2^T \\ a_1 & A_{22} \end{pmatrix}$ 其中 $a_1^T=(a_{21},a_{31},\cdots a_{n1})$，$a_2^T=(a_{12},a_{13},\cdots a_{1n})$

选取 $H_1=\begin{pmatrix} 1 & 0 \\ 0 & \tilde{H}_1 \end{pmatrix}$，其中 $\tilde{H}_1$ 满足 $\tilde{H}_1 a_1=\alpha e_1$

$H_1AH_1=\begin{pmatrix} a_{11} & a_2^T \\ \alpha e_1 & \tilde{H}_1A_{22} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \tilde{H}_1 \end{pmatrix}=\begin{pmatrix} a_{11} & a_2^T\tilde{H}_1 \\ \alpha e_1 & \tilde{H}_1A_{22}\tilde{H}_1 \end{pmatrix}\triangleq A^{(2)}$

对 $\tilde{A}_{22}=\tilde{H}_1 A_{22}\tilde{H}_1$ 进行同样的变换.

$$H_2=\begin{pmatrix}1 & 0\\ 0 & \tilde{H}_2\end{pmatrix}\quad s.t.\ \tilde{H}_2\tilde{A}_{22}\tilde{H}_2=\begin{pmatrix}* & * & \cdots & *\\ * & * & \cdots & *\\ 0 & * & & \vdots\\ \vdots & \vdots & & \\ 0 & * & \cdots & *\end{pmatrix}$$

则 $H_2H_1AH_1H_2=\begin{pmatrix}* & * & * & \cdots & *\\ * & * & * & \cdots & *\\ 0 & * & * & \cdots & *\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & * & \cdots & *\end{pmatrix}$

如此进行 $n-2$ 步. 找到 $n-2$ 个 Householder 变换 $H_1\cdots H_{n-2}$

使得 $H_{n-2}H_{n-3}\cdots H_2H_1AH_1H_2\cdots H_{n-3}H_{n-2}=H$ 为 Hessenberg 矩阵.

$H=Q^TAQ$ 即 $A=QHQ^T$ 为 A 的上 Hessenberg 分解.

算法 6.4.1 ①运算量 $O(n^3)$

②舍入误差: 上述算法得到 Hessenberg 矩阵 $\hat{H}$ 满足 $\hat{H}=Q^T(A+E)Q$, 其中 Q 为正交矩阵.

则 $\|E\|_F\le cn^2\|A\|_F u$.

③ A 稀疏, 可通过 Givens 变换将 A 正交相似化为 Hessenberg.

(四)"唯一"性 (上 Hessenberg 分解是不唯一的)

定理: $A\in\mathbb{R}^{n\times n}$, 有如下两个上 Hessenberg 分解. $U^TAU=H$, $V^TAV=G$.

其中 $U=[u_1,u_2,\cdots u_n]$ 和 $V=[v_1,v_2\cdots v_n]$ 是 n 阶正交矩阵. $H=(h_{ij})$, $G=(g_{ij})$ 是上 Hessenberg 矩阵.

若 $u_1=v_1$ 且 H 不可约, 则存在对角元均为 1 或 -1 的对角阵 D. s.t. $U=VD$, $H=DGD$.

即: A 的上 Hessenberg 分解 $Q^TAQ=H$ 不可约, Q 和 H 完全由 Q 的第一列来确定.

在相差一个正负号的意义下唯一.

(三) H 的 QR 方法.

(1) H 的 QR 分解 (由于 H 的特殊性, 可用 $n-1$ 个 Givens 变换来完成)

$$H^{(1)}=H=\begin{pmatrix}h_{11}^{(1)} & h_{12}^{(1)} & \cdots\cdots & h_{1n}^{(1)}\\ h_{21}^{(1)} & h_{22}^{(1)} & \cdots\cdots & h_{2n}^{(1)}\\ & \ddots & \ddots & \\ & & h_{n,n-1}^{(1)} & h_{nn}^{(1)}\end{pmatrix}\quad 作 Givens 矩阵\ G(1,2,\theta_1)=\begin{pmatrix}c_1 & s_1 & \\ -s_1 & c_1 & \\ & & I_{n-2}\end{pmatrix}$$

记 $H_2^{(2)}=G(1,2,\theta_1)H^{(1)}$. 若 θ_1 满足 $\tan\theta_1=\frac{h_{21}^{(1)}}{h_{11}^{(1)}}\Rightarrow h_{21}^{(2)}=0$

经过 $n-1$ 次 Givens 变换后, $R=H^{(n)}$ 为上三角阵. $R=H^{(n)}=Q^TH$. $\Rightarrow H=QR$.

