高等工程數學 (10)

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黄 煒
博士，副教授
Dr. Wei Huang

方阵函数与函数矩阵

矩阵论（10）

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黄 炜
博士·副教授
Dr. Wei Huang

4 方阵函数与函数矩阵

我们已经提到过方阵 $\boldsymbol{A}$ 的多项式，它是 $\boldsymbol{A}$ 的一种特殊的函数，为了得到一般的函数，就要考虑高等数学中通过幂级数定义函数的方法。由于幂级数 $\sum\limits_{k=0}^{\infty} c_k\lambda^k$ 在其收敛域内有和函数 $f(\lambda)$，即在收敛域内，有

$$\sum_{k=0}^{\infty} c_k\lambda^k = f(\lambda),$$

所以自然会考虑定义

$$f(\boldsymbol{A}) = \sum_{k=0}^{\infty} c_k\boldsymbol{A}^k.$$

问题是这有没有意义，从而涉及方阵幂级数的敛散性问题。

本章从讨论一般的矩阵序列和矩阵级数的敛散性问题出发，进而给出方阵幂级数收敛与发散的条件及方阵函数的定义和计算，并举例说明其应用。

4.1 矩阵序列与矩阵级数

设 $\boldsymbol{A}_0,\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k,\cdots$ 是线性空间 $\mathbf{C}^{m\times n}$ 中的无穷序列，记为 $\{\boldsymbol{A}_k\}_{k=0}^{\infty}$，其中 $\boldsymbol{A}_k=[a_{ij}^{(k)}]$.

定义 4.1-1 设 $\{\boldsymbol{A}_k\}_{k=0}^{\infty}$ 是一个 $m\times n$ 矩阵序列，如果存在 $m\times n$ 矩阵 $\boldsymbol{A}=[a_{ij}]$，使

$$\lim_{k\to\infty}a_{ij}^{(k)}=a_{ij}\quad(i=1,2,\cdots,m;j=1,2,\cdots,n),$$

则称矩阵序列 $\{\boldsymbol{A}_k\}_{k=0}^{\infty}$ 收敛于 $\boldsymbol{A}$，称 $\boldsymbol{A}$ 为此序列的**极限**，记做 $\lim\limits_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}$. 否则，称此序列是**发散的**.

由定义可知，一个 $m\times n$ 矩阵序列收敛等价于 mn 个数列 $\{a_{ij}^{(k)}\}_{k=0}^{\infty}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 同时收敛. 因此，矩阵序列的敛散性，有许多是与数列的敛散性相类似的. 例如，若 $\lim\limits_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}$，$\lim\limits_{k\to\infty}\boldsymbol{B}_k=\boldsymbol{B}$，则有

$$\lim_{k\to\infty}c\boldsymbol{A}_k=c\lim_{k\to\infty}\boldsymbol{A}_k=c\boldsymbol{A},\quad c\text{ 是常数；}\qquad \lim_{k\to\infty}(\boldsymbol{A}_k+\boldsymbol{B}_k)=\lim_{k\to\infty}\boldsymbol{A}_k+\lim_{k\to\infty}\boldsymbol{B}_k=\boldsymbol{A}+\boldsymbol{B};$$

$$\lim_{k\to\infty}\boldsymbol{A}_k\boldsymbol{B}_k=\lim_{k\to\infty}\boldsymbol{A}_k\cdot\lim_{k\to\infty}\boldsymbol{B}_k=\boldsymbol{A}\boldsymbol{B};$$

$$\lim_{k\to\infty}(\boldsymbol{A}_k)^{-1}=\boldsymbol{A}^{-1}=(\lim_{k\to\infty}\boldsymbol{A}_k)^{-1}\text{（这里要求 }\boldsymbol{A}\text{ 及 }\boldsymbol{A}_k,k=0,1,\cdots\text{都是可逆矩阵）.}$$

定义 4.1-2 设 $\{\boldsymbol{A}_k\}_{k=0}^{\infty}$ 是一个 $m\times n$ 矩阵序列，则由它形成的无穷和

$$\boldsymbol{A}_0+\boldsymbol{A}_1+\cdots+\boldsymbol{A}_k+\cdots\tag{4.1-1}$$

叫做由 $\{\boldsymbol{A}_k\}_{k=0}^{\infty}$ 生成的**无穷级数**，记做 $\sum\limits_{k=0}^{\infty}\boldsymbol{A}_k$. 对于任一有限的 l，将 $\boldsymbol{S}_l=\sum\limits_{k=0}^{l}\boldsymbol{A}_k$ 叫做级数(4.1-1)的前 $l+1$ 项构成的**部分和**，而序列 $\{\boldsymbol{S}_l\}_{k=0}^{\infty}$ 叫做级数(4.1-1)的**部分和序列**. 如果 $\lim\limits_{l\to\infty}\boldsymbol{S}_l=\boldsymbol{S}$，则称级数(4.1-1)**收敛**，且称 $\boldsymbol{S}$ 为它的和，记做 $\sum\limits_{k=0}^{\infty}\boldsymbol{A}_k=\boldsymbol{S}$.

4.1 矩阵序列与矩阵级数

例 1 设 $A_k=\begin{bmatrix}\frac{1}{2^k} & \frac{\pi}{3\times 4^k}\\ 0 & \frac{1}{k(k+1)}\end{bmatrix}$, $k=1,2,\cdots$, 讨论序列 $\{A_k\}_{k=1}^{\infty}$ 和级数 $\sum\limits_{k=1}^{\infty}A_k$ 的敛散性.

解 $\lim\limits_{k\to\infty}A_k=O$,

由于
$$S_l=\sum_{k=1}^{l}A_k=\begin{bmatrix}\sum\limits_{k=1}^{l}\frac{1}{2^k} & \sum\limits_{k=1}^{l}\frac{\pi}{3\times 4^k}\\ 0 & \sum\limits_{k=1}^{l}\frac{1}{k(k+1)}\end{bmatrix},$$

且 $\lim\limits_{l\to\infty}\sum\limits_{k=1}^{l}\frac{1}{2^k}=1$, $\lim\limits_{l\to\infty}\sum\limits_{k=1}^{l}\frac{\pi}{3\cdot 4^k}=\frac{\pi}{9}$, $\lim\limits_{l\to\infty}\sum\limits_{k=1}^{l}\frac{1}{k(k+1)}=1$, 所以

$$\lim_{l\to\infty}S_l=\begin{bmatrix}1 & \frac{\pi}{9}\\ 0 & 1\end{bmatrix},\quad \sum_{k=1}^{\infty}A_k=\begin{bmatrix}1 & \frac{\pi}{9}\\ 0 & 1\end{bmatrix}.$$
■

容易看出,无穷级数的许多性质可以推广到矩阵级数,这里就不多说了.

利用范数可以把矩阵序列的敛散性问题化为一个数列的敛散性,这样既简便又更深入.事实上,矩阵序列 $\{A_k\}_{k=0}^{\infty}$ 收敛于 A 的充要条件是,序列 $\{A_k-A\}_{k=0}^{\infty}$ 收敛于 O. 而 $\lim\limits_{k\to\infty}(A_k-A)=O$ 的充要条件是 $\lim\limits_{k\to\infty}\|A_k-A\|=0$.

4.1 矩阵序列与矩阵级数

定理 4.1-1 设 $\boldsymbol{A}$ 是方阵，且 $\|\boldsymbol{A}\|<1$，则

$$\lim_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{O} \tag{4.1-2}$$

证 $\|\boldsymbol{A}^k\|=\|\boldsymbol{A}\boldsymbol{A}^{k-1}\|\leqslant\|\boldsymbol{A}\|\ \|\boldsymbol{A}^{k-1}\|\leqslant\cdots\leqslant\|\boldsymbol{A}\|^k.$

当 $\|\boldsymbol{A}\|<1$ 时，$\lim\limits_{k\to\infty}\|\boldsymbol{A}\|^k=0$，从而 $\lim\limits_{k\to\infty}\|\boldsymbol{A}^k\|=0$. 因此(4.1-2)式成立. ∎

但是，$\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{O}$ 未必能推出 $\|\boldsymbol{A}\|<1$. 例如 $\boldsymbol{A}=\begin{bmatrix}0.8 & 1\\ 0 & 0.8\end{bmatrix}$（它是 Jordan 块），则

$$\lim_{k\to\infty}\boldsymbol{A}^k=\lim_{k\to\infty}\begin{bmatrix}(0.8)^k & k(0.8)^{k-1}\\ 0 & (0.8)^k\end{bmatrix}=\boldsymbol{O},$$

而 $\|\boldsymbol{A}\|_\infty=\|\boldsymbol{A}\|_1=1.8>1$.

