

第4章 矩阵分解

4.1 矩阵的三角分解

4.2 矩阵的QR分解

4.3 矩阵的满秩分解

4.4 矩阵的奇异值分解

4.1 矩阵的三角分解

1. Gauss 消去法的矩阵形式
2. 矩阵的三角分解
3. 其他三角分解及算法
4. 分块矩阵的拟LU分解与拟LDU分解

1. Gauss 消去法的矩阵形式

- ❖ 我们知道对矩阵进行一次初等变换，就相当于用相应的初等矩阵去左乘原来的矩阵。因此我们从这个观点来考察Gauss消元法用矩阵乘法来表示，即可得到求解线性方程组的另一种直接法：矩阵的三角分解。

1 .Gauss消元过程

第1步等价于: $a_{11}^{(1)} \neq 0$ 时, 将 $a_{21}^{(1)}, a_{31}^{(1)}, \dots, a_{n1}^{(1)}$ 清零, 令 $l_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}$

则 (1)行 $\times (-l_{i1}) + (i)$ 行 $i = 2, 3, \dots, n$, 其矩阵形式为

$$\begin{bmatrix} 1 & & & & \\ -l_{21} & 1 & & & \\ -l_{31} & 0 & 1 & & \\ \dots & & & \dots & \\ -l_{n1} & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & \dots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \dots & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \dots & \dots & a_{nn}^{(1)} \end{bmatrix} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & \dots & a_{1n}^{(1)} \\ a_{22}^{(2)} & \dots & \dots & \dots & a_{2n}^{(2)} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n2}^{(2)} & \dots & \dots & \dots & a_{nn}^{(2)} \end{bmatrix}$$

L_1
 $A^{(1)}$
 $A^{(2)}$

同理第2步等价于：若 $a_{22}^{(2)} \neq 0$ 时，用矩阵

$$L_2 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & -l_{32} & 1 & \dots & 0 \\ \dots & \dots & \dots & \ddots & \vdots \\ 0 & -l_{n2} & 0 & \dots & 1 \end{bmatrix} \quad l_{i2} = \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} \quad (i = 3, 4, \dots, n)$$

$$\text{左乘 } A^{(2)}, \text{ 即有: } L_2 A^{(2)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \dots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \dots & a_{3n}^{(3)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & a_{n3}^{(3)} & \dots & a_{nn}^{(3)} \end{bmatrix} = A^{(3)}$$

以此类推可得

$$L_{n-1}L_{n-2}\cdots L_2L_1A = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} \end{bmatrix} = U$$

$$L_{n-1}L_{n-2}\cdots L_2L_1b = b^{(n)}$$

初等矩阵 $L_k = I - l_k e_k^T$.

其中, $l_k = (0, \cdots, 0, m_{k+1,k}, m_{k+2,k}, \cdots, m_{n,k})^T$,

$$m_{ik} = a_{ik}^{(k)} / a_{kk}^{(k)}, i = k+1, \cdots, n$$

我们称 m_{ik} 为消元因子, $a_{kk}^{(k)}$ 为主元素;

消元过程的一个重要性质是：

消元过程不改变矩阵的顺序主子矩阵的行列式(顺序主子式)的值。

例

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 3 & 1 \\ -2 & 1 & 0 \end{bmatrix}, \text{顺序主子式为, } 1, 5, -10$$

$$\xrightarrow{(2)+(1), (3)+2*(1)} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 5 & 0 \\ 0 & 5 & -2 \end{bmatrix}, \text{顺序主子式为, } 1, 5, -10$$

$$\xrightarrow{(3)-(2)} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 5 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \text{顺序主子式为, } 1, 5, -10$$

引理：约化的主元素 $a_{ii}^{(i)} \neq 0$ 的充要条件是

矩阵 A 的顺序主子式 $D_i \neq 0$ ($i=1, \dots, k$);

推论：若矩阵 A 的顺序主子式 $D_i \neq 0$

($i=1, \dots, k$), 则

$$a_{11}^{(1)} = D_1, \quad a_{ii}^{(i)} = D_i / D_{i-1}, i = 1, 2, \dots, k;$$

由此有若 A 对称正定或严格对角占优，而它们的顺序主子矩阵也是对称正定或严格对角占优，从而顺序主子式不为 0，顺序高斯消去过程可进行；

2.回代过程:

$$\begin{cases} x_n = b_n^{(n)} / a_{nn}^{(n)} \\ x_k = (b_k^{(k)} - \sum_{j=k+1}^n a_{kj}^{(k)} x_j) / a_{kk}^{(k)}, \\ k = n-1, n-2, \dots, 1 \end{cases}$$

$$\text{设 } \mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 3 & 1 \\ -2 & 1 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix},$$

用高斯消去法解线性方程 $\mathbf{Ax}=\mathbf{b}$.

$$\text{增广矩阵为 } \begin{bmatrix} 1 & 2 & -1 & -1 \\ -1 & 3 & 1 & 1 \\ -2 & 1 & 0 & 3 \end{bmatrix} \xrightarrow{(2)+(1), (3)+2*(1)} \begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & 5 & 0 & 0 \\ 0 & 5 & -2 & 1 \end{bmatrix}$$

$$\xrightarrow{(3)-(2)} \begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix} \xrightarrow{-(3)/2} \begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & -1/2 \end{bmatrix} \xrightarrow{(2)/5} \begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1/2 \end{bmatrix}$$

$$\xrightarrow{(1)+(3)-2*(2)} \begin{bmatrix} 1 & 0 & 0 & -3/2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1/2 \end{bmatrix}, \text{ 因此, 问题的解为 } \mathbf{x} = \begin{bmatrix} -3/2 \\ 0 \\ -1/2 \end{bmatrix}$$

3. 数值稳定性

1)选列主元;

2)选全主元;

3)高斯若当(Gauss-Jordan)消去法, 求矩阵的逆;

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 3 & 1 \\ -2 & 1 & 0 \end{bmatrix}, \text{ 求 } \mathbf{A}^{-1}.$$

$$\text{增广矩阵为 } \begin{bmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 & 1 & 0 \\ -2 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{(2)+(1), (3)+2*(1)} \begin{bmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 5 & 0 & 1 & 1 & 0 \\ 0 & 5 & -2 & 2 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{(3)-(2), (2)/5} \begin{bmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1/5 & 1/5 & 0 \\ 0 & 0 & -2 & 1 & -1 & 1 \end{bmatrix} \xrightarrow{-(3)/2, (1)+(3)} \begin{bmatrix} 1 & 2 & 0 & 1/2 & 1/2 & -1/2 \\ 0 & 1 & 0 & 1/5 & 1/5 & 0 \\ 0 & 0 & 1 & -1/2 & 1/2 & -1/2 \end{bmatrix}$$

$$\xrightarrow{(1)-(2)*2} \begin{bmatrix} 1 & 0 & 0 & 1/10 & 1/10 & -1/2 \\ 0 & 1 & 0 & 1/5 & 1/5 & 0 \\ 0 & 0 & 1 & -1/2 & 1/2 & -1/2 \end{bmatrix} \quad \text{从而}$$

$$\mathbf{A}^{-1} = \frac{1}{10} \begin{bmatrix} 1 & 1 & -5 \\ 2 & 2 & 0 \\ -5 & 5 & -5 \end{bmatrix}$$

4. 高斯顺序消元法解方程的计算量

1) 乘除次数: $n^3 / 3 + n^2 - n / 3$

2) 加减次数: $n^3 / 3 + n^2 / 2 - 5n / 6$

3) 求矩阵的逆的计算量为 $O(n^4)$

(II) 顺序消元过程与矩阵的三角分解

$$(1) \quad L_k^{-1} = (I - l_k e_k^T)^{-1} = I + l_k e_k^T$$

(2) 若 $i \leq j$, 则有 $e_i^T l_j = 0$, 从而

$$(I + l_i e_i^T)(I + l_j e_j^T) = I + l_i e_i^T + l_j e_j^T ,$$

$$L_1^{-1} L_2^{-1} \cdots L_{n-1}^{-1} = I + l_1 e_1^T + l_2 e_2^T + \cdots + l_{n-1} e_{n-1}^T$$

$$(3) \quad \text{由 } A^{(k+1)} = L_k A^{(k)}$$

$$\text{有 } A = A^{(1)} = L_1^{-1} L_2^{-1} \cdots L_{n-1}^{-1} A^{(n)}$$

故有 $A=LU$,

其中 $U = A^{(n)}$,

$$L = I + l_1 e_1^T + l_2 e_2^T + \cdots + l_{n-1} e_{n-1}^T .$$

这时 L 为单位下三角矩阵。

由此解线性方程组 $A\mathbf{x} = \mathbf{b}$ 就等价于解两个三角方程：

$$L(U\mathbf{x}) = \mathbf{b} \Rightarrow \begin{cases} L\mathbf{y} = \mathbf{b} \\ U\mathbf{x} = \mathbf{y} \end{cases}$$

因此，关键问题在于能否对矩阵 A 直接进行 LU 分解。

2 矩阵的三角分解

定义 4.1 如果方阵 A 可分解成一个下三角矩阵 L 和一个上三角矩阵 U 的乘积, 则称 A 可作三角分解或 $LU(LR)$ 分解. 如果方阵 A 可分解成 $A = LDU$, 其中 L 是单位下三角矩阵, D 是对角矩阵, U 是单位上三角矩阵, 则称 A 可作 LDU 分解.

定理 4.1 设 $A = (a_{ij})$ 是 n 阶矩阵, 则当且仅当 A 的顺序主子式 $\Delta_k \neq 0$ ($k = 1, 2, \dots, n-1$) 时, A 可唯一地分解为 $A = LDU$, 其中 L 是单位下三角矩阵, U 是单位上三角矩阵, D 是对角矩阵

$$D = \text{diag}(d_1, d_2, \dots, d_n)$$

其中, $d_k = \frac{\Delta_k}{\Delta_{k-1}}$ ($k = 1, 2, \dots, n; \Delta_0 = 1$).

推论 n 阶非奇异矩阵 A 有三角分解 $A = LU$ 的充要条件是 A 的顺序主子式 $\Delta_k \neq 0$ ($k = 1, 2, \dots, n-1$).

矩阵 A 的 LDU 与 LU 两种分解都需要假设 A 的前 $n-1$ 阶顺序主子式非零. 如果这个条件不满足, 可以给 A 左(或右)乘以置换矩阵, P (以 n 阶单位矩阵的 n 个列向量为列作成的 n 阶矩阵), 就把 A 的行(或列)的次序重新排列, 使之满足这个条件. 从而就有如下的带行交换的矩阵分解定理.

定理 4.2 设 A 是 n 阶非奇异矩阵, 则存在置换矩阵 P 使 PA 的 n 个顺序主子式非零.

推论 设 A 是 n 阶非奇异矩阵, 则存在置换矩阵 P , 使

$$PA = L\hat{U} = LDU \quad (4.1.21)$$

其中 L 是单位下三角矩阵, \hat{U} 是上三角矩阵, U 是单位上三角矩阵, D 是对角矩阵.

2 矩阵的其他三角分解

定义 4.2 设矩阵 A 有唯一的 LDU 分解. 若把 $A = LDU$ 中的 D 与 U 结合起来, 并且用 \hat{U} 来表示, 就得到唯一的分解

$$A = L(DU) = L\hat{U} \quad (4.1.23)$$

称为 A 的 **Doolittle** 分解; 若把 $A = LDU$ 中的 L 与 D 结合起来, 并且用 \hat{L} 来表示, 就得到唯一的分解

$$A = (LD)U = \hat{L}U \quad (4.1.24)$$

称为 A 的 **Crout** 分解.

Doolittle分解

令

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{bmatrix}$$

用比较等式两边元素的方法逐行逐列求解 L, U 各元素

Doolittle分解

$k = 1$ 时:

由 $a_{1j} = 1 \times u_{1j}$ 得 $u_{1j} = a_{1j} \ (j = 1, 2, \dots, n)$;

再由 $a_{i1} = u_{11}l_{i1}$ 得 $l_{i1} = \frac{a_{i1}}{u_{11}} \ (i = 2, 3, \dots, n)$ 。

$k = 2$ 时:

由 $a_{2j} = l_{21}u_{1j} + 1 \times u_{2j}$ 得 $u_{2j} = a_{2j} - l_{21}u_{1j} \ (j = 2, 3, \dots, n)$;

再由 $a_{i2} = l_{i1}u_{12} + l_{i2}u_{22}$ 得 $l_{i2} = \frac{a_{i2} - l_{i1}u_{12}}{u_{22}} \ (i = 3, 4, \dots, n)$ 。

Doolittle分解

第 k 步时：计算 $u_{kk}, u_{kk+1}, \dots, u_{kn}$ $j \geq k$

$$a_{kj} = [l_{k1} l_{k2} \dots l_{kk-1} 1 0 \dots 0] \begin{bmatrix} u_{1j} \\ u_{2j} \\ \vdots \\ u_{jj} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sum_{t=1}^{k-1} l_{kt} u_{tj} + u_{kj}$$

有 $u_{kj} = a_{kj} - \sum_{t=1}^{k-1} l_{kt} u_{tj} \quad (j = k, k+1, \dots, n)$

Doolittle分解

计算 l_{k+1k}, \dots, l_{nk} 由于 $i \geq k$, 于是由

$$a_{ik} = [l_{i1}, \dots, l_{ik-1}, 1, 0, \dots, 0] \begin{bmatrix} u_{1k} \\ \vdots \\ u_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sum_{t=1}^{k-1} l_{it} u_{tk} + l_{ik} u_{kk}$$

得
$$l_{ik} = (a_{ik} - \sum_{t=1}^{k-1} l_{it} u_{tk}) / u_{kk} \quad (i = k+1, \dots, n)$$

Doolittle分解

即
$$\begin{cases} u_{kj} = a_{kj} - \sum_{t=1}^{k-1} l_{kt} u_{tj} & (j = k, \dots, n; i = k + 1, \dots, n) \\ l_{ik} = (a_{ik} - \sum_{t=1}^{k-1} l_{it} u_{tk}) / u_{kk} & k = 1, 2, \dots, n \end{cases}$$

各元素在计算机内存于A的各位

$$\begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ l_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ l_{n1} & l_{n2} & \dots & u_{nn} \end{bmatrix}$$