注：相对于对A用Householder矩阵进行QR分解，用Givens矩阵对与A正交相似的Hessenberg矩阵H进行QR分解

其运算大为减少

(Ⅲ) QR迭代

$H=QR$ 则 $\tilde{H}=RQ$ 仍为上Hessenberg矩阵.

步骤总结：

1° 对A作上Hessenberg分解 $Q^TAQ=H$

2° 令 $H_0=H$. QR迭代构造 $\{H_k\}$. $H_k=Q_kR_k$ $H_{k+1}=R_kQ_k$.

(四) 实用的QR方法 (对H)

(Ⅰ) 带原点位移的QR方法

$H_k-\mu_kI=Q_kR_k$, $H_{k+1}=R_kQ_k+\mu_kI$. $H_{k+1}=Q_k^TQ_kR_kQ_k+\mu_kQ_k^TQ_k$

$=Q_k^T(Q_kR_k+\mu_kI)Q_k=Q_k^TH_kQ_k$ 正交相似.

适当选取 μ_k, 加快 $\{H_k\}$ 的收敛速度.

取 $\mu_k=H_k(n,n)$ (线性收敛 ⟶ 二次收敛)

$QR=H-h_{nn}I$

$$H-h_{nn}I=\begin{pmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ & \ddots & \ddots & \\ & & a & b \\ & & \varepsilon & 0 \end{pmatrix}$$

$\tilde{H}=RQ+h_{nn}I$

(Ⅱ) 双重步位移的QR方法 (A有复共轭特征值，实位移一般不能起到加速作用)

思想：将两步带原点位移的QR迭代合为一步

1° $H=Q_0^TAQ_0$ (上Hessenberg分解)

2° $H-\mu_1I=Q_1R_1$ $H_1=R_1Q_1+\mu_1I$ $H_1=Q_1^THQ_1$

$H_1-\mu_2I=Q_2R_2$ $H_2=R_2Q_2+\mu_2I$ $H_2=Q_2^TH_1Q_2$

$M=(H-\mu_1 I)(H-\mu_2 I)=H^2-sH+tI$ 其中 $s=\mu_1+\mu_2$ $t=\mu_1\cdot\mu_2$

$M=Q_1R_1(H-\mu_2 I)=Q_1R_1(Q_1R_1+\mu_1 I-\mu_2 I)$

$=Q_1R_1Q_1R_1+Q_1(\mu_1-\mu_2)IR_1$

$=Q_1(H_1-\mu_1 I)R_1+Q_1(\mu_1-\mu_2 I)R_1$

$=Q_1(H_1-\mu_2 I)R_1=\underbrace{Q_1Q_2}_{Q}\underbrace{R_2R_1}_{R}$

$H_2=Q_2^T H_1 Q_2=Q_2^T Q_1^T H Q_1 Q_2=Q^T H Q.$

注：① M是实矩阵. $M=QR$

② $H_2=Q^THQ$ 仍是上 Hessenberg 矩阵.

步骤：① $M=H^2-sH+tI$

② $M=QR$

③ $H_2=Q^THQ$.

(Ⅲ) 隐式 QR 方法 目标：$H\to H_2$ $O(n^2)$

若H不可约

1° $Me_1=(\xi_1,\xi_2,\xi_3,0,0\cdots0)^T$. 其中 $\xi_1=h_{11}^2+h_{12}\cdot h_{21}-sh_{11}+t$

$\xi_2=h_{21}h_{11}+h_{22}h_{21}-sh_{21}$ $\xi_3=h_{32}h_{21}$

2° 若有Householder变换 P_0 将 $Me_1\longrightarrow\alpha e_1$

则 $P_0Me_1=\alpha e_1\Rightarrow Me_1=\alpha P_0e_1\Rightarrow$ M的第一列与 P_0 的第一列共线.

另：$M=QR$. $Me_1=r_{11}q_1$ M的第一列与Q的第一列共线.

可把 P_0 的第一列作为Q的第一列

$P_0=\begin{pmatrix}\tilde{P}_0 & \\ & I_{n-3}\end{pmatrix}$ $\tilde{P}_0=I_3-\beta vv^T$. $v=\begin{pmatrix}\xi_1-\alpha\\ \xi_2\\ \xi_3\end{pmatrix}$ $\alpha=\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$ $\beta=\frac{2}{v^Tv}$.