定理 4.1-2 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{O}$ 的充要条件是 $\rho(\boldsymbol{A})<1$.

证 **充分性**：由于 $\rho(\boldsymbol{A})<1$，故存在 $\varepsilon>0$ 使 $\rho(\boldsymbol{A})+\varepsilon<1$. 对这样的 ε，由 3.3 节中定理 3.3-2 知，存在矩阵范数 $\|\cdot\|_*$ 使

$$\|A\|_*\leqslant\rho(\boldsymbol{A})+\varepsilon<1.$$

于是由定理 4.1-1 推出 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{O}$.

必要性：由于存在可逆矩阵 $\boldsymbol{P}$，使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\boldsymbol{J}_1,\cdots,\boldsymbol{J}_s),$$

从而 $\boldsymbol{P}^{-1}\boldsymbol{A}^k\boldsymbol{P}=\mathrm{diag}(\boldsymbol{J}_1^k,\cdots,\boldsymbol{J}_s^k)$.

于是，$\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{O}$ 等价于 $\lim\limits_{k\to\infty}\boldsymbol{J}_i^k=\boldsymbol{O}(i=1,\cdots,s)$. 因为 $\boldsymbol{J}_i^k$ 的主对角线元素是 λ_i^k，故由 $\lim\limits_{k\to\infty}\lambda_i^k=0$ 知 $|\lambda_i|<1$. 由于 $\lambda_1,\cdots,\lambda_s$ 包含了 $\boldsymbol{A}$ 的所有不同特征值，所以 $\rho(\boldsymbol{A})<1$. ∎

4.1 矩阵序列与矩阵级数

下面讨论一类最重要的矩阵级数，即方阵 $\boldsymbol{A}$ 的幂级数 $\sum\limits_{k=0}^{\infty} c_k \boldsymbol{A}^k$，它的敛散性与对应的复变量幂级数 $\sum\limits_{k=0}^{\infty} c_k \lambda^k$ 的敛散性有密切的联系.

定理 4.1-3 设幂级数 $\sum\limits_{k=0}^{\infty} c_k \lambda^k$ 的收敛半径是 R，用方阵 $\boldsymbol{A}$ 替换该幂级数中的 λ，用 $\boldsymbol{I}$ 替换 $\lambda^0 = 1$ 得到方阵幂级数

$$\sum_{k=0}^{\infty} c_k \boldsymbol{A}^k, \tag{4.1-3}$$

则当 $\rho(\boldsymbol{A}) < R$ 时，方阵幂级数(4.1-3)收敛，而当 $\rho(\boldsymbol{A}) > R$ 时，方阵幂级数(4.1-3)发散.

证 由于任一方阵都相似于一个 Jordan 矩阵，故不妨设 $\boldsymbol{A}$ 是一个典型的 Jordan 块 $\boldsymbol{J}_i$：

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}_{r_i \times r_i}.$$

令

$$S_l(\lambda) = \sum_{k=0}^{l} c_k \lambda^k,$$

4.1 矩阵序列与矩阵级数

则 $\boldsymbol{S}_l(\boldsymbol{J}_i)=\sum_{k=0}^{l}c_k\boldsymbol{J}_i^k=\begin{bmatrix} S_l(\lambda) & S_l'(\lambda) & \frac{1}{2!}S_l''(\lambda) & \cdots & \frac{1}{(r_i-1)!}S_l^{(r_i-1)}(\lambda) \\ & S_l(\lambda) & S_l'(\lambda) & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2!}S_l''(\lambda) \\ & & & \ddots & S_l'(\lambda) \\ & & & & S_l(\lambda) \end{bmatrix}_{\lambda=\lambda_i}$ (4.1-4)

由幂级数理论知，当 $|\lambda_i|<R$ 时，部分和序列 $\{S_l(\lambda_i)\}_{l=0}^{\infty}$，$\{S_l'(\lambda_i)\}_{l=0}^{\infty}$，…，$\{S_l^{(r_i-1)}(\lambda_i)\}_{l=0}^{\infty}$ 都收敛；而当 $|\lambda_i|>R$ 时，$\{S_l(\lambda_i)\}_{l=0}^{\infty}$ 发散. 于是，如果 $\rho(\boldsymbol{A})<R$，则 $\boldsymbol{A}$ 的任一特征值的模都小于 $\boldsymbol{R}$，故方阵幂级数(4.1-3) 收敛；而当 $\rho(\boldsymbol{A})>R$ 时，必有一个特征值的模大于 R，不妨设为 $|\lambda_i|>1$，那么方阵幂级数 $\sum_{k=0}^{\infty}c_k\boldsymbol{J}_i^k$ 发散，从而方阵幂级数(4.1-3) 发散. ∎

4.1 矩阵序列与矩阵级数

例 2 讨论方阵幂级数

$$\sum_{k=1}^{\infty}\frac{1}{k}\begin{bmatrix}1 & 4\\ -2 & -3\end{bmatrix}^k$$

的敛散性.

解 与所给方阵幂级数对应的复变量幂级数是 $\sum_{k=1}^{\infty}\frac{1}{k}\lambda^k$，它的收敛半径

$$R=\lim_{k\to\infty}\frac{|c_k|}{|c_{k+1}|}=\lim_{k\to\infty}\frac{k+1}{k}=1.$$

而方阵 $\boldsymbol{A}=\begin{bmatrix}1 & 4\\ -2 & -3\end{bmatrix}$ 的特征值是 $-1+2\mathrm{i}$，$-1-2\mathrm{i}$，故其谱半径 $\rho(\boldsymbol{A})=\sqrt{5}$. 由于 $\rho(\boldsymbol{A})>R$，故所给方阵幂级数发散. ∎

例 3 讨论方阵幂级数

$$\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}1 & 4\\ -1 & -3\end{bmatrix}^k$$

的敛散性.

4.1 矩阵序列与矩阵级数

解 对应的复变量幂级数为$\sum\limits_{k=1}^{\infty}\frac{1}{k^2}\lambda^k$，它的收敛半径

$$R=\lim_{k\to\infty}\frac{(k+1)^2}{k^2}=1,$$

而方阵$\boldsymbol{A}=\begin{bmatrix}1 & 4\\ -1 & -3\end{bmatrix}$的谱半径$\rho(\boldsymbol{A})$也是1，因此不能用定理4.1-3来判断方阵幂级数是否收敛。但可应用$\boldsymbol{A}$的Jordan标准形判断它的敛散性。事实上，取$\boldsymbol{P}=\begin{bmatrix}2 & 1\\ -1 & 0\end{bmatrix}$，则有

$$\boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}-1 & 1\\ & -1\end{bmatrix}\boldsymbol{P}^{-1}.$$

由于

$$\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}-1 & 1\\ & -1\end{bmatrix}^k=\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}(-1)^k & k(-1)^{k-1}\\ & (-1)^k\end{bmatrix}$$

$$=\begin{bmatrix}\sum\limits_{k=1}^{\infty}\frac{(-1)^k}{k^2} & \sum\limits_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\\ & \sum\limits_{k=1}^{\infty}\frac{(-1)^k}{k^2}\end{bmatrix},$$

且$\sum\limits_{k=1}^{\infty}\frac{(-1)^k}{k^2}$，$\sum\limits_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}$都收敛，所以$\sum\limits_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}-1 & 1\\ & -1\end{bmatrix}^k$收敛。又因

4.1 矩阵序列与矩阵级数

$$\sum_{k=1}^{\infty}\frac{1}{k^2}\boldsymbol{A}^k=\sum_{k=1}^{\infty}\frac{1}{k^2}\left(\boldsymbol{P}\begin{bmatrix}-1 & 1\\ & -1\end{bmatrix}^k\boldsymbol{P}^{-1}\right)=\boldsymbol{P}\left(\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}-1 & 1\\ & -1\end{bmatrix}^k\right)\boldsymbol{P}^{-1},$$

故方阵幂级数$\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{bmatrix}1 & 4\\ -1 & -3\end{bmatrix}^k$收敛. ∎

4.2　方阵函数及其计算

众所周知，在收敛域内幂级数的和函数是解析函数（即任意阶导数都存在）. 因此我们可以用方阵幂级数定义方阵函数.