二、紧凑格式算法: $A = LDU = \tilde{L}U$ (Crout分解)

$$L = \tilde{L} = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}, U = \begin{bmatrix} 1 & u_{12} & \cdots & u_{1n} \\ & 1 & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

$$(i,1)\text{元: } a_{i1} = l_{i1} \bullet 1 \Rightarrow l_{i1} = a_{i1} \quad (i = 1, \cdots, n)$$

$$(1,j)\text{元: } a_{1j} = l_{11} \bullet u_{1j} \Rightarrow u_{1j} = \frac{a_{1j}}{l_{11}} \quad (j = 2, \cdots, n)$$

$$(i,k)\text{元: } a_{ik} = l_{i1} \bullet u_{1k} + \cdots + l_{i,k-1} \bullet u_{k-1,k} + l_{ik} \bullet 1 \quad (i \geq k)$$
$$\Rightarrow l_{ik} = a_{ik} - (l_{i1} \bullet u_{1k} + \cdots + l_{i,k-1} \bullet u_{k-1,k})$$

$$(k,j)\text{元: } a_{kj} = l_{k1} \bullet u_{1,j} + \cdots + l_{k,k-1} \bullet u_{k-1,j} + l_{kk} \bullet u_{kj} \quad (j > k)$$
$$\Rightarrow u_{kj} = \frac{1}{l_{kk}} \left[a_{kj} - (l_{k1} \bullet u_{1j} + \cdots + l_{k,k-1} \bullet u_{k-1,j}) \right]$$

计算框图:

l_{11}	u_{12}	u_{13}	u_{14}	\cdots	第1框
l_{21}	l_{22}	u_{23}	u_{24}	\cdots	第2框
l_{31}	l_{32}	l_{33}	u_{34}	\cdots	第3框
l_{41}	l_{42}	l_{43}	l_{44}	\cdots	第4框
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots

$$A = \begin{bmatrix} 5 & 2 & -4 & 0 \\ 2 & 1 & -2 & 1 \\ -4 & -2 & 5 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

计算框图:

$$\begin{array}{c|c|c|c} 5 & 2/5 & -4/5 & 0 \\ \hline 2 & 1/5 & -2 & 5 \\ \hline -4 & -2/5 & 1 & 2 \\ \hline 0 & 1 & 2 & -7 \end{array}$$

$$\tilde{L} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 2 & 1/5 & 0 & 0 \\ -4 & -2/5 & 1 & 0 \\ 0 & 1 & 2 & -7 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ 2/5 & 1 & & \\ -4/5 & -2 & 1 & \\ 0 & 5 & 2 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 1/5 \\ 1 \\ -7 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 2/5 & -4/5 & 0 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$A = \tilde{L}U = LDU$$

3 对称矩阵的Cholesky分解

- ❖ 在应用数学中，线性方程组大多数的系数矩阵为对称正定这一性质，因此利用对称正定矩阵的三角分解式求解对称正定方程组的一种有效方法，且分解过程无需选主元，有良好的数值稳定性。

定义 4.3 称式(4.1.32)为实对称正定矩阵的 Cholesky 分解
(平方根分解、对称三角分解).

$$\mathbf{A} = \mathbf{L}\hat{\mathbf{D}}^2\mathbf{L}^T = (\mathbf{L}\hat{\mathbf{D}})(\mathbf{L}\hat{\mathbf{D}})^T = \mathbf{G}\mathbf{G}^T \quad (4.1.32)$$

这里 $\mathbf{G} = \mathbf{L}\hat{\mathbf{D}}$ 是下三角矩阵.

对称矩阵的Cholesky分解

- ❖ 定理3.2.4 设 A 为对称正定矩阵，则存在唯一分解 $A=LDL^T$ ，其中 L 为单位下三角阵， $D=\text{diag}(d_1, d_2, \dots, d_n)$ 且 $d_i > 0 (i=1, \dots, n)$

推论： 设A为对称正定矩阵， 则存在唯一分解

$$A = LL^T$$

其中L为具有主对角元素为正数的下三角矩阵。

Cholesky分解的求法

设A对称正定, 则 $A = LL^T$

$$\text{令 } L = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \dots & \dots & \ddots & \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix}$$

如何求 l_{ij} ? 以 $n=3$ 为例。

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & & \\ l_{21} & l_{22} & \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ & l_{22} & l_{23} \\ & & l_{33} \end{bmatrix}$$

Cholesky分解的求法

$k = 1$ 时：由 $a_{11} = l_{11}^2$ ，得 $l_{11} = \sqrt{a_{11}}$ ；

由 $a_{21} = l_{21}l_{11}$ ，得 $l_{21} = \frac{a_{21}}{l_{11}}$ ；同理得 $l_{31} = \frac{a_{31}}{l_{11}}$ 。

$k = 2$ 时：由 $a_{22} = l_{21}^2 + l_{22}^2$ ，得 $l_{22} = \sqrt{a_{22} - l_{21}^2}$ ；

由 $a_{32} = l_{31}l_{21} + l_{32}l_{22}$ ，得 $l_{32} = \frac{a_{32} - l_{31}l_{21}}{l_{22}}$ 。

Cholesky分解的求法

$$k = 3 \text{时: 由 } a_{33} = l_{31}^2 + l_{32}^2 + l_{33}^2, \text{ 得 } l_{33} = \sqrt{a_{33} - \sum_{i=1}^2 l_{3i}^2}$$

推广到 n 阶行列式,有

$$\begin{cases} l_{jj} = (a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2)^{\frac{1}{2}} \\ l_{ij} = (a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}) / l_{jj} \end{cases} \quad i = j+1, \dots, n, j = 1, 2, \dots, n$$

Cholesky分解法

用 $cholesky$ 分解法解线性方程组

$$AX = b \Leftrightarrow \begin{cases} Ly = b \\ L^T x = y \end{cases} \quad \text{其中 } A = LL^T$$

Cholesky分解法缺点及优点

优点：可以减少存储单元。

缺点：存在开方运算，可能会出现根号下负数。

改进Cholesky分解法

❖ 改进的cholesky分解 $A=LDL^T$

$$L = \begin{bmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \dots & \dots & \dots & \ddots & \\ l_{n1} & l_{n2} & \dots & l_{nn-1} & 1 \end{bmatrix} \quad D = \begin{bmatrix} d_1 & & & & \\ & d_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & d_{nn} \end{bmatrix}$$

由 $A = L(DL^T)$

$$= \begin{bmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \dots & \dots & \dots & \ddots & \\ l_{n1} & l_{n2} & \dots & l_{nn-1} & 1 \end{bmatrix} \begin{bmatrix} d_1 & d_1 l_{21} & d_1 l_{31} & \dots & d_1 l_{n1} \\ & d_2 & d_2 l_{32} & \dots & d_2 l_{n2} \\ & & d_3 & \dots & d_3 l_{n3} \\ \dots & \dots & \dots & \ddots & \\ & & & & d_n \end{bmatrix}$$

改进的cholesky分解

逐行相乘，并注意*i* > *j*有

$$a_{ij} = \sum_{k=1}^{j-1} l_{ik} d_k l_{jk} + l_{ij} d_j \quad (j = 1, 2, \dots, i-1)$$

$$a_{ii} = \sum_{k=1}^{j-1} l_{ik}^2 d_k + d_i \quad (i = 1, 2, \dots, n)$$

由此可得

$$\begin{cases} l_{ij} = (a_{ij} - \sum_{k=1}^{j-1} l_{ik} d_k l_{jk}) / d_j & (j = 1, 2, \dots, i-1) \\ d_i = a_{ii} - \sum_{k=1}^{j-1} l_{ik}^2 d_k & i = 1, 2, \dots, n \end{cases}$$

改进的cholesky分解

为减少计算量,可令 $c_{ij} = l_{ij}d_j$, 则 $l_{ij} = \frac{c_{ij}}{d_j}$

所以可将上述公式改写成

$$\left\{ \begin{array}{l} c_{ij} = a_{ij} - \sum_{k=1}^{j-1} c_{ik} l_{jk} \\ l_{ij} = \frac{c_{ij}}{d_j} \\ d_i = a_{ii} - \sum_{k=1}^{i-1} c_{ik} l_{ik} \end{array} \right. \quad (i = 2, 3, \dots, n, \quad j = 1, 2, \dots, i-1)$$

改进的cholesky分解算法

1.输入 $a_{ij} (i, j = 1, 2, \dots, n)$,

2.分解 $A = LDL^T$

对 $i = 2, 3, \dots, n$ 做

(1) 对 $j = 1, 2, \dots, i-1$ 做

$$c_{ij} = a_{ij} - \sum_{k=1}^{j-1} c_{ik} l_{ik}; \quad a_{ij} \leftarrow l_{ij} = \frac{c_{ij}}{d_j};$$

$$(2) a_{ii} \leftarrow d_i = a_{ii} - \sum_{k=1}^{i-1} c_{ik} l_{ik};$$

例题

$$\begin{aligned}
 A &= \begin{bmatrix} 4 & -1 & 1 \\ -1 & 2 & -2 \\ 1 & -2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & -1 & 1 \\ -0.25 & 2 & -2 \\ 0.25 & -2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & -1 & 1 \\ -0.25 & 1.75 & -1.75 \\ 0.25 & -1 & 3 \end{bmatrix} \\
 &\rightarrow \begin{bmatrix} 4 & -1 & 1 \\ -0.25 & 1.75 & -1.75 \\ 0.25 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -0.25 & & \\ 0.25 & -1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 & 1 \\ 1.75 & -1.75 & \\ & & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & & \\ -0.25 & & \\ 0.25 & -1 & 1 \end{bmatrix} \begin{bmatrix} 4 & & \\ 1.75 & & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & -0.25 & 0.25 \\ & 1 & -1 \\ & & 1 \end{bmatrix} = LDL^T
 \end{aligned}$$

- ❖ $A=LDL^T$ 分解，既适合于解对称正定方程组，也适合求解 A 为对称，而各阶顺序主子式不为零的方程组
- ❖ 而对 $A=LL^T$ 只适合于对称正定方程组

四、分块矩阵的拟 LU 分解和拟 LDU 分解

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

若 A_{11} 可逆, 作 $L = \begin{bmatrix} I_{n1} & 0 \\ -A_{21}A_{11}^{-1} & I_{n2} \end{bmatrix}$

则 $LA = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}$

从而 $\det(A) = \det(A_{11}) \cdot \det(A_{22} - A_{21}A_{11}^{-1}A_{12})$

同样若 A_{22} 可逆, 可得类似结果。

例 4.3 设 $\mathbf{A} \in \mathbf{R}^{m \times n}$, $\mathbf{B} \in \mathbf{R}^{n \times m}$, 则有

$$\det(\mathbf{I}_m + \mathbf{AB}) = \det(\mathbf{I}_n + \mathbf{BA}) \quad (4.1.42)$$

特别是对于 $\mathbf{a} \in \mathbf{R}^{n \times 1}$, $\mathbf{b} \in \mathbf{R}^{n \times 1}$, 有 $\det(\mathbf{I}_n + \mathbf{ba}^T) = 1 + \mathbf{a}^T \mathbf{b}$.

证 分别使用式(4.1.40)和式(4.1.41)来计算下面的分块矩阵

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{B} \\ -\mathbf{A} & \mathbf{I}_m \end{bmatrix}$$

的行列式, 便有

$$\det \begin{bmatrix} \mathbf{I}_n & \mathbf{B} \\ -\mathbf{A} & \mathbf{I}_m \end{bmatrix} = \det(\mathbf{I}_n + \mathbf{BA}) = \det(\mathbf{I}_m + \mathbf{AB})$$

类似地又有

$$\det \begin{bmatrix} \mathbf{I}_n & \mathbf{b} \\ -\mathbf{a}^T & 1 \end{bmatrix} = \det(\mathbf{I}_n + \mathbf{ba}^T) = \det(1 + \mathbf{a}^T \mathbf{b}) = 1 + \mathbf{a}^T \mathbf{b}$$

矩阵求逆引理(Woodbury 公式)

$$(\mathbf{A}+\mathbf{BC})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{B}(\mathbf{I}+\mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}$$

证明: 求方程 $(\mathbf{A}+\mathbf{BC})\mathbf{x}=\mathbf{b}$,

令 $\mathbf{y}=\mathbf{C}\mathbf{x}$, 则有
$$\begin{cases} \mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{y}=\mathbf{b} \\ -\mathbf{C}\mathbf{x}+\mathbf{y}=\mathbf{0} \end{cases}$$

写成矩阵为

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix},$$

利用高斯消去法有

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{b} \\ -\mathbf{C} & \mathbf{I} & 0 \end{bmatrix}$$

$$\xrightarrow{(2)+\mathbf{CA}^{-1}*(1)} \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{b} \\ 0 & \mathbf{I} + \mathbf{CA}^{-1}\mathbf{B} & \mathbf{CA}^{-1}\mathbf{b} \end{bmatrix}$$

$$\xrightarrow{(\mathbf{I}+\mathbf{CA}^{-1}\mathbf{B})^{-1}*(2)} \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{b} \\ 0 & \mathbf{I} & (\mathbf{I} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}\mathbf{b} \end{bmatrix}$$

$$\xrightarrow{(1)-\mathbf{B}*(2)} \begin{bmatrix} \mathbf{A} & 0 & (\mathbf{I} - \mathbf{B}(\mathbf{I} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1})\mathbf{b} \\ 0 & \mathbf{I} & (\mathbf{I} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}\mathbf{b} \end{bmatrix}$$

$$\xrightarrow{\mathbf{A}^{-1}*(1)} \begin{bmatrix} \mathbf{I} & 0 & (\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1})\mathbf{b} \\ 0 & \mathbf{I} & (\mathbf{I} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}\mathbf{b} \end{bmatrix}$$

因此可得

$$\mathbf{x} = (\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1})\mathbf{b},$$

而由 $(\mathbf{A} + \mathbf{BC})\mathbf{x} = \mathbf{b}$ 可得 $\mathbf{x} = (\mathbf{A} + \mathbf{BC})^{-1}\mathbf{b}$

由于 \mathbf{b} 的任意性可得

$$(\mathbf{A} + \mathbf{BC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}$$

从而命题得证。

推论: $(\mathbf{A} + \mathbf{BD}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}$

$$\text{例 } \mathbf{A} = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 5 & -1 \\ -1 & 2 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{bmatrix} + \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$$

由于

$$\begin{bmatrix} 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} = \frac{1}{6}$$

利用矩阵求逆引理有

$$\begin{aligned}\mathbf{A}^{-1} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{bmatrix}^{-1} - \\ &\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \left(1 + \frac{1}{6}\right)^{-1} [1 \quad -2 \quad 1] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & -1/2 \end{bmatrix} - \frac{6}{7} \begin{bmatrix} -1 \\ -1/3 \\ 1/2 \end{bmatrix} [1 \quad -2/3 \quad -1/2] \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & -1/2 \end{bmatrix} - \frac{6}{7} \begin{bmatrix} -1 \\ -1/3 \\ 1/2 \end{bmatrix} [1 \quad -2/3 \quad -1/2]\end{aligned}$$

Schur 补

设矩阵 $A \in \mathbb{C}^{m \times n}$ 为非奇异, $I, J \subset \{1, 2, \dots, n\}$ 为有序的指标集
 $I \neq \emptyset, I \neq \{1, 2, \dots, n\}$ 。令 $R = \{1, 2, \dots, n\} / I, S = \{1, 2, \dots, n\} / J$ 。

假定 A_I, A_{IJ} 非奇异,

则 Schur 补 A/A_I 定义为:

$$A/A_I = A_R - A_{R,I} (A_I)^{-1} A_{I,R}$$

$$A/A_{IJ} = A_{R,S} - A_{R,J} (A_{IJ})^{-1} A_{I,S}$$

(当 A_{IJ} 不可逆时, 使用广义逆 $(A_{IJ})^+$)

其中, A_I 表示由矩阵 A 的元素其行数和列数在 I 中
组成主子矩阵, 其余类似。

定理：假设 A_I 非奇异，那么 A 非奇异的充要条件为 A/A_I 非奇异，此时我们有：

$$(A^{-1})_R = (A/A_I)^{-1} \quad (1)$$

$$(A^{-1})_{R,I} = -(A/A_I)^{-1} A_{R,I} (A_I)^{-1} \quad (2)$$

$$(A^{-1})_{I,R} = - (A_I)^{-1} A_{I,R} (A/A_I)^{-1} \quad (3)$$

$$(A^{-1})_I = (A_I)^{-1} + (A_I)^{-1} A_{I,R} (A/A_I)^{-1} A_{R,I} (A_I)^{-1} \quad (4)$$

从定理的条件看我们发现 I 和 R 的位置可以完全互换，因此交换 I 和 R 时等式一定成立。

由公式(1)和(4)可以得到Woodbury公式.