令 $B=P_0HP_0$

3° 约化 B 为 Hessenberg 矩阵.

$$P_1 = \begin{pmatrix} 1 & & \\ & \widetilde{P}_1 & \\ & & I_{n-4} \end{pmatrix} \xrightarrow{P_1 B P_1} P_2 = \begin{pmatrix} I_2 & & \\ & \widetilde{P}_2 & \\ & & I_{n-5} \end{pmatrix} \cdots\cdots P_k = \begin{pmatrix} I_k & & \\ & \widetilde{P}_k & \\ & & I_{n-k-3} \end{pmatrix} \cdots P_{n-2} = \begin{pmatrix} I_{n-2} & \\ & \underbrace{\widetilde{P}_{n-2}}_{2\times 2} \end{pmatrix}$$

$\widetilde{H}_2 = P_{n-2} \cdots P_1 P_0 H P_0 P_1 \cdots P_{n-2}$

令 $P = P_0 P_1 \cdots P_{n-2}$　$\widetilde{H}_2 = P^T H P$　$H_2 = Q^T H Q$

$P_k e_1 = e_1$　$k = 1, 2, \cdots n-2$　$Pe_1 = P_0 e_1$

又由于 P_0 与 Q 有相同的第一列. 由定理: $\widetilde{H}_2 = H_2$

算法 6.4.2 (双重步位移的 QR 迭代)

运算量 $10n^2$.

隐式 QR 算法 (The overall process)

1° A. 用算法 6.4.1 计算 A 的上 Hessenberg 分解 $H = U_0^T A U_0$　$Q = U_0$

2° 收敛性判定.

2.1 If $|h_{i,i-1}| \le (|h_{ii}| + |h_{i-1,i-1}|) \cdot tol$　$i = 2, \cdots n$

then $h_{i,i-1} = 0$

2.2 确定最大的非负整数 m 和最小的非负整数 l. 使得

$$H = \begin{array}{c} \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ 0 & H_{22} & H_{23} \\ 0 & 0 & H_{33} \end{bmatrix} \end{array} \begin{array}{l} l \\ n-l-m \\ m \end{array}$$

其中 H_{33} 为拟上三角阵, H_{22} 为不可约 Hessenberg 矩阵.

若 $m = n$, 则输出相关信息, 结束.

否则: i) 对 H_{22} 用算法 6.4.2 迭代一次. $H_{22} = P^T H_{22} P$.　$P = P_0 \cdots P_{n-m-l-2}$

ii) 计算 $Q = Q \begin{pmatrix} I_l & & \\ & P & \\ & & I_m \end{pmatrix}$　$H_{12} = H_{12} P$　$H_{23} = P^T H_{23}$

转为 2°

Chap 7 对称特征值问题的计算方法.

回顾：$A\in\mathbb{R}^{n\times n}$, $A=A^T$. 有以下性质

1° A的特征值都是实数.

2° 存在正交阵Q. s.t. $Q^TAQ=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$

$\lambda_1\cdots\lambda_n$ 为A的特征值，Q的第j列恰为λ_j的特征向量

§7.1 QR方法

(一) 三对角化

$A\in\mathbb{R}^{n\times n}$, $A=A^T$, A的上Hessenberg分解 $Q^TAQ=T=\begin{pmatrix}\alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_n\end{pmatrix}$

$\because T=T^T \Rightarrow T$ 必为三对角阵.

过程：(归纳法)

$$A=\begin{array}{c} \begin{pmatrix}\alpha_1 & v_0^T \\ v_0 & A_0\end{pmatrix}\begin{array}{c}1\\ n-1\end{array} \\ \begin{array}{cc}1 & n-1\end{array}\end{array} \xrightarrow[H_1=\begin{pmatrix}1 & \\ & \tilde{H}_1\end{pmatrix}]{H_1AH_1} \begin{pmatrix}\alpha_1 & v_0^T\tilde{H}_1 \\ \tilde{H}_1v_0 & \tilde{H}_1A_0\tilde{H}_1\end{pmatrix}$$

利用Householder变换经过k-1次. 即 $A^{(k-1)}=(H_1\cdots H_{k-1})^TAH_1\cdots H_{k-1}$ (前$(k-1)\times(k-1)$三对角)

$$=\begin{pmatrix}B_{11} & B_{12} & 0 \\ B_{21} & B_{22} & B_{23} \\ 0 & B_{32} & B_{33}\end{pmatrix} \quad B_{21}=(0\cdots 0,*)^T$$