定义 4.2-1　设幂级数 $\sum_{k=0}^{\infty} c_k \lambda^k$ 的收敛半径为 R，且在收敛域内 $\sum_{k=0}^{\infty} c_k \lambda^k = f(\lambda)$. 当方阵 $\boldsymbol{A}$ 的谱半径 $\rho(\boldsymbol{A}) < R$ 时，定义 $\boldsymbol{f}(\boldsymbol{A}) = \sum_{k=0}^{\infty} c_k \boldsymbol{A}^k$，并称 $\boldsymbol{f}(\boldsymbol{A})$ 为 $\boldsymbol{A}$ **的函数.**

根据这个定义，可以得到形式上与高等数学中一些类似的方阵函数，例如

$$e^{\boldsymbol{A}} = \sum_{k=0}^{\infty} \frac{1}{k!} \boldsymbol{A}^k, \quad \rho(\boldsymbol{A}) < +\infty;$$

$$\sin \boldsymbol{A} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \boldsymbol{A}^{2k+1}, \quad \rho(\boldsymbol{A}) < +\infty;$$

$$\cos \boldsymbol{A} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \boldsymbol{A}^{2k}, \quad \rho(\boldsymbol{A}) < +\infty;$$

$$\ln(\boldsymbol{I} + \boldsymbol{A}) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \boldsymbol{A}^k, \quad \rho(\boldsymbol{A}) < 1;$$

$$(\boldsymbol{I} - \boldsymbol{A})^{-1} = \sum_{k=0}^{\infty} \boldsymbol{A}^k, \quad \rho(\boldsymbol{A}) < 1.$$

4.2 方阵函数及其计算

如果 $\boldsymbol{A}$ 的谱半径 $\rho(\boldsymbol{A})<+\infty$，则 $\rho(\boldsymbol{A}t)=|t|\rho(\boldsymbol{A})<+\infty$，其中 t 是参变量，因而对任意的 $t\in(-\infty,+\infty)$，

$$e^{\boldsymbol{A}t}=\sum_{k=0}^{\infty}\frac{1}{k!}(\boldsymbol{A}t)^k=\sum_{k=0}^{\infty}\frac{t^k}{k!}\boldsymbol{A}^k. \qquad (4.2\text{-}1)$$

类似地，可以定义 $\sin(\boldsymbol{A}t)$，$\cos(\boldsymbol{A}t)$ 等等.

既然方阵函数的定义与高等数学中的类似，因此可以将一些函数的许多性质平行地推广到方阵函数，例如，$\sin(2\boldsymbol{A})=2\sin\boldsymbol{A}\cos\boldsymbol{A}$，$\cos(2\boldsymbol{A})=\cos^2\boldsymbol{A}-\sin^2\boldsymbol{A}$，$e^{2\boldsymbol{A}}=(e^{\boldsymbol{A}})^2$，等等. 但由于矩阵乘积一般不可交换，所以涉及两个不同方阵时，有些性质是不能照搬的. 譬如，一般来说，$e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}}\neq e^{\boldsymbol{A}+\boldsymbol{B}}$，$\sin(\boldsymbol{A}+\boldsymbol{B})\neq\sin\boldsymbol{A}\cos\boldsymbol{B}+\cos\boldsymbol{A}\sin\boldsymbol{B}$，$\cos(\boldsymbol{A}+\boldsymbol{B})\neq\cos\boldsymbol{A}\cos\boldsymbol{B}-\sin\boldsymbol{A}\sin\boldsymbol{B}$，等等. 不过，当 $\boldsymbol{AB}=\boldsymbol{BA}$ 时，上述诸式都能成为等式.

如何计算方阵函数 $f(\boldsymbol{A})=\sum\limits_{k=0}^{\infty}c_k\boldsymbol{A}^k$ 呢？这里介绍两种常用的方法.

1. 利用方阵 $\boldsymbol{A}$ 的 Jordan 标准形

我们仍以典型的 Jordan 块为例说明.

由(4.1-4)式可知，当 $|\lambda_i|<R$ 时，

$$\lim_{l\to\infty}S_l(\lambda_i)=f(\lambda_i),$$

$$\lim_{l\to\infty}S_l^{(j)}(\lambda_i)=f^{(j)}(\lambda_i)\quad(j=1,\cdots,r_i-1),$$

故有

4.2 方阵函数及其计算

$$
f(\boldsymbol{J}_i) = \begin{bmatrix} f(\lambda) & \frac{\mathrm{d}}{\mathrm{d}\lambda} f(\lambda) & \frac{1}{2!}\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} f(\lambda) & \cdots & \frac{1}{(r-1)!}\frac{\mathrm{d}^{r_i-1}}{\mathrm{d}\lambda^{r_i-1}} f(\lambda) \\ & f(\lambda) & \frac{\mathrm{d}}{\mathrm{d}\lambda} f(\lambda) & \ddots & \vdots \\ & & f(\lambda) & \ddots & \frac{1}{2!}\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} f(\lambda) \\ & & & \ddots & \frac{\mathrm{d}}{\mathrm{d}\lambda} f(\lambda) \\ & & & & f(\lambda) \end{bmatrix}_{\lambda=\lambda_i} \tag{4.2-2}
$$

因此，若 $\boldsymbol{A}$ 的 Jordan 标准形是

$$
\boldsymbol{J} = \mathrm{diag}(J_1(\lambda_1), J_2(\lambda_2), \cdots, J_s(\lambda_s)),
$$

即存在可逆矩阵 $\boldsymbol{P}$，使

$$
\boldsymbol{A} = \boldsymbol{P}\mathrm{diag}(J_1(\lambda_1), J_2(\lambda_2), \cdots, J_s(\lambda_s))\boldsymbol{P}^{-1},
$$

则有

$$
f(\boldsymbol{A}) = \boldsymbol{P}\mathrm{diag}(f[J_1(\lambda_1)], f[J_2(\lambda_2)], \cdots, f[J_s(\lambda_s)])\boldsymbol{P}^{-1} \tag{4.2-3}
$$

4.2 方阵函数及其计算

例 1 $\boldsymbol{A}=\begin{bmatrix}0 & 1\\0 & -2\end{bmatrix}$，求 $e^{\boldsymbol{A}}$，$\cos\boldsymbol{A}$，$\sin\boldsymbol{A}$.

解 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\lambda(\lambda+2)$，$\boldsymbol{A}$ 的特征值是 $\lambda_1=0$ 和 $\lambda_2=-2$. $\boldsymbol{x}_1=[1\quad 0]^{\mathrm{T}}$，$\boldsymbol{x}_2=[1\quad -2]^{\mathrm{T}}$ 分别是属于 λ_1、λ_2 的特征向量. 因此，取 $\boldsymbol{P}=\begin{bmatrix}1 & 1\\0 & -2\end{bmatrix}$，则 $\boldsymbol{P}^{-1}=\begin{bmatrix}1 & 1/2\\0 & -1/2\end{bmatrix}$. 于是，有

$$\boldsymbol{A}=\begin{bmatrix}1 & 1\\0 & -2\end{bmatrix}\begin{bmatrix}0 & \\ & -2\end{bmatrix}\begin{bmatrix}1 & 1/2\\0 & -1/2\end{bmatrix},$$

$$e^{\boldsymbol{A}}=\boldsymbol{P}\begin{bmatrix}e^{0} & \\ & e^{-2}\end{bmatrix}\boldsymbol{P}^{-1}=\begin{bmatrix}1 & 1\\0 & -2\end{bmatrix}\begin{bmatrix}1 & \\ & e^{-2}\end{bmatrix}\begin{bmatrix}1 & 1/2\\0 & -1/2\end{bmatrix}=\begin{bmatrix}1 & \frac{1}{2}(1-e^{-2})\\0 & e^{-2}\end{bmatrix},$$