Schur 补的商性质:

$$\text{设 } M = \begin{pmatrix} A & B & E \\ C & D & F \\ G & H & L \end{pmatrix}, A' = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

$$B' = \begin{pmatrix} B & E \\ D & F \end{pmatrix}, C' = \begin{pmatrix} C & D \\ G & H \end{pmatrix}, D' = \begin{pmatrix} D & F \\ H & L \end{pmatrix},$$

假设 D'/D 非奇异, 则商性质为

$$M/D' = ((M/D)/(D'/D)) = A'/D - (B'/D)(D'/D)^{-1}(C'/D)$$

$$\text{证明: } M/D' = A - (B, E)(D')^{-1} \begin{pmatrix} C \\ G \end{pmatrix}$$

$$(D')^{-1} = \begin{pmatrix} D^{-1} + D^{-1}F(D'/D)^{-1}HD^{-1} & -D^{-1}F(D'/D)^{-1} \\ -(D'/D)^{-1}HD^{-1} & (D'/D)^{-1} \end{pmatrix}$$

$$\text{从而 } M/D' = A - (B, E)(D')^{-1} \begin{pmatrix} C \\ G \end{pmatrix} = A - B\{D^{-1} + D^{-1}F(D'/D)^{-1}HD^{-1}\}C$$

$$\begin{aligned} & + BD^{-1}F(D'/D)^{-1}G + E(D'/D)^{-1}HD^{-1}C - E(D'/D)^{-1}G \\ = & A - BD^{-1}C - BD^{-1}F(D'/D)^{-1}HD^{-1}C \\ & + BD^{-1}F(D'/D)^{-1}G + E(D'/D)^{-1}HD^{-1}C - E(D'/D)^{-1}G \\ = & A'/D - BD^{-1}F(D'/D)^{-1}(HD^{-1}C - G) + E(D'/D)^{-1}(HD^{-1}C - G) \\ = & A'/D + (E - BD^{-1}F)(D'/D)^{-1}(HD^{-1}C - G) \\ = & A'/D - (E - BD^{-1}F)(D'/D)^{-1}(G - HD^{-1}C) \\ = & A'/D - (B'/D)(D'/D)^{-1}(C'/D) \end{aligned}$$

从而得证.

$$M/D = \begin{pmatrix} A'/D & B'/D \\ C'/D & D'/D \end{pmatrix}$$

当 A,E,G,L 为数时, 这时等式的两边为标量

从而 $A'/D = \det(A')/\det(D)$, $B'/D = \det(B')/\det(D)$,

$C'/D = \det(C')/\det(D)$, $D'/D = \det(D')/\det(D)$,

$M/D' = \det(M)/\det(D')$.

这时我们有

$$\det(M) \cdot \det(D) = \det(A') \det(D') - \det(B') \cdot \det(C')$$

这就是所谓的 Sylvester 行列式恒等式.

4.2 矩阵的QR分解

一、Givens 变换与 Householder 变换

二、矩阵的 QR(正交三角)分解

三、矩阵与 Hessenberg 矩阵的正交相似问题

1. Givens 矩阵和 Givens 变换

定义 4.4 设实数 c 与 s 满足 $c^2 + s^2 = 1$, 称

$$T_{ij} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & c & & s \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \\ & & & -s & & c & & \\ & & & & & & 1 & \\ & & & & & & & \ddots \\ & & & & & & & & 1 \end{bmatrix} \quad \begin{matrix} (i) \\ \\ \\ (i < j) \\ (j) \end{matrix} \quad (4.2.1)$$

为 Givens 矩阵 (初等旋转矩阵), 也可记作 $T_{ij} = T_{ij}(c, s)$. 由 Givens 矩阵确定的线性变换称为 Givens 变换 (初等旋转变换).

性质 1. Givens 矩阵是正交矩阵, 且有

$$[T_{ij}(c,s)]^{-1}=[T_{ij}(c,s)]^T=T_{ij}(c,-s)$$

$$\det[T_{ij}(c,s)]=1$$

性质 2 设 $\mathbf{x}=(\xi_1,\xi_2,\dots,\xi_n)^T, \mathbf{y}=T_{ij}\mathbf{x}=(\eta_1,\eta_2,\dots,\eta_n)^T$, 则有

$$\left. \begin{aligned} \eta_i &= c\xi_j + s\xi_j \\ \eta_j &= -s\xi_j + c\xi_j \\ \eta_k &= \xi_k \quad (k \neq i, j) \end{aligned} \right\}$$

定理 4.3 设 $\mathbf{x}=(\xi_1,\xi_2,\dots,\xi_n)^T\in\mathbb{R}^n$, $\mathbf{x}\neq\mathbf{0}$, 则存在有限个 Givens 矩阵的乘积, 记作 \mathbf{T} , 使得 $\mathbf{T}\mathbf{x}=|\mathbf{x}|\mathbf{e}_1$.

证 先考虑 $\xi_1 \neq 0$ 的情形. 对 x 构造 Givens 矩阵 $T_{12}(c, s)$:

$$c = \frac{\xi_1}{\sqrt{\xi_1^2 + \xi_2^2}}, \quad s = \frac{\xi_2}{\sqrt{\xi_1^2 + \xi_2^2}}$$

$$T_{12}x = (\sqrt{\xi_1^2 + \xi_2^2}, 0, \xi_3, \dots, \xi_n)^T$$

再对 $T_{12}x$ 构造 Givens 矩阵 $T_{13}(c, s)$:

$$c = \frac{\sqrt{\xi_1^2 + \xi_2^2}}{\sqrt{\xi_1^2 + \xi_2^2 + \xi_3^2}}, \quad s = \frac{\xi_3}{\sqrt{\xi_1^2 + \xi_2^2 + \xi_3^2}}$$

$$T_{13}(T_{12}x) = (\sqrt{\xi_1^2 + \xi_2^2 + \xi_3^2}, 0, 0, \xi_4, \dots, \xi_n)^T$$

如此继续下去, 最后对 $T_{1,n-1} \cdots T_{12}x$ 构造 Givens 矩阵 $T_{1n}(c, s)$:

$$c = \frac{\sqrt{\xi_1^2 + \cdots + \xi_{n-1}^2}}{\sqrt{\xi_1^2 + \cdots + \xi_{n-1}^2 + \xi_n^2}}, \quad s = \frac{\xi_n}{\sqrt{\xi_1^2 + \cdots + \xi_{n-1}^2 + \xi_n^2}}$$

$$T_{1n}(T_{1,n-1} \cdots T_{12}x) = (\sqrt{\xi_1^2 + \cdots + \xi_{n-1}^2 + \xi_n^2}, 0, \dots, 0)^T$$

令 $T = T_{1n}T_{1,n-1} \cdots T_{12}$, 则有 $Tx = |x| e_1$.

如果 $\xi_1 = 0$, 考虑 $\xi_1 = \cdots = \xi_{k-1} = 0, \xi_k \neq 0 (1 < k \leq n)$ 的情形. 此时 $|x| = \sqrt{\xi_k^2 + \cdots + \xi_n^2}$, 上面的步骤由 T_{1k} 开始进行即得结论亦成立.

证毕

推论. 设非零列向量 $\mathbf{x} \in \mathbb{R}^n$ 及单位列向量 $\mathbf{z} \in \mathbb{R}^n$, 则存在有限个 Givens 变换矩阵的乘积, 记作 \mathbf{T} , 使得 $\mathbf{T}\mathbf{x} = |\mathbf{x}| \mathbf{z}$.

例: $x = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$ 求G-矩阵之积 T , 使得 $Tx = |x|e_1$

解: $T_{12}(c, s)$ 中, $c = \frac{3}{5}, s = \frac{4}{5}$. $T_{12}x = \begin{bmatrix} 5 \\ 0 \\ 5 \end{bmatrix}$

$T_{13}(c, s)$ 中, $c = \frac{1}{\sqrt{2}}, s = \frac{1}{\sqrt{2}}$. $T_{13}(T_{12}x) = \begin{bmatrix} 5\sqrt{2} \\ 0 \\ 0 \end{bmatrix} = |x|e_1$

$$T = T_{13}T_{12} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{bmatrix} \cdot \frac{1}{5} \begin{bmatrix} 3 & 4 & 0 \\ -4 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$= \frac{1}{5\sqrt{2}} \begin{bmatrix} 3 & 4 & 5 \\ -4\sqrt{2} & 3\sqrt{2} & 0 \\ -3 & -4 & 5 \end{bmatrix}$$

$$Tx = 5\sqrt{2}e_1$$

2. Householder 矩阵和 Householder 变换

定义 4.5 设单位列向量 $\mathbf{u} \in \mathbb{R}^n$, 称

$$\mathbf{H} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T$$

为 Householder 矩阵(初等反射矩阵), 由 Householder 矩阵确定的线性变换为 Householder 变换(初等反射变换).

- (1). $\mathbf{H}^T = \mathbf{H}$ (对称矩阵)
- (2). $\mathbf{H}^T \mathbf{H} = \mathbf{I}$ (正交矩阵)
- (3). $\mathbf{H}^2 = \mathbf{I}$ (对合矩阵)
- (4) $\mathbf{H}^{-1} = \mathbf{H}$ (自逆矩阵)
- (5) $\det(\mathbf{H}) = -1$

定理 4.4 任意给定非零列向量 $\mathbf{x} \in \mathbb{R}^n (n > 1)$ 及单位列向量 $\mathbf{z} \in \mathbb{R}^n$, 则存在 Householder 矩阵 \mathbf{H} , 使得 $\mathbf{H}\mathbf{x} = |\mathbf{x}| \mathbf{z}$.

证明: (1) $\mathbf{x} = |\mathbf{x}| \mathbf{z} : n > 1$ 时, 取单位向量 \mathbf{u} 使得 $\mathbf{u} \perp \mathbf{x}$,

于是 $\mathbf{H}_u = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T : \mathbf{H}_u \mathbf{x} = \mathbf{I}\mathbf{x} - 2\mathbf{u}\mathbf{u}^T \mathbf{x} = \mathbf{x} = |\mathbf{x}| \mathbf{z}$

(2) $\mathbf{x} \neq |\mathbf{x}| \mathbf{z} :$ 取 $\mathbf{u} = \frac{\mathbf{x} - |\mathbf{x}| \mathbf{z}}{|\mathbf{x} - |\mathbf{x}| \mathbf{z}|}$, 有

$$\begin{aligned} \mathbf{H}_u \mathbf{x} &= \left[\mathbf{I} - 2 \frac{(\mathbf{x} - |\mathbf{x}| \mathbf{z})(\mathbf{x} - |\mathbf{x}| \mathbf{z})^T}{|\mathbf{x} - |\mathbf{x}| \mathbf{z}|^2} \right] \mathbf{x} = \mathbf{x} - \frac{2(\mathbf{x} - |\mathbf{x}| \mathbf{z}, \mathbf{x})}{|\mathbf{x} - |\mathbf{x}| \mathbf{z}|^2} (\mathbf{x} - |\mathbf{x}| \mathbf{z}) \\ &= \mathbf{x} - 1 \times (\mathbf{x} - |\mathbf{x}| \mathbf{z}) = |\mathbf{x}| \mathbf{z} \end{aligned}$$

例2: $x = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, 求H-矩阵 H 使得 $Hx = |x|e_1$

解: $|x| = 3, x - |x|e_1 = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}, \quad u = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

$$H = I - \frac{2}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$$

$$Hx = 3e_1$$

定理 4.5 初等旋转矩阵(Givens 变换)是两个初等反射矩阵(Householder 变换)的乘积。

证明：对初等旋转矩阵 T_{ij} , 取

$$\mathbf{u}=(0,\dots,0,\sin(\theta/4),0,\dots,0,\cos(\theta/4),0,\dots,0)^T$$

$$\mathbf{v}=(0,\dots,0,\sin(3\theta/4),0,\dots,0,\cos(3\theta/4),0,\dots,0)^T$$

直接计算可得

$$T_{ij}=H_v H_u$$

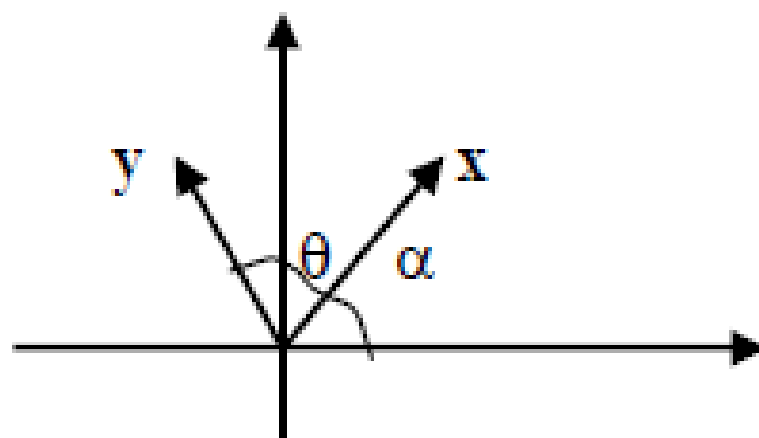
[注] \mathbf{H} -矩阵不能由若干个 \mathbf{G} 矩阵的乘积来表示。