构造 $\tilde{H}_k\in\mathbb{R}^{(n-k)\times(n-k)}$ $\tilde{H}_kB_{32}=\alpha e_1$ $H_k=\begin{pmatrix}I_k & \\ & \tilde{H}_k\end{pmatrix}$

$(H_1\cdots H_{n-2})^TAH_1H_2\cdots H_{n-2}=T$

从 $A^{(k-1)}\rightarrow A^{(k)}$，第k步的工作量. 计算 $\tilde{H}_kB_{33}\tilde{H}_k$

利用对称性：$\tilde{H}_k=I-\beta vv^T$ $\beta=\frac{2}{v^Tv}$

令 $u=\beta B_{33}v$ $w=u-\frac{1}{2}\beta(v^Tu)v$

$\tilde{H}_kB_{33}\tilde{H}_k=B_{33}-vw^T-wv^T$

$A^{(k-1)}\rightarrow A^{(k)}$ 计算量 $4(n-k)^2$.

算法7.2.1（利用Householder变换计算三对角分解）

注：①算法计算量为 $\frac{4n^3}{3}$ ② $\hat{T}=\tilde{Q}^T(A+E)\tilde{Q}$，$\|E\|_F\le c\|A\|_F u$.

(二)对T作QR方法

带原点位移的QR迭代

$$T-\mu I=QR \quad \tilde{T}=RQ+\mu I \quad \tilde{T}=Q^TTQ.$$

由于QR迭代保持上Hessenberg形+对称性 $\Rightarrow$ $\tilde{T}$ 仍是三对角的对称阵.

①最简单的做法 $\mu=T_{nn}$ $\quad O(\varepsilon^2)$

②更好的做法：$T(n-1:n,\ n-1:n)=\begin{pmatrix}\alpha_{n-1} & \beta_{n-1}\\ \beta_{n-1} & \alpha_n\end{pmatrix}$

取 μ 为 $T(n-1:n,n-1:n)$ 的两个特征值中最靠近 α_n 的一个.

$$\mu=\alpha_n+\delta-\mathrm{sgn}(\delta)\sqrt{\delta^2+\beta_{n-1}^2} \qquad \delta=\frac{\alpha_{n-1}-\alpha_n}{2}$$

↓

Wilkinson位移. $O(\varepsilon^3)$

③更漂亮的做法：隐式对称QR算法

隐含的方式实现 $T_k\to T_{k+1}$

考虑一次迭代：$T-\mu I=QR$ $\quad \tilde{T}=RQ+\mu I$ $\quad \tilde{T}=Q^TTQ$ $\quad T\to\tilde{T}$

1°用Givens变换实现 $T-\mu I$ 下的QR分解.

$$\begin{pmatrix}c & s\\ -s & c\end{pmatrix}\begin{pmatrix}\alpha_1-\mu\\ \beta\end{pmatrix}=\begin{pmatrix}*\\ 0\end{pmatrix}$$

$G_1=G(1,2,\theta_1)$ $\quad G_1^Te_1=Qe_1$ 让 G_1^T 的第一列作为Q的第一列.

$$2^\circ\ B=G_1TG_1^T=\begin{pmatrix}c & s & & \\ -s & c & & \\ & & 1 & \\ & & & 1\end{pmatrix}\begin{pmatrix}\times & \times & & \\ \times & \times & \times & \\ & \times & \times & \times\\ & & \times & \times\end{pmatrix}\begin{pmatrix}c & s & & \\ -s & c & & \\ & & 1 & \\ & & & 1\end{pmatrix}=\begin{pmatrix}\times & \times & + & \\ \times & \times & \times & \\ + & \times & \times & \times\\ & & \times & \times\end{pmatrix}$$

3°将B用Givens变换化为三对角.