$$\cos\boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}\cos 0 & \\ & \cos(-2)\end{bmatrix}\boldsymbol{P}^{-1}=\begin{bmatrix}1 & 1\\0 & -2\end{bmatrix}\begin{bmatrix}1 & \\ & \cos 2\end{bmatrix}\begin{bmatrix}1 & 1/2\\0 & -1/2\end{bmatrix}=\begin{bmatrix}1 & \frac{1}{2}(1-\cos 2)\\0 & \cos 2\end{bmatrix},$$

$$\sin\boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}\sin 0 & \\ & \sin(-2)\end{bmatrix}\boldsymbol{P}^{-1}=\begin{bmatrix}1 & 1\\0 & -2\end{bmatrix}\begin{bmatrix}0 & \\ & -\sin 2\end{bmatrix}\begin{bmatrix}1 & 1/2\\0 & -1/2\end{bmatrix}=\begin{bmatrix}1 & \frac{1}{2}\sin 2\\0 & -\sin 2\end{bmatrix}.\ \blacksquare$$

4.2 方阵函数及其计算

例 2 设 $\boldsymbol{A}=\begin{bmatrix}0&1&0\\0&0&1\\2&3&0\end{bmatrix}$，求 $\mathrm{e}^{\boldsymbol{A}}$ 和 $\mathrm{e}^{\boldsymbol{A}t}$.

解 因

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=\begin{vmatrix}-\lambda&1&0\\0&-\lambda&1\\2&3&-\lambda\end{vmatrix}=\begin{vmatrix}2-\lambda&4-2\lambda&2-\lambda\\0&-\lambda&1\\2&3&-\lambda\end{vmatrix}=(2-\lambda)\begin{vmatrix}1&2&1\\0&-\lambda&1\\2&3&-\lambda\end{vmatrix}$$

$$=(2-\lambda)\begin{vmatrix}1&2&1\\0&-\lambda&1\\0&-1&-2-\lambda\end{vmatrix}=(2-\lambda)(1+\lambda)^2,$$

故 $\boldsymbol{A}$ 有两个不同的特征值：2 和 -1，且 -1 是二重特征值.

对特征值 2，解齐次线性方程组 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$，即解

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

得 $\boldsymbol{A}$ 的属于 2 的一个特征向量为 $\boldsymbol{x}_1=[1,2,4]^{\mathrm{T}}$.

4.2 方阵函数及其计算

对二重特征值-1，考虑线性方程组$(\boldsymbol{A}-(-1)\boldsymbol{I})\boldsymbol{x}=\boldsymbol{y}$，即

$$\begin{bmatrix}1&1&0\\0&1&1\\2&3&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}.$$

对它的增广矩阵施以行初等变换

$$\left[\begin{array}{ccc:c}1&1&0&y_1\\0&1&1&y_2\\2&3&1&y_3\end{array}\right]\to\left[\begin{array}{ccc:c}1&1&0&y_1\\0&1&1&y_2\\0&0&0&y_3-2y_1-y_2\end{array}\right].$$

由于$r(\boldsymbol{A}+\boldsymbol{I})=2$，故只有一个线性无关的1级根向量，取其为$\boldsymbol{x}_2=[1,-1,1]^{\mathrm{T}}$. 因为$x_3-2x_1-x_2=1-2-(-1)=0$，所以令$\boldsymbol{y}=\boldsymbol{x}_2$，则方程组

$$(\boldsymbol{A}+\boldsymbol{I})\boldsymbol{x}=\boldsymbol{x}_2$$

有解，它的解即是关于特征值-1的2级根向量，取其中的一个为$\boldsymbol{x}_3=[1,0,-1]^{\mathrm{T}}$.

令$\boldsymbol{P}=\begin{bmatrix}1&1&1\\2&-1&0\\4&1&-1\end{bmatrix}$，则$\boldsymbol{P}^{-1}=\dfrac{1}{9}\begin{bmatrix}1&2&1\\2&-5&2\\6&3&-3\end{bmatrix}$，且有

$$\boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}\boxed{2}&&\\&-1&1\\&&-1\end{bmatrix}\boldsymbol{P}^{-1}.$$

4.2 方阵函数及其计算

由于对应于 $f(\boldsymbol{A})=\mathrm{e}^{\boldsymbol{A}}$ 的函数是 $f(\lambda)=\mathrm{e}^{\lambda}$，且 $f'(\lambda)=\mathrm{e}^{\lambda}$，所以由(4.2-2)式和(4.2-3)式得

$$\mathrm{e}^{\boldsymbol{A}}=\begin{bmatrix}1 & 1 & 1\\ 2 & -1 & 0\\ 4 & 1 & -1\end{bmatrix}\begin{bmatrix}\mathrm{e}^{2} & & \\ & \mathrm{e}^{-1} & \mathrm{e}^{-1}\\ & & \mathrm{e}^{-1}\end{bmatrix}\cdot\frac{1}{9}\begin{bmatrix}1 & 2 & 1\\ 2 & -5 & 2\\ 6 & 3 & -3\end{bmatrix}$$

$$=\frac{1}{9}\begin{bmatrix}\mathrm{e}^{2}+14\mathrm{e}^{-1} & 2\mathrm{e}^{2}+\mathrm{e}^{-1} & \mathrm{e}^{2}-4\mathrm{e}^{-1}\\ 2\mathrm{e}^{2}-8\mathrm{e}^{-1} & 4\mathrm{e}^{2}+2\mathrm{e}^{-1} & 2\mathrm{e}^{2}+\mathrm{e}^{-1}\\ 4\mathrm{e}^{2}+2\mathrm{e}^{-1} & 8\mathrm{e}^{2}-5\mathrm{e}^{-1} & 4\mathrm{e}^{2}+2\mathrm{e}^{-1}\end{bmatrix}.$$

又因对应于 $f(\boldsymbol{A})=\mathrm{e}^{\boldsymbol{A}t}$ 的函数是 $f(\lambda)=\mathrm{e}^{\lambda t}$，且 $f'(\lambda)=t\mathrm{e}^{\lambda t}$，故由(4.2-2)式和(4.2-3)式得

$$\mathrm{e}^{\boldsymbol{A}t}=\frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 2 & -1 & 0\\ 4 & 1 & -1\end{bmatrix}\begin{bmatrix}\mathrm{e}^{2t} & & \\ & \mathrm{e}^{-t} & t\mathrm{e}^{-t}\\ & & \mathrm{e}^{-t}\end{bmatrix}\begin{bmatrix}1 & 2 & 1\\ 2 & -5 & 2\\ 6 & 3 & -3\end{bmatrix}$$

$$=\frac{1}{9}\begin{bmatrix}\mathrm{e}^{2t}+(8+6t)\mathrm{e}^{-t} & 2\mathrm{e}^{2t}-(2-3t)\mathrm{e}^{-t} & \mathrm{e}^{2t}-(1+3t)\mathrm{e}^{-t}\\ 2\mathrm{e}^{2t}-(2+6t)\mathrm{e}^{-t} & 4\mathrm{e}^{2t}+(5-3t)\mathrm{e}^{-t} & 2\mathrm{e}^{2t}-(2-3t)\mathrm{e}^{-t}\\ 4\mathrm{e}^{2t}-(4-6t)\mathrm{e}^{-t} & 8\mathrm{e}^{2t}-(8-3t)\mathrm{e}^{-t} & 4\mathrm{e}^{2t}+(5-3t)\mathrm{e}^{-t}\end{bmatrix}.$$

■

4.2 方阵函数及其计算

2. 利用方阵 A 的最小多项式或特征多项式

设 n 阶方阵 A 的最小多项式为

$$m_A(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s}$$
$$=\lambda^m+a_{m-1}\lambda^{m-1}+\cdots+a_1\lambda+a_0,$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 A 的所有不同特征值，$m_1+m_2+\cdots+m_s=m$，则由 $m_A(A)=O$ 得

$$A^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0I=O,$$

从而
$$A^m=-a_0I-a_1A-\cdots-a_{m-1}A^{m-1}.$$

这就是说，A 的 m 次幂 A^m 可表示为 $I,A,\cdots,A^{m-1}$ 的线性组合. 因此，A 的 m 次和高于 m 次的幂 $A^k(k\geqslant m)$ 都能表示为 $I,A,\cdots,A^{m-1}$ 的线性组合，亦即表示为 A 的一个次数不高于 $m-1$ 的多项式. 故而，$f(A)=\sum\limits_{k=0}^{\infty}c_kA^k$ 可望表示为 A 的某个次数不高于 $m-1$ 的多项式 $g(A)$.