因为 $\det \mathbf{H} = -1$, 而 $\det \mathbf{G} = 1$

\mathbf{x} 为变换前的任意向量，用它与正半实轴的夹角 α 表示。

\mathbf{y} 为变换后的向量， 用它与正半实轴的夹角表示为

$$\alpha+2(\pi/2+\theta-\alpha)=\pi+2\theta-\alpha。$$



$$G_{\theta}(\alpha)=\theta+\alpha$$

其中， θ 为旋转角度；

\mathbf{x} 为变换前的任意向量，用它与正半实轴的夹角 α 表示；

\mathbf{y} 为变换后的向量， 用它与正半实轴的夹角表示为 $\theta+\alpha$ 。

我们需要证明的目标就是找到两个向量 \mathbf{u}_1 和 \mathbf{u}_2 使得它们对应的 Householder 变换的乘积等于给定的一个旋转变换 $\mathbf{G}_\theta(\alpha)$ 。

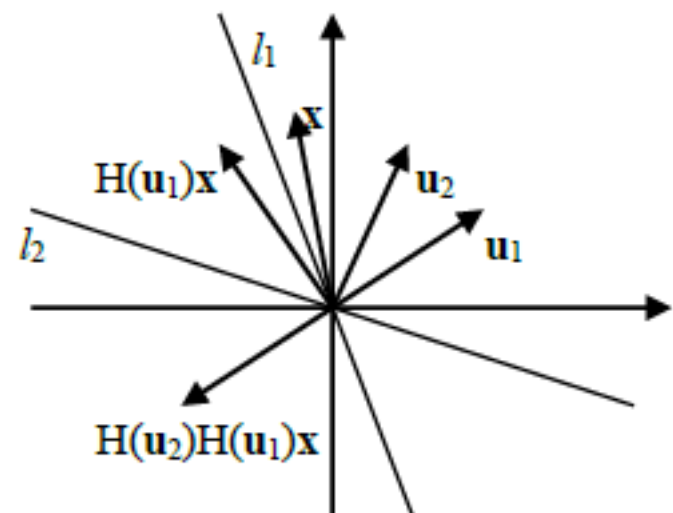
根据前面的讨论我们对于 H 变换和 G 变换都可以用向量的角度表示。因此有

$$\mathbf{H}_{\theta_2}\mathbf{H}_{\theta_1}(\alpha)=\pi+2\theta_2-(\pi+2\theta_1-\alpha)=2(\theta_2-\theta_1)+\alpha=\theta+\alpha=\mathbf{G}_\theta(\alpha)$$

可见 $\theta=2(\theta_2-\theta_1)$ 成立即可。

这个等式在几何成立的意思：

\mathbf{x} 和 $\mathbf{H}(\mathbf{u}_2)\mathbf{H}(\mathbf{u}_1)\mathbf{x}$ 之间的夹角等于二倍的 \mathbf{u}_1 和 \mathbf{u}_2 之间的夹角。



这在几何上是显然的。因为 \mathbf{u}_1 对应的垂线 l_1 平分 \mathbf{x} 和 $H(\mathbf{u}_1)\mathbf{x}$ 的夹角；

同样，因为 \mathbf{u}_2 对应的垂线 l_2 平分 $H(\mathbf{u}_1)\mathbf{x}$ 和 $H(\mathbf{u}_2)H(\mathbf{u}_1)\mathbf{x}$ 的夹角。

注意到 l_1 和 l_2 的夹角和向量 \mathbf{u}_1 和 \mathbf{u}_2 的夹角相等。因此结论是显然的。

换句话说，对于给定的旋转变换，仅需找到夹角为旋转角度的二分之一的两个向量，用它们作反射变换，就可以得到旋转变换。

例3: G-矩阵 $T_{ij}(0,1) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ 中, $c = 0, s = 1 \Rightarrow \theta = \pi/2$

$$H_u = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}, H_v = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}$$

$$\Rightarrow H_u H_v = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

定理(Householder矩阵的乘积公式)

$$(I - U_1 V_1^T) \cdot (I - U_2 V_2^T) \cdots (I - U_k V_k^T) = I - U(I + R)^{-1} V^T$$

其中 $U = (U_1, U_2, \dots, U_k)$, $V = (V_1, V_2, \dots, V_k)$

$$R = \begin{pmatrix} 0 & V_1^T U_2 & \cdots & V_1^T U_{k-1} & V_1^T U_k \\ 0 & 0 & \cdots & V_2^T U_{k-1} & V_2^T U_k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & V_{k-1}^T U_k \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

为 $V^T U$ 分块乘积矩阵的严格上三角部分。

若记 $U = (U_k, U_{k-1}, \dots, U_1)$, $V = (V_k, V_{k-1}, \dots, V_1)$

$$(I - U_1 V_1^T) \cdot (I - U_2 V_2^T) \cdots (I - U_k V_k^T)$$

$$= I - U(I + L)^{-1} V^T$$

其中 L 为为 $V^T U$ 分块乘积矩阵的严格下三角部分。
证明:可以使用数学归纳法.

二、矩阵的 QR(正交三角)分解

定义4.6 如果实(复)矩阵**A**能够化成正交(酉)矩阵**Q**与实(复)非奇异上三角矩阵**R**的乘积, 即

$$\mathbf{A}=\mathbf{QR}$$

则称为**A**的**QR**分解。

定理4.6 设**A**是**n** 阶实(复)非奇异矩阵, 则存在正交(酉)矩阵**Q**与实(复)非奇异上三角矩阵**R**使得

$$\mathbf{A}=\mathbf{QR}$$

且除去相差一个对角元素的绝对值(模)全等于1的对角矩阵外, 分解唯一,
即**QR**分解存在唯一。

证明: $A=(a_1,a_2,\cdots,a_n)$ 可逆 $\Rightarrow a_1,a_2,\cdots,a_n$ 线性无关,

正交化后可得:

$$\left\{ \begin{array}{l} b_1 = a_1 \\ b_2 = a_2 - k_{21}b_1 \\ \dots\dots\dots \\ b_n = a_n - k_{n,n-1}b_{n-1} - \dots - k_{n1}b_1 \end{array} \right. \quad \left\{ \begin{array}{l} a_1 = b_1 \\ a_2 = k_{21}b_1 + b_2 \\ \dots\dots\dots \\ a_n = k_{n1}b_1 + \dots + k_{n,n-1}b_{n-1} + b_n \end{array} \right.$$

$$(a_1, a_2, \cdots, a_n) = (b_1, b_2, \cdots, b_n)K$$

$$= (q_1, q_2, \cdots, q_n) \begin{bmatrix} |b_1| & & & \\ & |b_2| & & \\ & & \ddots & \\ & & & |b_n| \end{bmatrix} \begin{bmatrix} 1 & k_{21} & \cdots & k_{n1} \\ & 1 & \cdots & k_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

$$\text{令 } Q = (q_1, q_2, \cdots, q_n), R = \begin{bmatrix} |b_1| & & & \\ & |b_2| & & \\ & & \ddots & \\ & & & |b_n| \end{bmatrix} \begin{bmatrix} 1 & k_{21} & \cdots & k_{n1} \\ & 1 & \cdots & k_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

$$\text{则 } A = QR, \text{ 其中 } q_i = \frac{b_i}{|b_i|} \quad (i = 1, 2, \cdots, n)$$

唯一性(只证实数情形): $A=QR=Q_1R_1$ 从而
 $P=(Q_1)^TQ= R_1 R^{-1}$

注意到 $I=P^TP$ 所以 $P^{-1}=P^T$,而 P 为上三角矩阵, 其逆也为上三角矩阵, 而 P^T 为下三角矩阵, 从而 P 为对角矩阵.

又 $P^2=I$, 所以 P 的对角元只能为正负1。

$$Q= Q_1P, R= R_1P.$$

例4: 求 $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ 的QR分解。

解: $b_1 = a_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, b_2 = a_2 - 1 \times b_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, b_3 = a_3 - \frac{1}{3}b_2 - \frac{7}{6}b_1 = \begin{bmatrix} 1/2 \\ 2 \\ -1/2 \end{bmatrix}$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \quad R = \begin{bmatrix} \sqrt{6} & & \\ & \sqrt{3} & \\ & & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & \frac{7}{6} \\ & 1 & \frac{1}{3} \\ & & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{6} & \sqrt{6} & \frac{7}{\sqrt{6}} \\ & \sqrt{3} & \frac{1}{\sqrt{3}} \\ & & \frac{1}{\sqrt{2}} \end{bmatrix}$$

定理4.7 设 A 是 $m \times n$ 实(复)矩阵, 且其 n 个列线性无关, 则 A 有分解

$$A=QR$$

其中 Q 是 $m \times n$ 实(复)矩阵, 且满足 $Q^T Q = I$, R 是实(复)非奇异上三角矩阵且除去相差一个对角元素的绝对值(模)全等于1的对角矩阵外, 分解唯一。

定理4.8 任何n阶实非奇异矩阵A可通过左连乘初等Givens 旋转矩阵化为上三角矩阵。

证 第1步:由 $\det A \neq 0$ 知,A 的第1列 $\mathbf{b}^{(1)} = (a_{11}, a_{21}, \dots, a_{n1})^T \neq \mathbf{0}$.

根据定理 4.3,存在有限个 Givens 矩阵的乘积,记作 \mathbf{T}_1 , 使得

$$\mathbf{T}_1 \mathbf{b}^{(1)} = |\mathbf{b}^{(1)}| \mathbf{e}_1 \quad (\mathbf{e}_1 \in \mathbf{R}^n)$$

令 $a_{11}^{(1)} = |\mathbf{b}^{(1)}|$, 则有

$$\mathbf{T}_1 \mathbf{A} = \left[\begin{array}{c|ccc} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ \hline 0 & & & \\ \vdots & & \mathbf{A}^{(1)} & \\ 0 & & & \end{array} \right]$$

第2步:由 $\det A^{(1)} \neq 0$ 知, $A^{(1)}$ 的第1列 $b^{(2)} = (a_{22}^{(1)}, a_{32}^{(1)}, \dots, a_{n2}^{(1)})^T \neq 0$. 根据定理 4.3, 存在有限个 Givens 矩阵的乘积, 记作 T_2 , 使得

$$T_2 b^{(2)} = | b^{(2)} | e_1 \quad (e_1 \in \mathbb{R}^{n-1})$$

令 $a_{22}^{(2)} = | b^{(2)} |$, 则有

$$T_2 A^{(1)} = \left[\begin{array}{c|ccc} a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ \hline 0 & & & \\ \vdots & & A^{(2)} & \\ 0 & & & \end{array} \right]$$

.....

第 $n-1$ 步: 由 $\det A^{(n-2)} \neq 0$ 知, $A^{(n-2)}$ 的第 1 列 $b^{(n-1)} = (a_{n-1, n-1}^{(n-2)}, a_{n, n-1}^{(n-2)})^T \neq \mathbf{0}$. 根据定理 4.3, 存在 Givens 矩阵 T_{n-1} , 使得

$$T_{n-1} b^{(n-1)} = |b^{(n-1)}| e_1 \quad (e_1 \in \mathbf{R}^2)$$

令 $a_{n-1, n-1}^{(n-1)} = |b^{(n-1)}|$, 则有

$$T_{n-1} A^{(n-2)} = \begin{bmatrix} a_{n-1, n-1}^{(n-1)} & a_{n-1, n}^{(n-1)} \\ 0 & a_{nn}^{(n-1)} \end{bmatrix}$$

最后, 令

$$T = \begin{bmatrix} I_{n-2} & \mathbf{0} \\ \mathbf{0} & T_{n-1} \end{bmatrix} \cdots \begin{bmatrix} I_2 & \mathbf{0} \\ \mathbf{0} & T_3 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & T_2 \end{bmatrix} T_1$$

则 T 是有限个 Givens 矩阵的乘积, 使得

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

证毕

定理4.10 任何n阶实非奇异矩阵A可通过左连乘Householder矩阵化为上三角矩阵。

证 以 $n=4$ 为例.

$$(1) |A| \neq 0: \beta^{(0)} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{bmatrix} \neq 0 \Rightarrow \exists \text{H-矩阵 } H_0, \text{ 使得 } H_0 \beta^{(0)} = \begin{bmatrix} |\beta^{(0)}| \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$$a_{11}^{(1)} = |\beta^{(0)}| > 0$$

$$H_0 A = \left[\begin{array}{c|ccc} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ \hline 0 & & & \\ 0 & & A^{(1)} & \\ 0 & & & \end{array} \right], \quad A^{(1)} = \begin{bmatrix} a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \end{bmatrix}$$

$$(2) \quad |A^{(1)}| \neq 0: \quad \beta^{(1)} = \begin{bmatrix} a_{22}^{(1)} \\ a_{32}^{(1)} \\ a_{42}^{(1)} \end{bmatrix} \neq 0 \Rightarrow \exists \text{H-矩阵 } H_1, \text{ 使得 } H_1 \beta^{(1)} = \begin{bmatrix} |\beta^{(1)}| \\ 0 \\ 0 \end{bmatrix}.$$

$$a_{22}^{(2)} = |\beta^{(1)}| > 0$$

$$H_1 A^{(1)} = \left[\begin{array}{c|cc} a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ \hline 0 & a_{33}^{(2)} & a_{34}^{(2)} \\ 0 & a_{43}^{(2)} & a_{44}^{(2)} \end{array} \right], \quad A^{(2)} = \begin{bmatrix} a_{33}^{(2)} & a_{34}^{(2)} \\ a_{43}^{(2)} & a_{44}^{(2)} \end{bmatrix}$$

$$(3) \quad |A^{(2)}| \neq 0: \quad \beta^{(2)} = \begin{bmatrix} a_{33}^{(2)} \\ a_{43}^{(2)} \end{bmatrix} \neq 0 \Rightarrow \exists \mathbf{H}\text{-矩阵 } H_2, \text{ 使得 } H_2 \beta^{(2)} = \begin{bmatrix} |\beta^{(2)}| \\ 0 \end{bmatrix}.$$

$$a_{33}^{(3)} = |\beta^{(2)}| > 0, \quad H_2 A^{(2)} = \begin{bmatrix} a_{33}^{(3)} & a_{34}^{(3)} \\ 0 & a_{44}^{(3)} \end{bmatrix}$$

$$\diamond \quad S = \begin{bmatrix} I_2 & \\ & H_2 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{1} & \\ & H_1 \end{bmatrix} \cdot H_0$$

则 S 为有限个 \mathbf{H} -矩阵之积, 且有

$$SA = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ & & a_{33}^{(3)} & a_{34}^{(3)} \\ & & & a_{44}^{(3)} \end{bmatrix} \stackrel{\Delta}{=} R$$

[注] 设 $H_l = I_{n-l} - 2u_{n-l}u_{n-l}^T$ ($u_{n-l}^T u_{n-l} = 1$), 则

$$\begin{aligned} \begin{bmatrix} I_l & \\ & H_l \end{bmatrix} &= \begin{bmatrix} I_l & \\ & I_{n-l} \end{bmatrix} - 2 \begin{bmatrix} O & O \\ O & u_{n-l}u_{n-l}^T \end{bmatrix} = I_n - 2 \begin{bmatrix} 0 \\ \vdots \\ 0 \\ u_{n-l} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 & u_{n-l}^T \end{bmatrix} \\ &= I_n - 2u_n u_n^T \quad (u_n^T u_n = u_{n-l}^T u_{n-l} = 1) \end{aligned}$$

故 $\begin{bmatrix} I_l & \\ & H_l \end{bmatrix}$ 是 H-矩阵; 在 Th10 中, 当 $A_{n \times n}$ 不可逆时, 仍可得 $SA = R$,

但 R 是不可逆矩阵.