$$B\xrightarrow[\begin{pmatrix}1 & & & \\ & c & s & \\ & -s & c & \\ & & & 1\end{pmatrix}]{G_2BG_2^T}\begin{pmatrix}\times & \times & & \\ \times & \times & \times & +\\ & \times & \times & \times\\ & + & \times & \times\end{pmatrix}\xrightarrow[\begin{pmatrix}1 & & & \\ & 1 & & \\ & & c & s\\ & & -s & c\end{pmatrix}]{G_3G_2BG_2^TG_3^T}\begin{pmatrix}\times & \times & & \\ \times & \times & \times & \\ & \times & \times & \times\\ & & \times & \times\end{pmatrix}\triangleq\tilde{\tilde{T}}=G_3G_2G_1TG_1^TG_2^TG_3^T=G^TTG$$

$$G_2^Te_1=e_1\quad G_3^Te_1=e_1\ \Rightarrow Ge_1=G_1^Te_1=Qe_1\ \Rightarrow \tilde{\tilde{T}}=\tilde{T}.$$

隐式对称QR算法

1° A. 用算法7.2.1计算A的三对角分解 $T=U_0^TAU_0$ $Q=U_0$

2° 收敛性判定.(until $m=n$)

2.1 If $|t_{i+1,i}|\le(|t_{ii}|+|t_{i+1,i+1}|)\cdot tol$ $i=2,\dots n$

then $t_{i,i+1}=t_{i+1,i}=0$

2.2 确定最大的非负整数m和最小的非负整数l. 使得

$$H=\begin{array}{c}\left[\begin{array}{ccc}T_{11}&0&0\\0&T_{22}&0\\0&0&T_{33}\end{array}\right]\end{array}\begin{array}{l}l\\n-l-m\\m\end{array}$$

其中T_{33}为对角阵, T_{22}为不可约三对角阵

If$(m<n)$ 用算法7.2.2对T_{22}迭代一次

$$T_{22}=GT_{22}G^T\quad G=G_1^T\cdots G_{n-m-l-1}^T\quad Q=Q\begin{pmatrix}I_l&&\\&G&\\&&I_m\end{pmatrix}$$

end.

end.

§7.2 Jacobi方法 $A=A^T$ $A\xrightarrow[\text{相似于}]{\text{正交}}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}$

基本思想:对A作一系列正交相似变换,使其逐次"更对角化"

通过平面旋转变换,逐次加重对角线元素的比重,使非对角元素逐次约化为0.

(一)平面旋转变换(Givens变换, Jacobi rotations)

$A=(a_{ij})$ $A=A^T$.

记平面旋转变换 $J(p,q,\theta)=\begin{array}{c}\begin{array}{ccccc}&p&&q&\end{array}\\\left(\begin{array}{ccccc}1&&&&\\&\cos\theta&\cdots&\sin\theta&\\&\vdots&&\vdots&\\&-\sin\theta&\cdots&\cos\theta&\\&&&&1\end{array}\right)\end{array}\begin{array}{l}\\\\p\\\\q\\\\\end{array}$ 为正交阵.

$B=J^TAJ\triangleq(b_{ij})$

$$\left.\begin{array}{ll} b_{ip}=b_{pi}=a_{pi}\cos\theta-a_{qi}\sin\theta,\ i\neq p,q & b_{iq}=b_{qi}=a_{pi}\sin\theta+a_{qi}\cos\theta,\ i\neq p,q \\ b_{pp}=a_{pp}\cos^2\theta+a_{qq}\sin^2\theta-a_{pq}\sin 2\theta & b_{qq}=a_{pp}\sin^2\theta+a_{qq}\cos^2\theta+a_{pq}\sin 2\theta. \\ b_{pq}=b_{qp}=a_{pq}\cos 2\theta+\frac{a_{pp}-a_{qq}}{2}\sin 2\theta & b_{ij}=a_{ij},\ i\neq p,q,\ j\neq p,q. \end{array}\right\}\circledast$$

若$a_{pq}\neq 0$，可选取θ，使得$b_{pq}=0$，即 $\cot 2\theta=\frac{a_{qq}-a_{pp}}{2a_{pq}}$ $|2\theta|<\frac{\pi}{2}$.

注：$a_{pq}=0$，取$\theta=0$即可.

令 $\tau=\frac{a_{qq}-a_{pp}}{2a_{pq}}$，$t=\tan\theta$.

由 $\tan^2\theta+2\tan\theta\cot 2\theta-1=0 \Rightarrow t^2+2\tau t-1=0$.