设对应于 $g(A)$ 的多项式为

$$g(\lambda)=b_0\lambda^{m-1}+b_1\lambda^{m-2}+\cdots+b_{m-2}\lambda+b_{m-1},$$

其中 $b_0,b_1,\cdots,b_{m-1}$ 是待定的未知系数，又设 A 的 Jordan 标准形是

$$J=\mathrm{diag}(J_1,J_2,\cdots,J_s),$$

其中 $J_i(i=1,2,\cdots,s)$ 是关于特征值 λ_i 的子 Jordan 矩阵，且 J_i 中 Jordan 块的最大阶数是 m_i，那么，$g(A)=f(A)$ 的充要条件是

4.2　方阵函数及其计算

$$
\begin{bmatrix}
g(\lambda_i) & g'(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}g^{(m_i-1)}(\lambda_i) \\
 & g(\lambda_i) & \ddots & \vdots \\
 & & \ddots & g'(\lambda_i) \\
 & & & g(\lambda_i)
\end{bmatrix}
$$

$$
= \begin{bmatrix}
f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\
 & f(\lambda_i) & \ddots & \vdots \\
 & & \ddots & f'(\lambda_i) \\
 & & & f(\lambda_i)
\end{bmatrix},
$$

亦即

$$
g(\lambda_i)=f(\lambda_i),\ g'(\lambda_i)=f'(\lambda_i),\ \cdots,\ g^{(m_i-1)}(\lambda_i)=f^{(m_i-1)}(\lambda_i)
$$

$$
(i=1,\cdots,s). \tag{4.2-4}
$$

关于 $b_0, b_1, \cdots, b_{m-1}$ 的线性方程组(4.2-4)有 $m_1+m_2+\cdots+m_s=m$ 个方程，由于 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是互异的，故这 m 个方程是相互独立的，从而可唯一地解得 m 个未知系数 $b_0, b_1, \cdots, b_{m-1}$.①

---

① 实际上，(4.2-4)式是多项式插值条件，在第二部分数值计算方法中多项式插值将证明此结论.

4.2 方阵函数及其计算

例 3 设 $\boldsymbol{A}=\begin{bmatrix}2 & -1\\ 1 & 0\end{bmatrix}$，求 e^{A} 和 e^{At}.

解 由于

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=\begin{vmatrix}2-\lambda & -1\\ 1 & -\lambda\end{vmatrix}=(1-\lambda)^2,$$

且因 $\boldsymbol{A}-\boldsymbol{I}=\begin{bmatrix}1 & -1\\ 1 & -1\end{bmatrix}\neq\boldsymbol{O}$，故 $\boldsymbol{A}$ 的最小多项式是

$$m_A(\lambda)=(\lambda-1)^2.$$

因此设 $g(\lambda)=b_0\lambda+b_1$. 则由(4.2-4)式得

$$\begin{cases}g(1)=f(1),\\ g'(1)=f'(1),\end{cases}\quad 即 \quad \begin{cases}b_0+b_1=f(1),\\ b_0\qquad=f'(1).\end{cases}\tag{4.2-5}$$

由于对应于 $f(\boldsymbol{A})=e^{A}$ 的函数是 $f(\lambda)=e^{\lambda}$，且 $f'(\lambda)=e^{\lambda}$，故由(4.2-5)式得

$$\begin{cases}b_0+b_1=e,\\ b_0\qquad=e.\end{cases}$$

解之得 $b_0=e, b_1=0$. 于是 $e^{A}=e\cdot\boldsymbol{A}+0\cdot\boldsymbol{I}=\begin{bmatrix}2e & -e\\ e & 0\end{bmatrix}$.

4.2 方阵函数及其计算

由于对应于 $f(\boldsymbol{A})=\mathrm{e}^{\boldsymbol{A}t}$ 的函数是 $f(\lambda)=\mathrm{e}^{\lambda t}$，且 $f'(\lambda)=t\mathrm{e}^{\lambda t}$，故由(4.2-5)式得

$$\begin{cases} b_0+b_1 = \mathrm{e}^t, \\ b_0 \qquad = t\mathrm{e}^t, \end{cases}$$

解得 $b_0=t\mathrm{e}^t, b_1=(1-t)\mathrm{e}^t$. 于是

$$\mathrm{e}^{\boldsymbol{A}t}=t\mathrm{e}^t\boldsymbol{A}+(1-t)\mathrm{e}^t\boldsymbol{I}=\begin{bmatrix}(1+t)\mathrm{e}^t & -t\mathrm{e}^t \\ t\mathrm{e}^t & (1-t)\mathrm{e}^t\end{bmatrix}.$$

定义 4.2-2 满足条件(4.2-4)的两个函数 $f(\lambda)$ 与 $g(\lambda)$ 称为在 A 的谱 $Sp(\boldsymbol{A})=\{\lambda_1,\lambda_2,\cdots,\lambda_s\}$ 上一致.

由上述的讨论看出，$\boldsymbol{f}(\boldsymbol{A})=\boldsymbol{g}(\boldsymbol{A})$ 的充要条件是对应于它们的函数 $f(\lambda)$ 与 $g(\lambda)$ 在 $\boldsymbol{A}$ 的谱上一致.

由于求方阵 $\boldsymbol{A}$ 的最小多项式不是很容易的，所以通常用 $\boldsymbol{A}$ 的特征多项式替代最小多项式来计算方阵函数.

这时，设
$$g(\lambda)=b_0\lambda^{n-1}+b_1\lambda^{n-2}+\cdots+b_{n-2}\lambda+b_{n-1},$$
其中的未知系数 $b_0, b_1, \cdots, b_{n-1}$ 由条件

$$g(\lambda_i)=f(\lambda_i),\ g'(\lambda_i)=f'(\lambda_i),\ \cdots,\ g^{(n_i-1)}(\lambda_i)=f^{(n_i-1)}(\lambda_i),$$
$$(i=1,2,\cdots,s) \tag{4.2-6}$$

确定，这里的 n_i 是特征值 λ_i 的代数重数. 由多项式插值理论知，线性方程组(4.2-6)有解，从而可以确定多项式 $g(\lambda)$. 因为 $m_i\leqslant n_i(i=1,2,\cdots,s)$，故从(4.2-6)式可推出(4.2-4)式，所以有
$$\boldsymbol{f}(\boldsymbol{A})=\boldsymbol{g}(\boldsymbol{A}).$$

4.2 方阵函数及其计算

例 4 设 $\boldsymbol{A}=\begin{bmatrix}17 & 0 & -25\\ 0 & 3 & 0\\ 9 & 0 & -13\end{bmatrix}$，求 e^{At}.