例 7 用 H-变换求 $A = \begin{bmatrix} 3 & 14 & 9 \\ 6 & 43 & 3 \\ 6 & 22 & 15 \end{bmatrix}$ 的 QR 分解.

解 (1) $\beta^{(0)} = \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix}$, $\beta^{(0)} - |\beta^{(0)}|e_1 = \begin{bmatrix} -6 \\ 6 \\ 6 \end{bmatrix} = 6 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, $u = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

$$H_0 = I - 2uu^T = \frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}, \quad H_0 A = \begin{bmatrix} 9 & 48 & 15 \\ 0 & 9 & -3 \\ 0 & -12 & 9 \end{bmatrix}$$

(2) $A^{(1)} = \begin{bmatrix} 9 & -3 \\ -12 & 9 \end{bmatrix}$, $\beta^{(1)} = \begin{bmatrix} 9 \\ -12 \end{bmatrix}$

$$\beta^{(1)} - |\beta^{(1)}|e_1 = \begin{bmatrix} -6 \\ -12 \end{bmatrix} = (-6) \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad u = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$H_1 = I - 2uu^T = \frac{1}{5} \begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}, \quad H_1 A^{(1)} = \begin{bmatrix} 15 & -9 \\ 0 & -3 \end{bmatrix}$$

$$\diamond \quad S = \begin{bmatrix} 1 & \\ & H_1 \end{bmatrix} H_0$$

$$\text{则} \quad Q = S^{-1} = S^T = H_0 \begin{bmatrix} 1 & \\ & H_1 \end{bmatrix} = \frac{1}{15} \begin{bmatrix} 5 & -2 & -14 \\ 10 & 11 & 2 \\ 10 & -10 & 5 \end{bmatrix}$$

$$R = \begin{bmatrix} 9 & 48 & 15 \\ & 15 & -9 \\ & & -3 \end{bmatrix}: \quad A = QR$$

五、化方阵与 Hessenberg 矩阵相似

上 Hessenberg 矩阵: $F_{\text{上}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ & \ddots & \ddots & \ddots & \vdots \\ & & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ & & & a_{n,n-1} & a_{nn} \end{bmatrix}$

Th11 设 $A \in \mathbb{R}^{n \times n}$, 则存在有限个 G-矩阵之积 Q , 使得 $QAQ^T = F_{\perp}$.

证 (1) 对 A : 若 $\beta^{(0)} = \begin{bmatrix} a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} \neq 0$, 则存在有限个 G-矩阵之积 T_0 , 使得

$$T_0 \beta^{(0)} = |\beta^{(0)}| e_1 \stackrel{\Delta}{=} a_{21}^{(1)} e_1$$

$$\begin{bmatrix} 1 & & \\ & T_0 & \end{bmatrix} A \begin{bmatrix} 1 & & \\ & T_0 & \end{bmatrix}^T = \left[\begin{array}{c|cccc} a_{11} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ \hline a_{21}^{(1)} & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \end{array} \right] \begin{matrix} \\ \\ A^{(1)} \\ \\ \end{matrix}$$

若 $\beta^{(0)} = \theta$, 转入(2);

(2) 对 $A^{(1)}$: 若 $\beta^{(1)} = \begin{bmatrix} a_{32}^{(1)} \\ \vdots \\ a_{n2}^{(1)} \end{bmatrix} \neq 0$, 则存在有限个 G-矩阵之积 T_1 , 使得

$$T_1 \beta^{(1)} = |\beta^{(1)}| e_1 \stackrel{\Delta}{=} a_{32}^{(2)} e_1$$

$$\begin{bmatrix} 1 & \\ & T_1 \end{bmatrix} A^{(1)} \begin{bmatrix} 1 & \\ & T_1 \end{bmatrix}^T = \left[\begin{array}{c|cccc} a_{22}^{(1)} & a_{21}^{(2)} & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \hline a_{32}^{(2)} & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \end{array} \right] A^{(2)}$$

若 $\beta^{(1)} = \theta$, 转入(3);

(3) 对 $A^{(2)}: \dots$, 进行 $n-2$ 步结束.

$$\text{令 } Q = \begin{bmatrix} I_{n-2} & & \\ & \ddots & \\ & & T_{n-3} \end{bmatrix} \dots \begin{bmatrix} I_2 & \\ & T_1 \end{bmatrix} \begin{bmatrix} 1 & \\ & T_0 \end{bmatrix}, \text{ 则 } Q A Q^T = F_{\perp}.$$

定理4.12任何实方正 A 都可以通过初等反射变换正交相似与上Hessenberg矩阵。

证明:类似于定理4.11的证明.

推论: 任何实对称矩阵 A 都可以通过初等旋转变换(或初等反射变换)正交相似于实对称三对角矩阵。

证明:由定理4.12可知:

存在 $Q = H_{n-2} \cdots H_{n_1}$, 使得 $Q A Q^T = F_{\perp}$. 于是有

$$A^T = A \Rightarrow Q A Q^T = (F_{\perp})^T$$

即 $F_{\perp} = (F_{\perp})^T$, 故 F_{\perp} 是“实对称三对角矩阵”。

例 8 用 H-变换化 $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ 正交相似于“三对角矩阵”.

解 $\beta^{(0)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$; $\beta^{(0)} - |\beta^{(0)}| e_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $u = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$$H_0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & & \\ & H_0 & \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$QA = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \quad QAQ^T = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

4.3

满秩分解

复习矩阵的秩的概念和性质

定义:矩阵行或列向量的最大线性无关组的向量个数。

性质1: 矩阵的行秩和列秩相等。

性质2: $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$

定义4.8 设 $A \in C_r^{m \times n}$ ($r > 0$), 如果存在矩阵 $F \in C_r^{m \times r}$ 和 $G \in C_r^{r \times n}$ 使得 $A = FG$, 则称为 A 的满秩分解。

当 A 为行满秩或列满秩矩阵时, A 可分解为一个因子为单位矩阵, 另一个为 A 本身, 称为平凡分解。

应用: 当 r 远小于 m 和 n 时, 利用满秩分解可以去掉 A 中的冗余信息, 节省存储量和运算量。

定理： 设 $A \in C_r^{m \times n}$,则A的满秩分解存在。

证明过程即为构造过程。

证明： 设 $A=(a_1, a_2, \dots, a_n)$, 取A的列向量组的一个极大线性无关组, 不妨设为 $a_{i_1}, a_{i_2}, \dots, a_{i_r}$, 则对于A的任意*i*列 a_i 可以表示为 $a_{i_1}, a_{i_2}, \dots, a_{i_r}$ 的线性组合。即

$$a_i = g_{1i} \cdot a_{i_1} + g_{2i} \cdot a_{i_2} + \dots + g_{ri} \cdot a_{i_r}, \quad i=1, 2, \dots, n$$

写成矩阵形式有

$$A=FG$$

其中 $F=(a_{i_1}, a_{i_2}, \dots, a_{i_r})$, $G=(g_1, g_2, \dots, g_n)$, $g_i=(g_{1i}, g_{2i}, \dots, g_{ri})^T$, 显然F为列满秩的矩阵, G为行满秩的矩阵(为什么?)。

注:满秩分解不唯一。

由于前面的证法可计算性差，我们可以利用如下证法(2)：

$$\text{rank} A = r \Rightarrow A \xrightarrow{\text{行}} \text{阶梯形 } B = \begin{pmatrix} G \\ O \end{pmatrix} : G \in \mathbb{C}_r^{r \times n}$$

$$\Rightarrow \exists \text{有限个初等矩阵之积 } P_{m \times m}, \text{st. } PA = B$$

$$\Rightarrow A = P^{-1}B = \left(F_{m \times r} \mid S_{m \times (m-r)} \right) \begin{pmatrix} G \\ O \end{pmatrix} = FG : F \in \mathbb{C}_r^{m \times r}$$

例 1 $A = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ 2 & 2 & -2 & -1 \end{bmatrix}$, 求 $A = FG$.

解(1) $(A | I) = \left[\begin{array}{cccc|ccc} -1 & 0 & 1 & 2 & 1 & & \\ & 1 & 2 & -1 & & 1 & \\ & 2 & 2 & -2 & & & 1 \end{array} \right] \rightarrow \left[\begin{array}{cccc|ccc} -1 & 0 & 1 & 2 & 1 & 0 & 0 \\ & 0 & 2 & 0 & 3 & 1 & 1 & 0 \\ & 0 & 0 & 0 & 0 & 1 & -1 & 1 \end{array} \right]$

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 1 & 1 \end{bmatrix}, \quad F = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ -2 & 1 \end{bmatrix}$$

$$G = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & 2 & 0 & 3 \end{bmatrix}; \quad \text{满秩分解为 } A = FG.$$

$$\text{解(2)} \quad (A|I) \rightarrow \cdots \rightarrow \left[\begin{array}{cccc|ccc} 1 & 0 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 3/2 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 \end{array} \right]$$

$$A = P^{-1}B = (F|S) \begin{pmatrix} I_2 & B_{12} \\ O & O \end{pmatrix} = (F|FB_{12})$$

$$\text{故 } F = \text{“} A \text{ 的前 2 列”} = \begin{bmatrix} -1 & 0 \\ 1 & 2 \\ 2 & 2 \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 0 & 3/2 \end{bmatrix}.$$

定义4.9 Hermite标准形, $B \in C_r^{m \times n}$ 满足

- 1) B 的前 r 行中每一行至少含一个非零元素, 且第一个非零元素是1, 而后 $m-r$ 行的元素都为0;
- 2) 若 B 中第 i 行的第一个非零元素1在第 j_i 列($i=1,2,\dots,r$), 则 $j_1 < j_2 < \dots < j_r$;
- 3) $Be_{jk} = e_k$ 中的 j_1, j_2, \dots, j_r 列为单位矩阵 I_m 的前 r 列。

[注] 使用初等行变换可将任何非零矩阵化为拟 Hermite 标准形.

矩阵的 Hermite 标准形就是矩阵的行最简形 (初等变换意义下).

例如： $B = \begin{pmatrix} 0 & 1 & \times & \times & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 1 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

这样 $B \in C_2^{3 \times 7}$ 为 Hermite 标准形。

定理： 任意非零矩阵 $A \in C_r^{m \times n}$ 可通过初等行变换

化为Hermite 标准形B, 且B的前r行线性无关。

2. 置换矩阵: 划分单位矩阵 $I_n = (e_1, e_2, \dots, e_n)$, 称 $P_1 = (e_{j_1}, e_{j_2}, \dots, e_{j_n})$

为置换矩阵, 其中 $j_1 j_2 \dots j_n$ 是 $1, 2, \dots, n$ 的一个排列.

特点: 划分 $A_{m \times n} = (a_1, a_2, \dots, a_n)$, 那么 $AP_1 = (a_{j_1}, a_{j_2}, \dots, a_{j_n})$.

定理 4.14 设矩阵 $A \in C_r^{m \times n}$ 的 Hermite 标准形为 B ,

那么, 取 F 为 A 的 j_1, j_2, \dots, j_r 列构成 $m \times r$ 矩阵,

G 为 B 的前 r 行构成的 $r \times n$ 矩阵, $A = FG$.

证明: $A \rightarrow B$, 知存在可逆矩阵 P 使得 $PA = B$

从而 $A = P^{-1}B$

将 P^{-1} 分块为 $P^{-1} = [F, S]$ 可得 $A = FG$, 其中 G 为 B 的前 r 行构成的 $r \times n$ 矩阵。(为什么?)

注意到 B 为 Hermite 标准形, 从而它的后 $n-r$ 行为 0 向量.

作置换矩阵 P_1 使得 $BP_1 = \begin{bmatrix} I_r & B_{12} \\ 0 & 0 \end{bmatrix}$

从而 $AP_1 = P^{-1}BP_1 = [F, S] \begin{bmatrix} I_r & B_{12} \\ 0 & 0 \end{bmatrix} = [F, FB_{12}]$

从而 F 为 A 对应的列。

例 2 $A = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 2 & 0 & 2 & 2 \\ 4 & 2 & 4 & 2 \end{bmatrix}$, 求 $A = FG$.

解 $A \xrightarrow{\text{行}} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 \end{bmatrix} \xrightarrow{\text{行}} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = B$

(1) $c_1 = 1, c_2 = 2$: $F = \text{“} A \text{ 的第 } 1, 2 \text{ 列”} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 2 & 0 \\ 4 & 2 \end{bmatrix}$

$$G = \text{“} B \text{ 的第 } 1, 2 \text{ 行”} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

(2) $c_1 = 3, c_2 = 2$: $F =$ “ A 的第 3, 2 列” (同前)

$G =$ “ B 的第 1, 2 行” (同前)

$$(3) \quad A \xrightarrow{\text{行}} B \xrightarrow{\text{行}} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = C: \quad \text{可取 } c_1 = 1, c_2 = 4$$

$$F = \text{“} A \text{ 的第 1, 4 列”} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 2 \\ 4 & 2 \end{bmatrix}$$

$$G = \text{“} C \text{ 的第 1, 2 行”} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

例3 $A_1 \in C_{r_1}^{m \times n}, A_2 \in C_{r_2}^{m \times n} \Rightarrow \text{rank}(A_1 + A_2) \leq r_1 + r_2.$

证 $r_1 \cdot r_2 = 0$: 成立.

$r_1 \cdot r_2 \neq 0$: $A_1 = F_1 G_1, A_2 = F_2 G_2$ ($F_1 \in C_{r_1}^{m \times r_1}, F_2 \in C_{r_2}^{m \times r_2}$)

$$A_1 + A_2 = (F_1 | F_2) \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} \Rightarrow \text{秩}(A_1 + A_2) \leq \text{秩}(F_1 | F_2) \leq r_1 + r_2$$

例：设矩阵 $A \in C_r^{m \times n}$ ($r > 0$), 则必有分解

$$A=QR.$$

其中, Q 为 $m \times r$ 矩阵, $Q^H Q = I$, 而 R 为 $r \times n$ 矩阵, 它的行线性无关。

证明：作 A 的满秩分解

$$A=FG$$

其中 $F \in C_r^{m \times r}$, $G \in C_r^{r \times n}$. 由矩阵的 QR 分解定理有

$$F=QR_1$$

其中 Q 为 $m \times r$ 矩阵, $Q^H Q = I$, 而 R_1 为 r 阶非奇异矩阵。

于是 $A = FG = QR_1 G = QR.$

这里 $R = R_1 G$, 它的 r 个行线性无关。

定理(习题 4.3,第 4 题). 设 $\mathbf{F} \in C_r^{m \times r}$, $\mathbf{G} \in C_r^{r \times n}$, 则

$$\text{rank}(\mathbf{F}\mathbf{G})=r$$

证明: 显然 $\text{rank}(\mathbf{F}\mathbf{G}) \leq r$ (a)

如果 $\text{rank}(\mathbf{F}^H\mathbf{F}) = \text{rank}(\mathbf{G}\mathbf{G}^H) = r$.