取 $t=\tan\theta=\begin{cases} t^2+2\tau t-1=0 \text{ 的按模最小根}, & \tau\neq 0 \\ 1, & \tau=0 \end{cases}$

即 $t=\begin{cases} \frac{1}{\tau+\sqrt{1+\tau^2}}, & \tau\geq 0 \\ \frac{1}{\tau-\sqrt{1+\tau^2}}, & \tau<0 \end{cases}$

记 $c=\cos\theta$，$s=\sin\theta$ 有 $c=\frac{1}{\sqrt{1+t^2}}$，$s=\frac{t}{\sqrt{1+t^2}}$

此时$\circledast$可写为 $\begin{cases} b_{ip}=b_{pi}=ca_{pi}-sa_{qi},\ i\neq p,q \\ b_{iq}=b_{qi}=sa_{pi}+ca_{qi},\ i\neq p,q \\ b_{pp}=a_{pp}-ta_{qq} \quad b_{qq}=a_{pp}+ta_{qq} \\ b_{pq}=b_{qp}=0. \quad b_{ij}=a_{ij} \quad i\neq p,q \quad j\neq p,q. \end{cases}$

记 $\sigma_1(A)=\sum_{\substack{i,j=1\\i\neq j}}^{n}a_{ij}^2$ $\sigma_2(A)=\sum_{i=1}^{n}a_{ii}^2$，$\|A\|_F^2=\sigma_1(A)+\sigma_2(A)$

$\because B=J^TAJ \quad \therefore \|A\|_F=\|B\|_F$

$$b_{pp}^2+b_{qq}^2=(a_{pp}-ta_{pq})^2+(a_{qq}+ta_{pq})^2$$

$$=a_{pp}^2+a_{qq}^2+2a_{pq}^2 \Rightarrow \sigma_2(B)=\sigma_2(A)+2a_{pq}^2. \Rightarrow \sigma_1(B)=\sigma_1(A)-2a_{pq}^2.$$

(二) Jacobi方法基本迭代格式(构造正交相似矩阵序列$\{A^{(k)}\}$.

1° $A^{(0)}=A$

2° 已有$A^{(k)}=(a_{ij}^{(k)})$

2.1 求下标(p,q)，选$a_{pq}^{(k)}=\max\limits_{1\le i<j\le n}|a_{ij}|$

2.2 记$\tau^{(k)}=(a_{qq}^{(k)}-a_{pp}^{(k)})/2a_{pq}^{(k)}$

$$t^{(k)}=\begin{cases}\dfrac{1}{\tau^{(k)}+\sqrt{1+(\tau^{(k)})^2}}, & \tau^{(k)}\ge 0\\ \dfrac{1}{\tau^{(k)}-\sqrt{1+(\tau^{(k)})^2}}, & \tau^{(k)}<0\end{cases}\qquad c^{(k)}=\frac{1}{\sqrt{1+(t^{(k)})^2}},\quad s^{(k)}=t^{(k)}c^{(k)}.$$

2.3 作$A^{(k+1)}=J^T(p,q,\theta)A^kJ(p,q,\theta)$.

其中θ满足$\cos\theta=c^{(k)}$，$\sin\theta=s^{(k)}$.

$A^{(k+1)}$的元素为 $a_{ip}^{(k+1)}=a_{pi}^{(k+1)}=c^{(k)}a_{pi}^{(k)}-s^{(k)}a_{qi}^{(k)}$，$i\ne p,q$.

$$a_{iq}^{(k+1)}=a_{qi}^{(k+1)}=c^{(k)}a_{qi}^{(k)}+s^{(k)}a_{pi}^{(k)},\quad i\ne p,q.$$

$$a_{pp}^{(k+1)}=a_{pp}^{(k)}-t^{(k)}a_{pq}^{(k)}\qquad a_{qq}^{(k+1)}=a_{qq}^{(k)}+t^{(k)}a_{pq}^{(k)}.$$

$$a_{pq}^{(k+1)}=a_{qp}^{(k+1)}=0\qquad a_{ij}^{(k+1)}=a_{ij}^{(k)}\quad i\ne p,q,\ j\ne p,q.$$

以上Jacobi方法可得到正交相似的矩阵序列$\{A^{(k)}\}$.

且$\sigma_1(A^{(k+1)})=\sigma_1(A^{(k)})-2(a_{pq}^{(k)})^2$，$\sigma_2(A^{(k+1)})=\sigma_2(A^{(k)})+2(a_{pq}^{(k)})^2$

(三) 收敛性

定理：A为对称阵，则由$A^{(0)}=A$出发，利用Jacobi方法得到的正交相似矩阵

序列$\{A^{(k)}\}$有极限$\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots\lambda_n)$，$\lambda_1\cdots\lambda_n$是特征值.