解 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-3)(\lambda-2)^2$，$\boldsymbol{A}$ 的所有不同特征值是$\lambda_1=3$和 $\lambda_2=2$(二重特征值). 设

$$g(\lambda)=b_0\lambda^2+b_1\lambda+b_2,$$

由(4.2-6)式得

$$\begin{cases}g(3)=f(3),\\ g(2)=f(2),\\ g'(2)=f'(2),\end{cases}\quad 即\begin{cases}9b_0+3b_1+b_2=f(3),\\ 4b_0+2b_1+b_2=f(2),\\ 4b_0+b_1\qquad\quad=f'(2).\end{cases}$$

用 $f(\lambda)=e^{\lambda t}$ 代入上式，解得

$$b_0=e^{3t}-(1+t)e^{2t},\quad b_1=-4e^{3t}+(4+5t)e^{2t},$$

$$b_2=4e^{3t}-3(1+2t)e^{2t}.$$

因而

$$e^{At}=b_0\boldsymbol{A}^2+b_1\boldsymbol{A}+b_2\boldsymbol{I}=\begin{bmatrix}(1+15t)e^{2t} & 0 & -25te^{2t}\\ 0 & e^{3t} & 0\\ 9te^{2t} & 0 & (1-15t)e^{2t}\end{bmatrix}.$$ ∎

4.2 方阵函数及其计算

上面定义方阵函数 $f(\boldsymbol{A})$ 的实质是，先把 $f(\lambda)$ 展开成形如 $\sum_{k=0}^{\infty} c_k\lambda^k$ 的幂级数，然后用 $\boldsymbol{A}$ 替换 λ，得到 $f(\boldsymbol{A}) = \sum_{k=0}^{\infty} c_k\boldsymbol{A}^k$. 这样就要考虑方阵幂级数 $\sum_{k=0}^{\infty} c_k\boldsymbol{A}^k$ 的收敛性问题，从而要求 $\boldsymbol{A}$ 的谱半径小于幂级数 $\sum_{k=0}^{\infty} c_k\lambda^k$ 的收敛半径，于是限制了很大一类函数. 例如，$\ln\lambda$，λ^{-1} 按上述定义很难拓广到方阵的情况. 因此，我们用另一种方法定义方阵函数.

定义 4.2-3 设 n 阶方阵 $\boldsymbol{A}$ 的最小多项式为

$$m_A(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s},$$

而函数 $f(\lambda)$在 $\boldsymbol{A}$ 的谱 $S_p(\boldsymbol{A})=\{\lambda_1,\lambda_2,\cdots,\lambda_s\}$上的值

$$f(\lambda_i),f'(\lambda_i),\cdots,f^{(m_i-1)}(\lambda_i) \quad (i=1,2,\cdots,s)$$

都存在，则用(4.2-3)式定义 $f(\boldsymbol{A})$，即

$$f(\boldsymbol{A}) = \boldsymbol{P}\mathrm{diag}(f(J_1(\lambda_1)),f(J_2(\lambda_2)),\cdots,f(J_s(\lambda_s)))\boldsymbol{P}^{-1},$$

其中 $f(J_i(\lambda_i))$ 由(4.2-2) 式给出.

4.2 方阵函数及其计算

例 5 设 $\boldsymbol{A}=\begin{bmatrix}5 & -4 & 4\\ 1 & 0 & 5\\ 1 & -1 & 6\end{bmatrix}$，证明 $\ln\boldsymbol{A}$ 有定义，并求 $\ln\boldsymbol{A}$。

解 由于

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=\begin{vmatrix}5-\lambda & -4 & 4\\ 1 & -\lambda & 5\\ 1 & -1 & 6-\lambda\end{vmatrix}=\begin{vmatrix}5-\lambda & \lambda-5 & 5-\lambda\\ 1 & -\lambda & 5\\ 1 & -1 & 6-\lambda\end{vmatrix}$$

$$=(5-\lambda)\begin{vmatrix}1 & -1 & 1\\ 1 & -\lambda & 5\\ 1 & -1 & 6-\lambda\end{vmatrix}=(5-\lambda)\begin{vmatrix}1 & -1 & 1\\ 0 & 1-\lambda & 4\\ 0 & 0 & 5-\lambda\end{vmatrix}=(1-\lambda)(5-\lambda)^2,$$

且对于二重特征值 5， $r(\boldsymbol{A}-5\boldsymbol{I})=r\left(\begin{bmatrix}0 & -4 & 4\\ 1 & -5 & 5\\ 1 & -1 & 1\end{bmatrix}\right)=2$，

从而齐次线性方程组 $(\boldsymbol{A}-5\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$ 的线性无关解的个数为 $3-2=1$，故关于 5 存在 2 级根向量，所以 $\boldsymbol{A}$ 的最小多项式是

4.2 方阵函数及其计算

$$m_A(\lambda) = (\lambda - 1)(\lambda - 5)^2.$$

由于与 $f(\boldsymbol{A}) = \ln \boldsymbol{A}$ 对应的函数是 $f(\lambda) = \ln\lambda$，它在 $\boldsymbol{A}$ 的谱 $Sp(\boldsymbol{A}) = \{1, 5\}$ 上有

$$f(1) = \ln 1 = 0,\quad f(5) = \ln 5,\quad f'(5) = \frac{1}{5},$$

因此 $\ln\boldsymbol{A}$ 有定义.

若取 $\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$（$\boldsymbol{P}$ 由 $\boldsymbol{A}$ 的属于 1 的特征向量和关于 5 的 1 级与 2 级根向量构成），则

$$\boldsymbol{P}^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix},\quad \boldsymbol{A} = \boldsymbol{P}\begin{bmatrix} 1 & & \\ & 5 & 1 \\ & & 5 \end{bmatrix}\boldsymbol{P}^{-1}.$$

于是

$$\ln\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & & \\ & \ln 5 & \frac{1}{5} \\ & & \ln 5 \end{bmatrix}\begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} \ln 5 & -\ln 5 & \ln 5 \\ \frac{1}{5} & -\frac{1}{5} & \frac{1}{5} + \ln 5 \\ \frac{1}{5} & -\frac{1}{5} & \frac{1}{5} + \ln 5 \end{bmatrix}. \quad \blacksquare$$

4.3 函数矩阵及其应用

在研究线性微分方程组时，为了简化和运用矩阵来讨论，需要考虑以函数作为元素的矩阵，称这样的矩阵为函数矩阵.

如果矩阵 $\boldsymbol{A}(t)=[a_{ij}(t)]$ 的每个元素 $a_{ij}(t)$ 都是定义在同一区间 $[a,b]$（或 $[t_0,+\infty)$，$(-\infty,+\infty)$）上的函数，则称 $\boldsymbol{A}(t)$ 为定义在该区间上的一个**函数矩阵**. 又若每个 $a_{ij}(t)$ 在 $[a,b]$（或 $[t_0,+\infty)$，$(-\infty,+\infty)$）上有界、连续、可微、可积，则称 $\boldsymbol{A}(t)$ 在该区间上有界、连续、可微、可积. 从而可将高等数学的许多概念移植到矩阵论. 例如 $\boldsymbol{A}(t)$ 的导数和 $[t_0,t_1]$ 上的积分可分别定义为

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)=\left[\frac{\mathrm{d}}{\mathrm{d}t}a_{ij}(t)\right],\quad \int_{t_0}^{t_1}\boldsymbol{A}(\tau)\mathrm{d}\tau=\left[\int_{t_0}^{t_1}a_{ij}(\tau)\mathrm{d}\tau\right],$$

亦即 $\boldsymbol{A}(t)$ 的导数是 $\boldsymbol{A}(t)$ 的每个元素求导后组成的矩阵，$\boldsymbol{A}(t)$ 在 $[t_0,t_1]$ 上的积分是 $\boldsymbol{A}(t)$ 的每个元素在 $[t_0,t_1]$ 上积分后组成的矩阵.

因此，高等数学中微分和积分的许多性质对于函数矩阵仍是成立的. 但是，由于矩阵乘法不满足交换律，所以有些性质不能照搬. 例如 $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}^2(t)\neq 2\boldsymbol{A}(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)$，而是 $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}^2(t)=\frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{A}(t)\cdot\boldsymbol{A}(t)]=\boldsymbol{A}(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)+\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)\right)\boldsymbol{A}(t)$，仅当 $\boldsymbol{A}(t)$ 与 $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)$ 可交换时才有 $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}^2(t)=2\boldsymbol{A}(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)$.