则 $\text{rank}(\mathbf{F}^H\mathbf{F}\mathbf{G}\mathbf{G}^H) = r$, 从而

$$\text{rank}(\mathbf{F}\mathbf{G}) \geq r \quad (\text{b})$$

这样 $\text{rank}(\mathbf{F}\mathbf{G}) = r$.

下面我们证明 $\text{rank}(\mathbf{F}^H\mathbf{F}) = r$ (实际为习题 4.3 第 2 题)

这需要证明 $\mathbf{F}^H\mathbf{F}$ 为满秩矩阵。仅需证明对任意 $\mathbf{x} \in C^r$, 若 $\mathbf{F}^H\mathbf{F}\mathbf{x} = \mathbf{0}$ 则 $\mathbf{x} = \mathbf{0}$ 即可。

事实上, 由 $\mathbf{F}^H\mathbf{F}\mathbf{x} = \mathbf{0}$ 可得 $\mathbf{x}^H\mathbf{F}^H\mathbf{F}\mathbf{x} = 0$,

从而 $\|\mathbf{F}\mathbf{x}\| = 0$, 根据范数定义, $\mathbf{F}\mathbf{x} = \mathbf{0}$;

再由 \mathbf{F} 为列满秩矩阵，可得 $\mathbf{x}=\mathbf{0}$ 。

这样证明了。

思考题：设 $\mathbf{F} \in C_r^{m \times r}$ ， $\mathbf{G} \in C_r^{r \times m}$ ，则 $\text{rank}(\mathbf{GF})=r$ 成立吗？

若是，证明之；若否，研究成立的条件是什么？

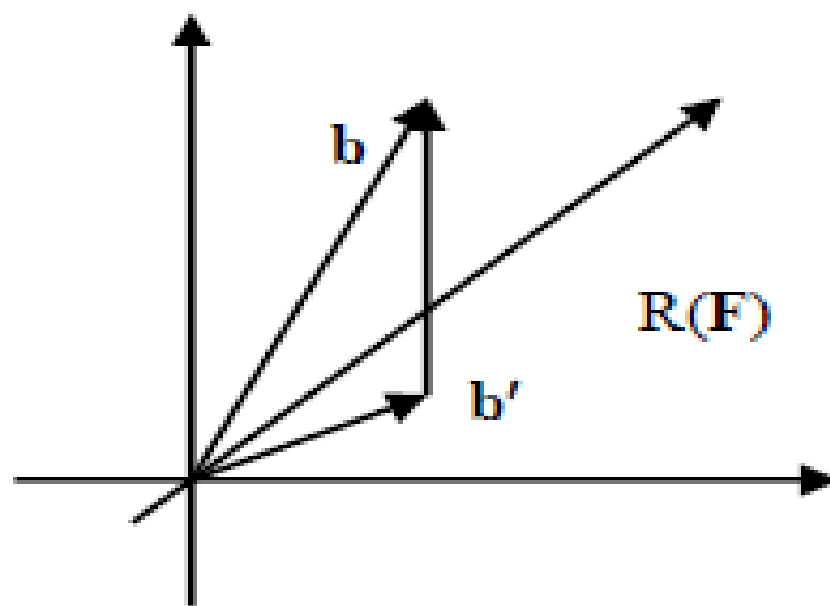
下面利用满秩分解讨论线性方程组求解。

$$\mathbf{Ax} = \mathbf{b},$$

设 \mathbf{A} 的满秩分解为 $\mathbf{A} = \mathbf{FG}$, 方程变为 $\mathbf{FGx} = \mathbf{b}$.

则若 $\mathbf{b} \in \mathbf{R}(\mathbf{F})$, 则存在相容解。

若 $\mathbf{b} \notin \mathbf{R}(\mathbf{F})$, 则不存在相容解, 这时我们可以将 \mathbf{b} 投影到 $\mathbf{R}(\mathbf{F})$ 得到 $\mathbf{b}' = \mathbf{F}(\mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H\mathbf{b}$



这时方程求解的实际就是所谓最小二乘解。

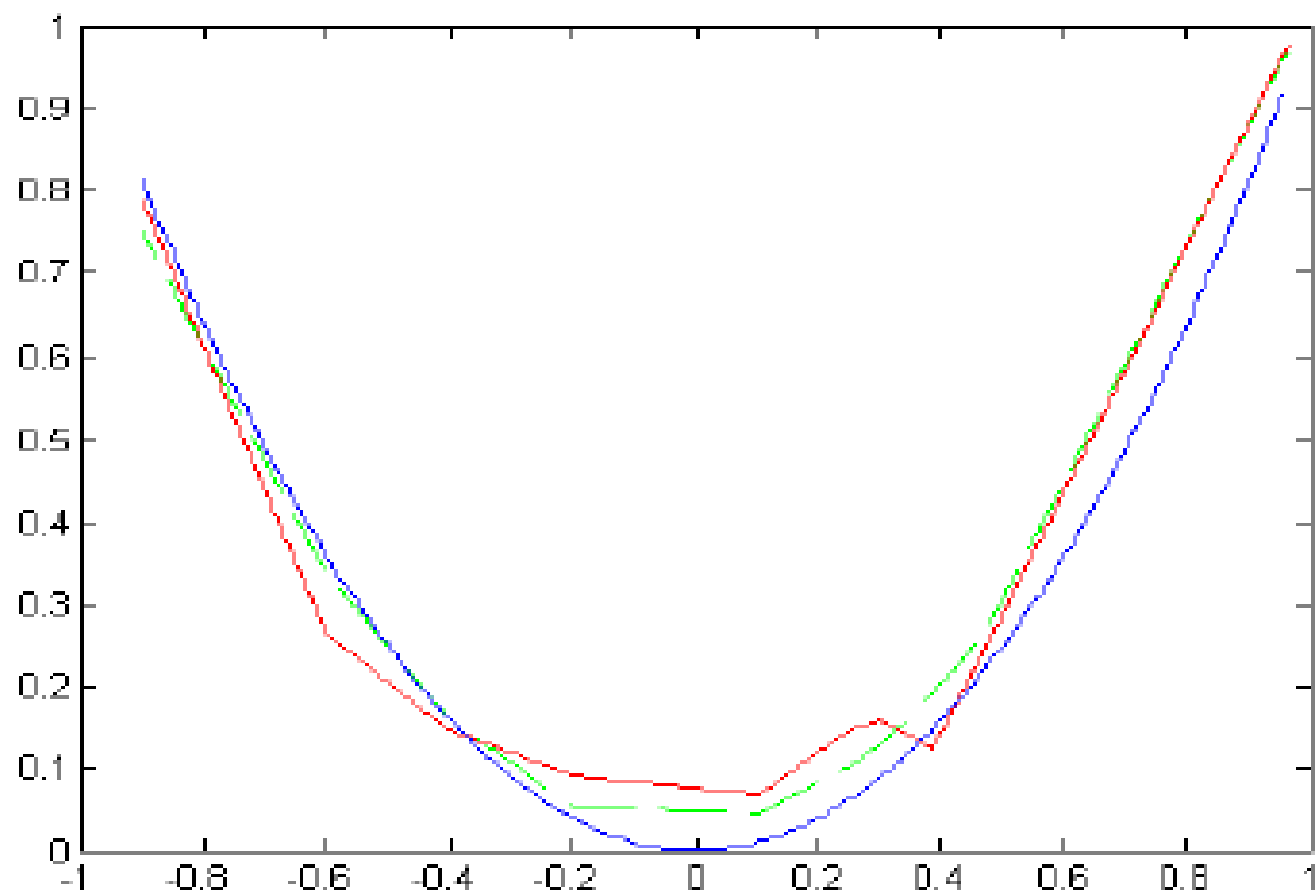
如果这样方程组变为 $\mathbf{F}\mathbf{G}\mathbf{x} = \mathbf{F}(\mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H\mathbf{b}$.

由于 \mathbf{G} 的秩为 r 因此方程组 $\mathbf{G}\mathbf{x} = (\mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H\mathbf{b}$ 肯定有解。如果 \mathbf{G} 不为方阵,

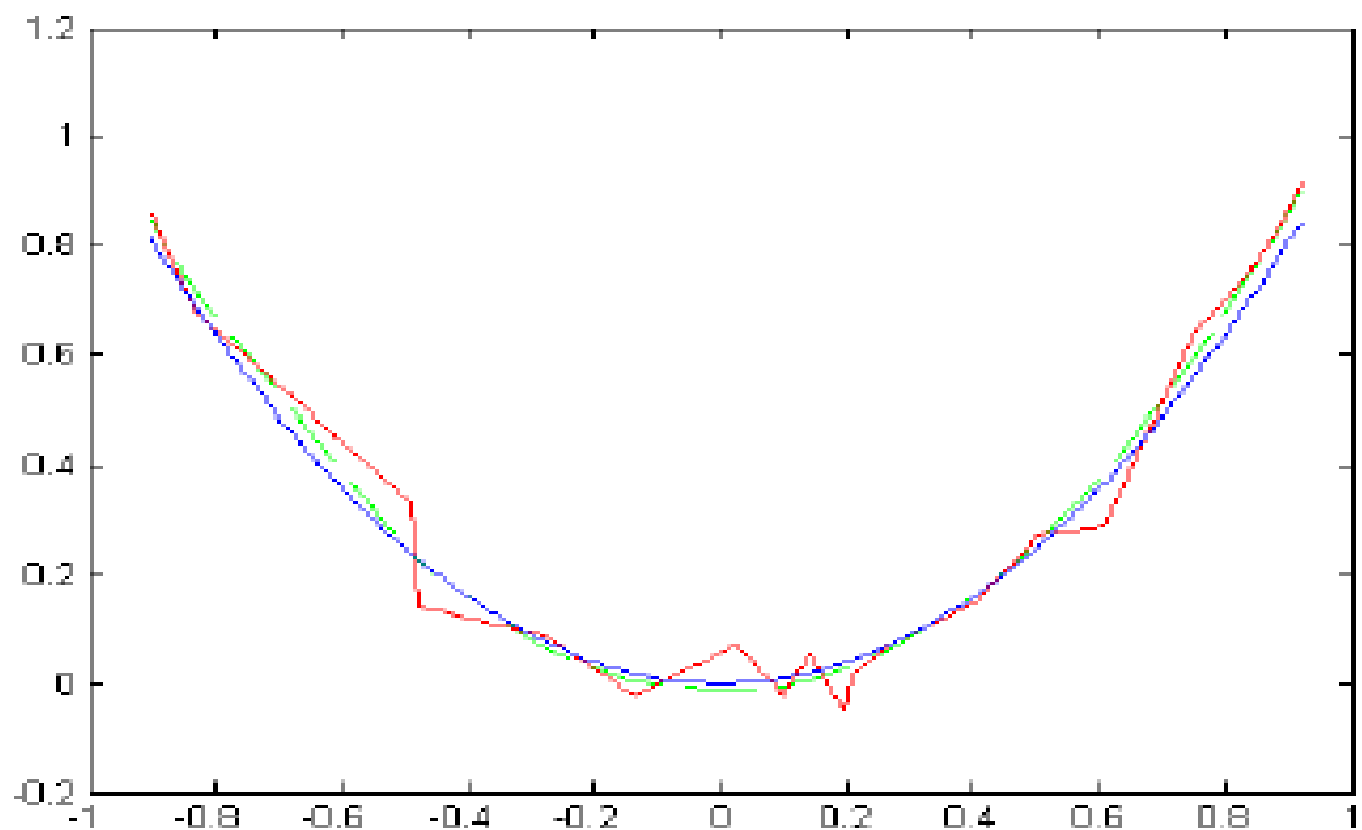
则解不惟一。这时需要求解所谓最小范数解。

即 $\mathbf{x} = \mathbf{G}^H(\mathbf{G}\mathbf{G}^H)^{-1}(\mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H\mathbf{b}$

这就是所谓最小范数最小二乘解。



10 个随机逼近点， $y = x^2 + r$ ，其中 r 为 $[-0.1, 0.1]$ 之间均匀分布的随机数。根据模型 $y = ax^2 + bx + c$ ，利用最小二乘解得到的逼近结果。红线所在点为逼近点
蓝线为 $y = x^2$ ，绿线为求得的逼近解。



20 个随机逼近点， $y = x^2 + r$ ，其中 r 为 $[-0.1, 0.1]$ 之间均匀分布的随机数。根据模型 $y = ax^2 + bx + c$ ，利用最小二乘解得到的逼近结果。红线所在点为逼近点，蓝线为 $y = x^2$ ，绿线为求得的逼近解。

4.4 奇异值分解

1. 厄米矩阵的谱分解
2. 非奇异矩阵的西对角分解
3. 一般矩阵的奇异值分解
4. 正交相抵
5. 矩阵奇异值在矩阵重建中的应用

0. 预备知识:

(1) $\forall A_{m \times n}, (A^H A)_{n \times n}$ 是 Hermite (半) 正定矩阵.

$$\forall x \neq 0, x^H A^H A x = (Ax)^H (Ax) = |Ax|^2 \geq 0$$

(2) 齐次方程组 $Ax = 0$ 与 $A^H Ax = 0$ 同解.

若 $Ax = 0$, 则 $A^H Ax = 0$;

反之, $A^H Ax = 0 \Rightarrow |Ax|^2 = (Ax)^H (Ax) = x^H (A^H Ax) = 0 \Rightarrow Ax = 0$

(3) $\text{rank } A = \text{rank}(A^H A)$

$$S_1 = \{x \mid Ax = 0\}, \quad S_2 = \{x \mid A^H Ax = 0\}$$

$$S_1 = S_2 \Rightarrow \dim S_1 = \dim S_2 \Rightarrow n - r_A = n - r_{A^H A} \Rightarrow r_A = r_{A^H A}$$

(4) $A = O_{m \times n} \Leftrightarrow A^H A = O_{n \times n}$

必要性. 左乘 A^H 即得; 充分性. $r_A = r_{A^H A} = 0 \Rightarrow A = O$.