(四) 算法

(1) $J(p,q,\theta)$

1° 确定p,q，$|a_{pq}|=\max\limits_{1\le i<j\le n}|a_{ij}|$

2° 求 $(c=\cos\theta,\ s=\sin\theta)$

function $[c, s, t] =$ symSchur2 (A, p, q)

If $A(p,q) \neq 0$

$\tau = (A(q,q) - A(p,p))/2A(p,q)$

If $\tau \geq 0$

$t = \dfrac{1}{\tau + \sqrt{1+\tau^2}}$

else

$t = \dfrac{1}{\tau - \sqrt{1+\tau^2}}$

end

$c = \dfrac{1}{\sqrt{1+t^2}}$, $s = ct$.

else $c=1$, $s=0$

end

(II) 循环 Jacobi 算法

$V = I_n$, $\delta = tol \cdot \|A\|_F$

while $\sigma_1(A) > \delta$

for $p = 1 : n-1$

for $q = p+1 : n$

$[c, s, t] =$ symSchur2 (A, p, q)

$A = J(p,q,\theta)^T A J(p,q,\theta)$

$V = VJ(p,q,\theta)$

end 易实现并行化.

end

end

§7.3 二分法 用Sturm定理对特征值进行隔离的方法

$$T=\begin{pmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_n \end{pmatrix}$$

(一) 记$P_i(\lambda)$为$T-\lambda I$的i阶顺序主子式. 即$P_i(\lambda)$为T的i阶主子阵B_i的特征多项式.

则 $P_0(\lambda)=1$, $P_1(\lambda)=\alpha_1-\lambda$, $P_i(\lambda)=(\alpha_i-\lambda)P_{i-1}(\lambda)-\beta_{i-1}^2P_{i-2}(\lambda)$

$\{P_i(\lambda)\}$有如下性质:

① $\lim\limits_{\lambda\to-\infty} P_i(\lambda)=\infty$

② 相邻的两个多项式没有公共根

③ 若$P_i(u)=0$, 则$P_{i-1}(u)P_{i+1}(u)<0$

④ 记$P_i(\lambda)$的i个实根按其大小排列为$\lambda_1^{(i)},\lambda_2^{(i)},\cdots,\lambda_i^{(i)}$则有

$\lambda_1^{(i+1)}<\lambda_1^{(i)}<\lambda_2^{(i+1)}<\lambda_2^{(i)}<\cdots<\lambda_i^{(i+1)}<\lambda_i^{(i)}<\lambda_{i+1}^{(i+1)}$

即$P_i(\lambda)$的根都是单重的, $P_i(\lambda)$的根严格分隔$P_{i+1}(\lambda)$的根.

⑤ $P_n'(\lambda)$在$(-\infty,\infty)$上有$n-1$个零点, 分别分布在$(\lambda_k^{(n)},\lambda_{k+1}^{(n)})$上. 且$P_n'(\lambda_k^{(n)})\cdot P_{n-1}(\lambda_k^{(n)})<0$, $k=1,2,\cdots n$

注: 满足① $P_0(\lambda)$不变号, 性质③及$P_n(\lambda)$有单零点, 性质④

序列$P_n(\lambda), P_{n-1}(\lambda),\cdots, P_0(\lambda)$为$(-\infty,\infty)$上Sturm序列.

(二) 变号函数

1° $a_0,a_1\cdots a_m$为全不为0的实数. 若有$a_ka_{k+1}<0$称为序列的一个变号

而序列的变号总数称为序列$\{a_0,a_1,\cdots a_m\}$的变号数. 记为

$V\{a_0,a_1,\cdots a_m\}$. 例$V\{-3,-2,1,-1,-1,5\}=3$

2° $f_0(x), f_1(x), f_2(x)\cdots f_n(x)$为不恒等于0的解析函数

若$x=\xi$是某个$f_k(x)$的零点, 则称$x=\xi$为序列$\{f_0(x),f_1(x)\cdots f_n(x)\}$的零点

定义 $V(x)=V\{f_0(x),\cdots f_n(x)\}$ 为函数列的变号函数.

注：① $V(x)$ 仅可能在某些孤立零点无定义

② 记 $V(x+0)$ 为 V 在 x 处右极限，$V(x-0)$ 为 V 在 x 处左极限.

间断点必是函数列零点. 反之不然.

③ 在函数列两个零点之间 $V(x)$ 值不变. 分段片状函数.