4.3 函数矩阵及其应用

从 4.2 节有关方阵函数的论述中看到，涉及带有参变量的方阵函数实际上是函数方阵，例如 e^{At}，$\sin(At)$，$\cos(At)$ 是带有参变量 t 的方阵函数，但由于它们的每个元素是 t 的函数，所以它们也是函数方阵。换句话说，它们关于方阵 A 是方阵函数，而关于 t 则是函数方阵。

定理 4.3-1 对于任何方阵 A，e^{At} 可逆，且 $(e^{At})^{-1}=e^{-At}$，并且有

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{At}=Ae^{At}=e^{At}\cdot A,\quad t\in(-\infty,+\infty). \tag{4.3-1}$$

证 由于 A 与 $-A$ 的乘积是可交换的，所以

$$e^{At}\cdot e^{-At}=e^{(A-A)t}=I,$$

故

$$(e^{At})^{-1}=e^{-At}.$$

又因 $(e^{At})_{ij}=\sum_{k=0}^{\infty}\frac{t^k}{k!}(A^k)_{ij}$，故有

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{At})_{ij}=\sum_{k=1}^{\infty}\frac{t^{k-1}}{(k-1)!}(A^k)_{ij}=\sum_{k=0}^{\infty}\frac{t^k}{k!}(A^{k+1})_{ij},$$

因此

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{At}=\sum_{k=0}^{\infty}\frac{t^k}{k!}A^{k+1}=\begin{cases}A\left(\sum_{k=0}^{\infty}\frac{t^k}{k!}A^k\right)=Ae^{At}\\ \left(\sum_{k=0}^{\infty}\frac{t^k}{k!}A^k\right)A=e^{At}\cdot A.\end{cases}$$

∎

4.3 函数矩阵及其应用

类似地可得 $$\frac{\mathrm{d}}{\mathrm{d}t}\sin(\boldsymbol{A}t)=\boldsymbol{A}\cos(\boldsymbol{A}t)=\cos(\boldsymbol{A}t)\cdot\boldsymbol{A},$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\cos(\boldsymbol{A}t)=-\boldsymbol{A}\sin(\boldsymbol{A}t)=-\sin(\boldsymbol{A}t)\cdot\boldsymbol{A}.$$

定理 4.3-2 若 $\boldsymbol{A}(t)$对任意的 $t\in[a,b]$(或$[t_0,+\infty)$,$(-\infty,+\infty)$)可逆,且 $\boldsymbol{A}(t)$和 $\boldsymbol{A}^{-1}(t)$都可微,则有

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}^{-1}(t)=-\boldsymbol{A}^{-1}(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)\cdot\boldsymbol{A}^{-1}(t). \qquad (4.3\text{-}2)$$

证 因为 $\boldsymbol{A}^{-1}(t)\boldsymbol{A}(t)=\boldsymbol{I}$,两边对 t 求导得

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}^{-1}(t)\cdot\boldsymbol{A}(t)+\boldsymbol{A}^{-1}(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)=\boldsymbol{O},$$

所以 $$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}^{-1}(t)\cdot\boldsymbol{A}(t)=-\boldsymbol{A}^{-1}(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t).$$

用 $\boldsymbol{A}^{-1}(t)$右乘上式两边便得(4.3-2)式. ∎

4.3　函数矩阵及其应用

例 1　设 $\boldsymbol{A}$、$\boldsymbol{B}$ 都是 n 阶方阵，验证 $\boldsymbol{X}(t)=\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{B}$，$\boldsymbol{Y}(t)=\boldsymbol{B}\mathrm{e}^{\boldsymbol{A}t}$ 分别满足

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t)=\boldsymbol{A}\boldsymbol{X}(t),\qquad \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{Y}(t)=\boldsymbol{Y}(t)\boldsymbol{A}.$$

证
$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t)=\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{B})=\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\boldsymbol{A}t}\cdot\boldsymbol{B}+\mathrm{e}^{\boldsymbol{A}t}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{B}=\boldsymbol{A}\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{B}=\boldsymbol{A}\boldsymbol{X}(t).$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{Y}(t)=\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{B}\mathrm{e}^{\boldsymbol{A}t})=\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{B}\cdot\mathrm{e}^{\boldsymbol{A}t}+\boldsymbol{B}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\boldsymbol{A}t}=\boldsymbol{B}\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{A}=\boldsymbol{Y}(t)\boldsymbol{A}.\qquad\blacksquare$$

这个例子说明函数方阵 $\boldsymbol{X}(t)=\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{B}$ 满足矩阵微分方程 $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t)-\boldsymbol{A}\boldsymbol{X}(t)=\boldsymbol{O}$，且 $\boldsymbol{B}$ 可为任何 n 阶方阵. 而 $\boldsymbol{Y}(t)=\boldsymbol{B}\mathrm{e}^{\boldsymbol{A}t}$ 满足的矩阵微分方程是 $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{Y}(t)-\boldsymbol{Y}(t)\boldsymbol{A}=\boldsymbol{O}$，$\boldsymbol{B}$ 也可为任何 n 阶方阵.

4.3 函数矩阵及其应用

例 2 验证函数向量 $\boldsymbol{x}(t)=\sin(\boldsymbol{A}t)\boldsymbol{b}_1+\cos(\boldsymbol{A}t)\boldsymbol{b}_2$ 满足方程

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{x}(t)+\boldsymbol{A}^2\boldsymbol{x}(t)=\boldsymbol{0}, \tag{4.3-3}$$

这里 $\boldsymbol{A}$ 是 n 阶方阵，$\boldsymbol{b}_1$、$\boldsymbol{b}_2$ 为任意的 n 维向量.

解 由于

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t)=\frac{\mathrm{d}}{\mathrm{d}t}[\sin(\boldsymbol{A}t)\cdot\boldsymbol{b}_1+\cos(\boldsymbol{A}t)\cdot\boldsymbol{b}_2]=\frac{\mathrm{d}}{\mathrm{d}t}\sin(\boldsymbol{A}t)\cdot\boldsymbol{b}_1+\frac{\mathrm{d}}{\mathrm{d}t}\cos(\boldsymbol{A}t)\cdot\boldsymbol{b}_2$$

$$=\boldsymbol{A}\cos(\boldsymbol{A}t)\cdot\boldsymbol{b}_1-\boldsymbol{A}\sin(\boldsymbol{A}t)\cdot\boldsymbol{b}_2,$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{x}(t)=\frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{A}\cos(\boldsymbol{A}t)\cdot\boldsymbol{b}_1-\boldsymbol{A}\sin(\boldsymbol{A}t)\cdot\boldsymbol{b}_2]=\boldsymbol{A}\frac{\mathrm{d}}{\mathrm{d}t}\cos(\boldsymbol{A}t)\cdot\boldsymbol{b}_1-\boldsymbol{A}\frac{\mathrm{d}}{\mathrm{d}t}\sin(\boldsymbol{A}t)\cdot\boldsymbol{b}_2$$

$$=-\boldsymbol{A}^2\sin(\boldsymbol{A}t)\cdot\boldsymbol{b}_1-\boldsymbol{A}^2\cos(\boldsymbol{A}t)\cdot\boldsymbol{b}_2=-\boldsymbol{A}^2\boldsymbol{x}(t),$$

所以 $\boldsymbol{x}(t)=\sin(\boldsymbol{A}t)\boldsymbol{b}_1+\cos(\boldsymbol{A}t)\boldsymbol{b}_2$ 满足方程(4.3-3). ∎

例 3 证明：若函数矩阵 $\boldsymbol{A}(t)$ 可积，则当 $t_1>t_0$ 时，

$$\left\|\int_{t_0}^{t_1}\boldsymbol{A}(\tau)\mathrm{d}\tau\right\|_1\leqslant\int_{t_0}^{t_1}\|\boldsymbol{A}(\tau)\|_1\mathrm{d}\tau. \tag{4.3-4}$$

证 由于 $\int_{t_0}^{t_1}\boldsymbol{A}(\tau)\mathrm{d}\tau$ 的第 i 行第 j 列的元素与 $\boldsymbol{A}(\tau)$ 的第 i 行第 j 列的元素有下述关系式：

$$\left(\int_{t_0}^{t_1}\boldsymbol{A}(\tau)\mathrm{d}\tau\right)_{ij}=\int_{t_0}^{t_1}(\boldsymbol{A}(\tau))_{ij}\mathrm{d}\tau,$$