1. 厄米矩阵的谱分解

A 为厄米矩阵, 则存在酉矩阵 **U**, 使

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \begin{bmatrix} \lambda_1 & & & \mathbf{0} \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_n \end{bmatrix} = \mathbf{\Lambda}$$

将 **U** 写成列向量形式, 即 $\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_n]$, 则

$$\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \mathbf{0} \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^H \\ \mathbf{u}_2^H \\ \vdots \\ \mathbf{u}_n^H \end{bmatrix} = \sum_{i=1}^n \lambda_i \mathbf{u}_i \mathbf{u}_i^H$$

由此可见：

我们知道实对称矩阵 A 或 Hermite 矩阵 A 都可正交相似于对角矩阵，即

$$A=Q^H D Q$$

由定理 4.12 任何实方阵 A 都可以正交相似于上 Hessenberg 矩阵。

即
$$A=Q^T H Q$$

进一步的思考可以发现，如果要进一步使用左乘和右乘以正交(酉)矩阵化简 A ，只有放弃这两个矩阵是相等的约束，这就是下面的定理。

2. 非奇异矩阵的西对角分解

定理：设 A 为 n 阶非奇异矩阵，则存在 n 阶酉矩阵 U 及 V ，使得

$$U^H A V = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{bmatrix}, \quad \sigma_i > 0 (i = 1, 2, \dots, n)$$

(若将 U, V 写成

$$U = [u_1 \quad u_2 \quad \dots \quad u_n], V = [v_1 \quad v_2 \quad \dots \quad v_n], \quad \text{则 } A = \sum_{i=1}^n \sigma_i u_i v_i^H)$$

证： $A^H A$ 也为 n 阶非奇异矩阵，而且是厄米，正定矩阵，故存在 n 阶

$$\text{酉矩阵 } V, \text{ 使 } V^H (A^H A) V = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_n^2 \end{bmatrix} \quad \sigma_i^2 \text{ 为 } A^H A \text{ 的特}$$

征值。

$$\text{令 } \Sigma = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{bmatrix}, \text{ 则 } V^H A^H A V = \Sigma^2$$

$$\text{令 } U^H = \Sigma^{-1} V^H A^H, \rightarrow U = A V \Sigma^{-1}, \text{ 则}$$

$$U^H U = \Sigma^{-1} (V^H A^H A V) \Sigma^{-1} = I_n$$

$$\text{即 } U \text{ 也是酉矩阵, 而且 } U^H A V = \Sigma^{-1} V^H A^H A V = \Sigma \quad \text{证毕}$$

酉对角分解的求法正如证明中所给: 先对 $A^H A$ 对角化 (酉对角化), 求出变换矩阵 V , 再令 $U = A V \Sigma^{-1}$ 即可。

3. 一般矩阵的奇异值分解

定理：设 $A \in \mathbb{C}_r^{m \times n}$ ，则存在 m 阶酉矩阵 U 及 n 阶酉矩阵 V ，使

$$U^H A V = \begin{bmatrix} \sigma_1 & & & O \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r & \\ O & & & & O \end{bmatrix} \quad \begin{matrix} r \text{ 行} \\ \\ \\ \\ (m-r) \text{ 行} \end{matrix} \quad \begin{matrix} \\ \\ \\ \\ (n-r) \text{ 列} \end{matrix} \quad \text{即}$$

$$A = U \begin{bmatrix} \sigma_1 & & & O \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r & \\ O & & & & O \end{bmatrix} V^H$$

证 对于 Hermite 半正定矩阵 $A^H A$, \exists 酉矩阵 $V_{n \times n}$, 使得

$$V^H (A^H A) V = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \begin{bmatrix} \Sigma^2 & O \\ O & O \end{bmatrix}_{n \times n}$$

划分 $V = (V_1 | V_2)$: V_1 是 V 的前 r 列, V_2 是 V 的后 $n-r$ 列.

$$A^H A V = V \begin{bmatrix} \Sigma^2 & O \\ O & O \end{bmatrix}_{n \times n} : (A^H A V_1 | A^H A V_2) = (V_1 \Sigma^2 | O)$$

$$(1) \quad A^H A V_1 = V_1 \Sigma^2 ;$$

$$V_1^H A^H A V_1 = \Sigma^2 \Rightarrow (A V_1 \Sigma^{-1})^H (A V_1 \Sigma^{-1}) = I_r$$

$$(2) \quad A^H AV_2 = O :$$

$$V_2^H A^H AV_2 = O \Rightarrow (AV_2)^H (AV_2) = O \Rightarrow AV_2 = O$$

$$\text{令 } U_1 \stackrel{\Delta}{=} AV_1 \Sigma^{-1} : \quad U_1^H U_1 = I_r$$

设 U_1 的列为 u_1, \cdots, u_r , 扩充为 \mathbb{C}^m 的标准正交基 $u_1, \cdots, u_r, u_{r+1}, \cdots, u_m$,

则 $U_2 = (u_{r+1}, \cdots, u_m)$ 满足 $U_2^H U_1 = O_{(m-r) \times r}$. 记 $U = (U_1 | U_2)$, 则有

$$U^H AV = U^H (AV_1 | AV_2) = \begin{pmatrix} U_1^H \\ U_2^H \end{pmatrix} (U_1 \Sigma | O)$$

$$= \begin{bmatrix} U_1^H U_1 \Sigma & O \\ U_2^H U_1 \Sigma & O \end{bmatrix} = \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix}_{m \times n}$$

[注] 称 $A = UDV^H$ 为 A 的奇异值分解.

- (1) U 与 V 不唯一;
- (2) U 的列为 AA^H 的特征向量, V 的列为 A^HA 的特征向量;
- (3) 称 U 的列为 A 的左奇异向量, 称 V 的列为 A 的右奇异向量.

例 1 $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, 求 $A = UDV^T$.

解 $A^T A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} = B, \quad |\lambda I - B| = \lambda(\lambda - 1)(\lambda - 3)$

$$\lambda_1 = 3: \quad 3I - B = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 1 \end{bmatrix}, \quad \xi_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

$$\lambda_2 = 1: \quad 1I - B = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix}, \quad \xi_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$\lambda_3 = 0: \quad 0I - B = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -1 & -2 \end{bmatrix}, \quad \xi_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$r_A = 2: \quad \Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}$$

$$V = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \end{bmatrix}, \quad V_1 = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{bmatrix}$$

$$U_1 = AV_1\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 \end{bmatrix}, \quad \text{取 } U_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \text{则 } U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$U^T AV = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = D, \quad A = UDV^T$$

奇异值分解的最佳逼近性

定理：考虑矩阵 A 的秩 k 近似，并将其记作 A_k ，其中， $k < r = \text{rank}(A)$ 。矩阵 A_k 定义如下：

$$A_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^H, \quad k < r$$

则 A 与秩为 k 的任一矩阵 B 之差的 l_1 和 Frobenius 范数分别为

$$\min_{\text{rank}(B)=k} \|A - B\|_1 = \|A - A_k\|_1 = \sigma_{k+1}$$

和

$$\min_{\text{rank}(B)=k} \|A - B\|_F^2 = \|A - A_k\|_F^2 = \sigma_{k+1}^2 + \sigma_{k+2}^2 + \cdots + \sigma_r^2$$

定理: 在 $A \in \mathbb{C}_r^{m \times n}$ ($r > 0$) 的奇异值分解 $A = U \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix} V^H$ 中, 划分

$U = (u_1, u_2, \dots, u_m), V = (v_1, v_2, \dots, v_n)$, 则有

$$(1) \quad N(A) = \text{span}\{v_{r+1}, v_{r+2}, \dots, v_n\};$$

$$(2) \quad R(A) = \text{span}\{u_1, u_2, \dots, u_r\};$$

$$(3) \quad A = \sigma_1 u_1 v_1^H + \sigma_2 u_2 v_2^H + \dots + \sigma_r u_r v_r^H.$$

证 $A = (U_1 | U_2) \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} \begin{pmatrix} V_1^H \\ V_2^H \end{pmatrix} = U_1 \Sigma V_1^H$

容易验证: $U_1 \Sigma V_1^H x = 0 \Leftrightarrow V_1^H x = 0$

$$\begin{aligned} (1) \quad N(A) &= \{ x \mid Ax = 0 \} = \{ x \mid U_1 \Sigma V_1^H x = 0 \} \\ &= \{ x \mid V_1^H x = 0 \} = N(V_1^H) = R^\perp(V_1) \\ &= R(V_2) = \text{span}\{ v_{r+1}, \dots, v_n \} \end{aligned}$$

$$\begin{aligned}
 (2) \quad R(A) &= \{ y \mid y = Ax \} = \{ y \mid y = U_1(\Sigma V_1^H x) \} \\
 &\subset \{ y \mid y = U_1 z \} = R(U_1)
 \end{aligned}$$

$$\begin{aligned}
 R(U_1) &= \{ y \mid y = U_1 z \} = \{ y \mid y = A(V_1 \Sigma^{-1} z) \} \\
 &\subset \{ y \mid y = Ax \} = R(A)
 \end{aligned}$$

$$\text{故 } R(A) = R(U_1) = \text{span}\{u_1, \cdots, u_r\}$$

$$(3) \quad A = (u_1, \cdots, u_r) \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^H \\ \vdots \\ v_r^H \end{bmatrix}$$

$$= \sigma_1 u_1 v_1^H + \cdots + \sigma_r u_r v_r^H \quad (\text{比较矩阵的谱分解})$$

四、正交相抵

$A_{m \times n}, B_{m \times n}$, 若有酉矩阵 $U_{m \times m}$ 及 $V_{n \times n}$ 使 $U^H A V = B$, 称 A 与 B 正交相抵.

性质: (1) A 与 A 正交相抵;

(2) A 与 B 正交相抵 $\Rightarrow B$ 与 A 正交相抵;

(3) A 与 B 正交相抵, B 与 C 正交相抵 $\Rightarrow A$ 与 C 正交相抵.

定理4.18 A 与 B 正交相抵 $\Rightarrow \sigma_A = \sigma_B$.

$$\begin{aligned} \text{证 } B = U^H A V &\Rightarrow B^H B = \cdots = V^{-1} (A^H A) V \\ &\Rightarrow \lambda_{B^H B} = \lambda_{A^H A} \geq 0 \Rightarrow \sigma_B = \sigma_A \end{aligned}$$

例 2 $A^H = A \Rightarrow \sigma_A = |\lambda_A| \quad (\because \lambda_{A^H A} = \lambda_{A^2} = (\lambda_A)^2)$

$$A^H = -A \Rightarrow \sigma_A = |\lambda_A| \quad (\because \lambda_{A^H A} = \lambda_{(jA)^2} = (j\lambda_A)^2)$$

($A^H = -A \Rightarrow \lambda_A$ 为 0 或纯虚数, $j\lambda_A$ 为实数)

5.矩阵奇异值在矩阵重建中的应用

- **矩阵重建是信号处理、人工智能和优化领域最近研究的热点。基于凸优化的矩阵重建问题衍生于近几年非常流行的压缩感知技术，主要分为矩阵填充和矩阵恢复问题，是一种重要的数据分析工具，在图像处理、计算机视觉、文本分析、推荐系统等方面已经找到重要的应用。**



- **人工智能分为强人工智能与弱人工智能。强人工智能观点认为有可能制造出真正能推理和解决问题的智能机器，并且这些机器将被认为是有知觉的、有自我意识的。弱人工智能观点认为不可能制造出能真正地推理和解决问题的智能机器，这些机器只不过是智能的，但是并不真正拥有智能，也不会有自主意识。**



- **压缩感知的概念：**
- **将未知的要获得的信号记为AK，它是一个波形。我们希望它越不连续越好，越扩散越好，而我所要做的是按照一定的顺序获得图像信号的信息。我们按照高斯分布来收集数据，而不是线性的，所以我们只会有少量的感测次数，而这些数据 的相关性非常低。**

- **压缩的意思，是指比较先前及当前数据的差异，并记录下这些差异而减少需储存的资料量。对于压缩感知的问题比压缩难多了，像是我们要收集怎样分布的数据、如何收集、保留一些什么系数，从而得到最佳的结果。我们会得到一些机变，保留最大的，最有意义的系数在里头。我们可以做一些抽样，把最重要的保留下来。一个信号，我们知道它是非常好的，但这个信号完全是稀疏的，可能并不是我们要损失掉的。**

- 当感兴趣的信号是可压缩的或者可稀疏表示的，那么我们可以通过极少的采样精确地获得该信号。压缩感知中，信号的获取并不是直接测量信号本身，而是采样测量信号一个感知矩阵相乘后的信号。

- 矩阵重建分为矩阵填充（ Matrix Completion ）和矩阵恢复（ Matrix Recovery ）两大类。前者主要研究如何在数据不完整的情况下将缺失数据进行填充，后者主要研究在某些数据受到严重损坏的情况下恢复出准确的矩阵。
- 研究主要集中在矩阵重建在何种情况下可以准确地实现、有没有快速的算法解决矩阵重建问题和矩阵重建的应用。

- **矩阵填充**
- **对于某个矩阵，我们只能采样得到矩阵的一部分元素，其它一部分或者大部分元素由于各种原因丢失了或无法得到，假设这个矩阵是有信息冗余的，比如是低秩的，也就是说其数据分布在一个低维的线性子空间上。**

- 可以通过如下优化问题来实现矩阵填充：

$$\min \text{rank}(X);$$

$$\text{subject to } X_{ij} = M_{ij}; (i, j) \in \Omega;$$

- 其中 Ω 是已知元素下标的集合。将空缺的元素填充之后使得矩阵的结构尽可能好，即秩尽可能低。

- 一个矩阵的秩 r 与它的非零奇异值的个数相同。于是有一个选择是用矩阵的奇异值的和，即核范数，来近似地替代矩阵的秩：

$$\min \|X\|_*;$$

$$\text{subject to } X_{ij} = M_{ij}; (i, j) \in \Omega;$$

- 矩阵填充的应用举例
- 矩阵填充的一个著名应用是Netflix推荐系统。比赛要求参赛者预测Netflix客户分别喜欢什么影片，要把预测的效率相对原推荐系统Cinematch提高10%以上。这是一个典型的矩阵填充问题，即矩阵的每一行对应某个用户对电影的评级，每一列表示某电影在所有用户中的评级，但是每个用户只可能对一部分电影进行评价，所以我们可以通过矩阵填充得出用户对每部电影的喜好程度。

- 矩阵恢复
- 当矩阵的某些元素被严重破坏后，自动识别出被破坏的元素，恢复出原矩阵。假定原矩阵有非常良好的结构，即是低秩的；另外，假定只有很少一部分元素被严重破坏，即噪声是稀疏的但大小可以任意。

$$\min \text{rank}(\mathbf{A}) + \|\mathbf{E}\|_0;$$

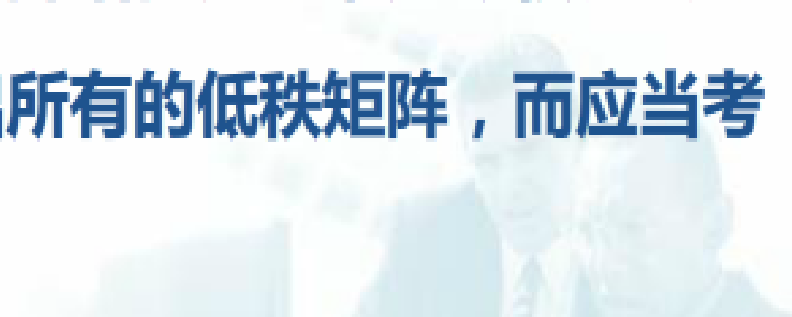
$$\text{subject to } \mathbf{A} + \mathbf{E} = \mathbf{D};$$

其中目标函数为矩阵A的秩以及噪声矩阵E的零范数，即E的非零元素的个数，表明噪声所占的权重。

- 矩阵恢复在背景建模、人脸图像处理等问题中的应用。其中，背景建模利用图片帧与帧之间的相似性，将每帧作为一行排列成一个矩阵，该矩阵理应具备相对较低的秩。于是利用稀疏与低秩矩阵分解技术可以将每帧中间相似的部分和特有的部分分开，即将背景与前景分离。



- 一、矩阵填充理论研究主要考虑的是矩阵填充的可行性，即究竟在什么情况下可以精确无误地把缺失数据填充完整。首先，我们需要理解的是，一个杂乱无章的矩阵是不可能进行重建的，所以我们假定矩阵结构良好，如低秩。我们不能指望通过采样可以重建出所有的低秩矩阵，而应当考虑有多大的概率可以重建。



- **二、并不是用所有的采样方式采样得到的矩阵均能够重建。**
例如，当矩阵的某一行完全没有被采样时，任何方法都不可能将这一行的元素准确地填充。于是，一般都假定矩阵的采样方式也是均匀采样，在这样的情况下来考虑矩阵精确重建的概率。
- **三、采样的矩阵元素数目必须大于一定范围时，才有可能将矩阵进行精确填充。**



- 一种简单的一阶方法，奇异值阈值（Singular Value Thresholding，简称SVT）算法，来求解矩阵填充问题。矩阵填充问题可以写为：

$$\min \|X\|_*;$$

$$\text{subject to } P_{\Omega}(X) = P_{\Omega}(M);$$

其中 Ω 表示所有的采样元素的坐标 $(i; j)$ 的集合， $P_{\Omega}(X)$ 表示一种投影算子，它将矩阵在 Ω 以外的元素置0， Ω 内部元素保持不变。

- SVT算法可以理解作为一种拉格朗日乘子法。首先，它求解的是原问题的一个近似问题：

$$\begin{aligned} \min \quad & \tau \|X\|_* + \frac{1}{2} \|X\|_F^2, \\ \text{subject to} \quad & P_\Omega(X) = P_\Omega(M), \end{aligned}$$

其拉格朗日乘子为

$$L(X, Y) = \tau \|X\|_* + \frac{1}{2} \|X\|_F^2 + \langle Y, P_\Omega(M - X) \rangle,$$

其中 $\langle A, B \rangle = \text{trace}(AB^T)$ 。然后每一步通过最小化这个拉格朗日函数，最终收敛到最优解。迭代序列为：

$$\begin{cases} X_k = \arg \min_X L(X, Y^{k-1}), \\ Y_k = Y_{k-1} + \delta_k P_\Omega(M - X_k). \end{cases}$$

由于

$$L(X, Y) = \tau \|X\|_* + \frac{1}{2} \|X - P_\Omega(Y)\|_F^2 + \langle P_\Omega(Y), M \rangle - \frac{1}{2} \|P_\Omega(Y)\|_F^2,$$

并且

$$\arg \min_X \left[\tau \|X\|_* + \frac{1}{2} \|X - P_\Omega(Y)\|_F^2 \right] = \mathcal{D}_\tau(P_\Omega(Y)),$$

其中

$$\mathcal{D}_\tau(X) := U \mathcal{S}_\tau(\Sigma) V^T, \text{ 如果 } X = U \Sigma V^T.$$

而 \mathcal{S}_τ 为收缩算子:

$$\mathcal{S}_\tau(x) = \begin{cases} x - \tau, & \text{如果 } x > \tau, \\ x + \tau, & \text{如果 } x < -\tau, \\ 0, & \text{其它.} \end{cases}$$

▪ Algorithm 1 (矩阵填充的SVT算法)

Algorithm 1 (矩阵填充的SVT算法)

1: 初始化 $Y_0 = 0$,
2: while not converged do
3: $X_k = D_r(Y^{k-1})$,
4: $Y_k = Y_{k-1} + \delta_k P_\Omega(M - X_k)$.
5: end while

- 序列 X_k 始终低秩，可以表示成两个瘦矩阵 U_k 与 V_k^T 之积，而 Y_k 始终是稀疏的，由于这些性质，该算法执行过程内存需求大大降低，因此适合大规模矩阵的计算。

APG (Accelerated Proximal Gradient) 算法[1]是一种利用Nesterov技巧的一阶算法, 其收敛速度很有竞争力。[42]中将矩阵填充问题转化为一个与原问题近似的无约束优化问题:

$$\min F(X) = \frac{1}{2} \|P_{\Omega}(X - M)\|_F^2 + \mu \|X\|_*, \quad (2.6)$$

APG算法每次最小化上述函数在某个点Y处的一个二阶近似:

$$Q(X, Y) = \frac{1}{2} \|P_{\Omega}(X - M)\|_F^2 + \langle P_{\Omega}(Y), X - Y \rangle + \frac{L_f}{2} \|X - Y\|_F^2 + \mu \|X\|_*, \quad (2.7)$$

其中 L_f 为 ∇f 的Lipschitz常数:

$$\|\nabla f(X_1) - \nabla f(X_2)\| \leq L_f \|X_1 - X_2\|. \quad (2.8)$$

由于

$$\arg \min_X Q(X, Y) = \arg \min_X \frac{L_f}{2} \left\| X - Y + \frac{1}{L_f} P_{\Omega}(Y) \right\|_F^2 + \mu \|X\|_*, \quad (2.9)$$

于是每次迭代:

$$X_{k+1} = \arg \min_X Q(X, Y_k) = \mathcal{D}_{\frac{\mu}{L_f}} \left(Y_k - \frac{1}{L_f} P_{\Omega}(Y_k) \right). \quad (2.10)$$

通常 Y_k 取以下序列时, 算法收敛速度可以达到 $O(k^{-2})$:

$$Y_k = X_k + \frac{t_{k-1} - 1}{t_k} (X_k - X_{k-1}), \quad t_{k+1}^2 - t_{k+1} \leq t_k^2. \quad (2.11)$$

由上面的推导, 矩阵填充的APG算法如算法2所示。

[42]中的实验表明, APG算法采用了continuation技术和line search技术后, 收敛速度比SVT等算法可以提高不少。

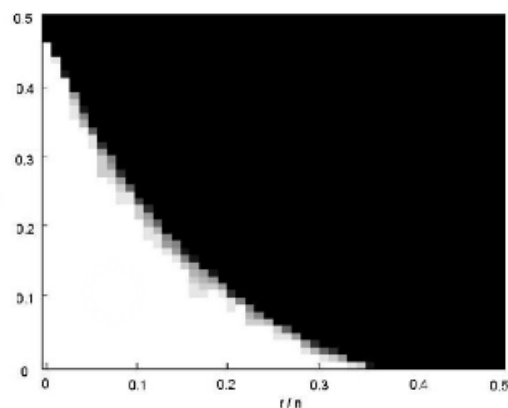
- APG (Accelerated Proximal Gradient) 算法是一种利用Nesterov技巧的一阶算法，其收敛速度很有竞争力。将矩阵填充问题转化为一个与原问题近似的无约束优化问题：

Algorithm 2 (矩阵填充的APG算法)

- 1: 初始化 X_0, X_{-1} ,
 - 2: while not converged do
 - 3: $Y_k = X_k + \frac{t_{k-1}-1}{t_k}(X_k - X_{k-1})$,
 - 4: $X_{k+1} = D_{\mu/L_f}(Y_k - \frac{1}{L_f}P_{\Omega}(Y_k))$,
 - 5: $t_{k+1} = \frac{1+\sqrt{t_k^2+1}}{2}, k = k + 1$.
 - 6: end while
-



- 矩阵恢复的可行性



- 横坐标为 r/n ，纵坐标为损坏元素所占比率。白色表示精确恢复，黑色表示不能恢复，灰色表示一定概率恢复。从图中可以看出，当矩阵的秩越低，被损坏的元素数目越少时，矩阵越容易精确恢复。

IT算法即迭代阈值 (Iterative Thresholding) 算法, 与矩阵填充中的SVT算法类似, 是求解矩阵恢复最早的算法[19]。它将原问题转化为如下近似问题:

$$\begin{aligned} \min_{A,E} \quad & \|A\|_* + \lambda \|E\|_1 + \frac{1}{2\tau} \|A\|_F^2 + \frac{1}{2\tau} \|E\|_F^2, \\ \text{subject to} \quad & A + E = D, \end{aligned} \quad (2.14)$$

其中 τ 取一个很大的正数, 以使此问题的解近似于原问题。其拉格朗日函数为:

$$L(A, E, Y) = \|A\|_* + \lambda \|E\|_1 + \frac{1}{2\tau} \|A\|_F^2 + \frac{1}{2\tau} \|E\|_F^2 + \frac{1}{\tau} \langle Y, D - A - E \rangle. \quad (2.15)$$

于是

$$\arg \min_{A,E} L(A, E, Y) = \arg \min_{A,E} \|A\|_* + \frac{1}{2\tau} \|A - Y\|_F^2 + \lambda \|E\|_1 + \frac{1}{2\tau} \|E - Y\|_F^2. \quad (2.16)$$

Algorithm 3 (矩阵恢复的迭代阈值法)

- 1: Input D, λ, τ .
 - 2: **while** not converged **do**
 - 3: $(U, S, V) = \text{svd}(Y_{k-1})$,
 - 4: $A_k = US_\tau[S]V^T$,
 - 5: $E_k = \mathcal{S}_{\lambda\tau}[Y_{k-1}]$,
 - 6: $Y_k = Y_{k-1} + \delta_k(D - A_k - E_k)$.
 - 7: **end while**
 - 8: Output A_k, E_k
-

类似于矩阵填充，矩阵恢复问题可用以下无约束问题来近似：

$$\min F(A, E) = \frac{1}{2} \|D - A - E\|_F^2 + \mu(\|A\|_* + \lambda\|E\|_1), \quad (2.17)$$

接下来最小化上述函数在某个点 (Y^A, Y^E) 处的二阶近似：

$$\begin{aligned} Q(X^A, X^E, Y^A, Y^E) &= \frac{1}{2} \|D - Y^A - Y^E\|_F^2 + \langle Y^A + Y^E - D, X^A - Y^A + X^E - Y^E \rangle \\ &+ \frac{L_f}{2} (\|X^A - Y^A\|_F^2 + \|X^E - Y^E\|_F^2) + \mu(\|X^A\|_* + \lambda\|X^E\|_1). \end{aligned} \quad (2.18)$$

由于

$$\begin{aligned} \arg \min_{X^A, X^E} Q(X^A, X^E, Y^A, Y^E) &= \\ \arg \min_{X^A, X^E} \frac{L_f}{2} \left\| X^A - Y^A + \frac{1}{L_f} (Y^A + Y^E - D) \right\|_F^2 &+ \mu\|X^A\|_* \\ + \frac{L_f}{2} \left\| X^E - Y^E + \frac{1}{L_f} (Y^A + Y^E - D) \right\|_F^2 &+ \mu\lambda\|X^E\|_1, \end{aligned} \quad (2.19)$$

于是得到如下迭代:

$$\begin{aligned}X_{k+1}^A &= S_{\mu/L_f} \left(Y_k^A - \frac{1}{L_f} (Y_k^A + Y_k^E - D) \right), \\X_{k+1}^E &= S_{\mu\lambda/L_f} \left(Y_k^E - \frac{1}{L_f} (Y_k^E + Y_k^E - D) \right).\end{aligned}$$

通常 Y_k 取以下序列时算法的收敛较快:

$$\begin{aligned}Y_k^A &= X_k^A + \frac{t_{k-1} - 1}{t_k} (X_k^A - X_{k-1}^A), \\Y_k^E &= Y_k^E + \frac{t_{k-1} - 1}{t_k} (X_k^E - X_{k-1}^E), \\t_{k+1}^2 - t_{k+1} &\leq t_k^2.\end{aligned}$$

Algorithm 4 (矩阵恢复的APG算法)

- 1: Input D, λ .
 - 2: $A_0 = A_{-1} = 0; E_0 = E_{-1} = 0; t_0 = t_{-1} = 1; \bar{\mu} > 0; 0 < \eta < 1$.
 - 3: while not converged do
 - 4: $Y_k^A = A_k + \frac{t_{k-1}-1}{t_k} (A_k - A_{k-1}), Y_k^E = E_k + \frac{t_{k-1}-1}{t_k} (E_k - E_{k-1}),$
 - 5: $G_k^A = Y_k^A - \frac{1}{2} (Y_k^A + Y_k^E - D),$
 - 6: $(U, S, V) = \text{svd}(G_k^A), A_{k+1} = US_{\frac{\mu_k}{2}}[S]V^T,$
 - 7: $G_k^E = Y_k^E - \frac{1}{2} (Y_k^A + Y_k^E - D),$
 - 8: $E_{k+1} = S_{\frac{\lambda\mu_k}{2}}[G_k^E],$
 - 9: $t_{k+1} = \frac{1+\sqrt{4t_k^2+1}}{2}, \mu_{k+1} = \max(\eta \mu_k, \bar{\mu}),$
 - 10: $k \leftarrow k + 1.$
 - 11: end while
 - 12: Output $A_k, E_k.$
-

- 矩阵重建是一种重要的数据分析工具，已经在图像处理、计算机视觉、推荐系统等领域找到了不少应用。随着其理论上不断完善，算法上不断优化，以及并行和分布式计算的不断普及，矩阵重建将会进一步在未来的科学研究和工程实践中找到更多应用。



[19]Amir Beck and MarcTeboulle, A Fast Iterative Shrinkage- Thresholding Algorithm for Linear Inverse Problems, SIAMJ.IMAGINGSCIENCESc. 2009, Society for Industrial and Applied Mathematics Vol.2,No.1,pp.183–202

[41]JF Cai, EJ Candès, Z Shen, A Singular Value Thresholding Algorithm for Matrix Completion, 《Siam Journal on Optimization》, 2010, 20(4):1956-1982

[42]Kim-Chuan Toh, Sangwoon Yun, An accelerated proximal gradient algorithm for nuclear norm regularized least squares problems , 《Pacific Journal of Optimization》, 2010, 6(3):615-640