4.3 函数矩阵及其应用

所以
$$\left|\left(\int_{t_0}^{t_1}\boldsymbol{A}(\tau)\mathrm{d}\tau\right)_{ij}\right| = \left|\int_{t_0}^{t_1}(\boldsymbol{A}(\tau))_{ij}\mathrm{d}\tau\right| \leqslant \int_{t_0}^{t_1}\left|(\boldsymbol{A}(\tau))_{ij}\right|\mathrm{d}\tau.$$

于是
$$\sum_{i=1}^{m}\left|\left(\int_{t_0}^{t_1}\boldsymbol{A}(\tau)\mathrm{d}\tau\right)_{ij}\right| \leqslant \sum_{i=1}^{m}\int_{t_0}^{t_1}\left|(\boldsymbol{A}(\tau))_{ij}\right|\mathrm{d}\tau,\ j = 1,\cdots,n,$$

从而
$$\max_{1\leqslant j\leqslant n}\sum_{i=1}^{m}\left|\left(\int_{t_0}^{t_1}\boldsymbol{A}(\tau)\mathrm{d}\tau\right)_{ij}\right| \leqslant \max_{1\leqslant j\leqslant n}\int_{t_0}^{t_1}\sum_{i=1}^{m}\left|(\boldsymbol{A}(\tau))_{ij}\right|\mathrm{d}\tau,$$

即(4.3-4)式成立. ■

类似地,可由矩阵的∞-范数定义得到

$$\left\|\int_{t_0}^{t_1}\boldsymbol{A}(\tau)\mathrm{d}\tau\right\|_{\infty} \leqslant \int_{t_0}^{t_1}\left\|\boldsymbol{A}(\tau)\right\|_{\infty}\mathrm{d}\tau \quad (t_1 > t_0). \tag{4.3-5}$$

一个高阶线性微分方程必能化为等价的一阶线性微分方程组.例如,对于线性微分方程

$$\frac{\mathrm{d}^3}{\mathrm{d}t^3}x + 6\frac{\mathrm{d}^2}{\mathrm{d}t^2}x + 11\frac{\mathrm{d}}{\mathrm{d}t}x - 4x = t + \mathrm{e}^{-2t},$$

令 $x_1 = x, x_2 = \frac{\mathrm{d}}{\mathrm{d}t}x, x_3 = \frac{\mathrm{d}^2}{\mathrm{d}t^2}x$,则有
$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}t}x_1 = x_2,\\[2mm] \dfrac{\mathrm{d}}{\mathrm{d}t}x_2 = x_3,\\[2mm] \dfrac{\mathrm{d}}{\mathrm{d}t}x_3 = \dfrac{\mathrm{d}^3}{\mathrm{d}t^3}x = -6\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}x - 11\dfrac{\mathrm{d}}{\mathrm{d}t}x + 4x + t + \mathrm{e}^{-2t}\\[2mm] \qquad\quad = 4x_1 - 11x_2 - 6x_3 + t + \mathrm{e}^{-2t}.\end{cases}$$

4.3 函数矩阵及其应用

用矩阵和向量记号，上述一阶线性方程组可以表示成如下形式

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{g}(t), \tag{4.3-6}$$

其中
$$\boldsymbol{x}(t)=\begin{bmatrix}x_1(t)\\x_2(t)\\x_3(t)\end{bmatrix},\ \boldsymbol{A}=\begin{bmatrix}0&1&0\\0&0&1\\4&-11&-6\end{bmatrix},\boldsymbol{g}(t)=\begin{bmatrix}0\\0\\t+\mathrm{e}^{-2t}\end{bmatrix}.$$

即使是高阶线性微分方程组，也可以化为等价的一阶线性微分方程组.

例 4 将线性微分方程组

$$\begin{cases}\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}u+3\dfrac{\mathrm{d}}{\mathrm{d}t}v+4u+v=8t,\\[2ex]\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}u-\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}v+\dfrac{\mathrm{d}}{\mathrm{d}t}u+2v=\cos t\end{cases}$$

化为一阶线性微分方程组.

解 首先通过代数运算将所给线性微分方程组化为

$$\begin{cases}\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}u+3\dfrac{\mathrm{d}}{\mathrm{d}t}v+4u+v=8t,\\[2ex]\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}v-\dfrac{\mathrm{d}}{\mathrm{d}t}u+3\dfrac{\mathrm{d}}{\mathrm{d}t}v+4u-v=8t-\cos t.\end{cases}$$

然后令 $x_1=u,x_2=\dfrac{\mathrm{d}}{\mathrm{d}t}u,x_3=v,x_4=\dfrac{\mathrm{d}}{\mathrm{d}t}v$，则有

4.3 函数矩阵及其应用

$$
\begin{cases}
\dfrac{\mathrm{d}}{\mathrm{d}t}x_1 = x_2, \\[2ex]
\dfrac{\mathrm{d}}{\mathrm{d}t}x_2 = \dfrac{\mathrm{d}^2 u}{\mathrm{d}t^2} = -4u - v - 3\dfrac{\mathrm{d}}{\mathrm{d}t}v + 8t = -4x_1 - x_3 - 3x_4 + 8t, \\[2ex]
\dfrac{\mathrm{d}}{\mathrm{d}t}x_3 = x_4, \\[2ex]
\dfrac{\mathrm{d}}{\mathrm{d}t}x_4 = \dfrac{\mathrm{d}^2 v}{\mathrm{d}t^2} = -4u + \dfrac{\mathrm{d}}{\mathrm{d}t}u + v - 3\dfrac{\mathrm{d}}{\mathrm{d}t}v + 8t - \cos t = -4x_1 + x_2 + x_3 - 3x_4 + 8t - \cos t.
\end{cases}
$$

于是,若令

$$
\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -4 & 0 & -1 & -3 \\ 0 & 0 & 0 & 1 \\ -4 & 1 & 1 & -3 \end{bmatrix}, \quad \boldsymbol{g}(t) = \begin{bmatrix} 0 \\ 8t \\ 0 \\ 8t - \cos t \end{bmatrix},
$$

则得到向量-矩阵微分方程(4.3-6). ▮

4.3 函数矩阵及其应用

为了求解形如(4.3-6)式的非齐次线性微分方程组，与高等数学中求解非齐次线性微分方程一样，先求对应的齐次线性微分方程组的通解，再用常数变易法求解非齐次线性微分方程组.

对应于非齐次线性微分方程(4.3-6)的齐次线性微分方程是

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}-\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}. \tag{4.3-7}$$

由例1可知，齐次线性微分方程(4.3-7)的通解应为 $\boldsymbol{x}(t)=\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{b}$，这里 $\boldsymbol{b}$ 是任意常数向量.

设非齐次线性微分方程(4.3-6)的解为

$$\boldsymbol{x}(t)=\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{b}(t), \tag{4.3-8}$$

则

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t)=\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{b}(t))=\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\boldsymbol{A}t}\cdot\boldsymbol{b}(t)+\mathrm{e}^{\boldsymbol{A}t}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{b}(t)$$

$$=\boldsymbol{A}\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{b}(t)+\mathrm{e}^{\boldsymbol{A}t}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{b}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\mathrm{e}^{\boldsymbol{A}t}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{b}(t).$$

将此代入(4.3-6)式，得

$$\mathrm{e}^{\boldsymbol{A}t}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{b}(t)=\boldsymbol{g}(t).$$

因 $\mathrm{e}^{\boldsymbol{A}t}$ 可逆，故有

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{b}(t)=\mathrm{e}^{-\boldsymbol{A}t}\boldsymbol{g}(t).$$

两边积分得

$$\boldsymbol{b}(t)=\boldsymbol{b}(t_0)+\int_{t_0}^{t}\mathrm{e}^{-\boldsymbol{A}\tau}\boldsymbol{g}(\tau)\mathrm{d}\tau.$$

4.3 函数矩阵及其应用

将它代入(4.3-8)式，得

$$\boldsymbol{x}(t)=\mathrm{e}^{At}\boldsymbol{b}(t_0)+\mathrm{e}^{At}\int_{t_0}^{t}\mathrm{e}^{-A\tau}\boldsymbol{g}(\tau)\mathrm{d}\tau=\mathrm{e}^{At}\boldsymbol{b}(t_0)+\int_{t_0}^{t}\mathrm{e}^{A(t-\tau)}\boldsymbol{g}(\tau)\mathrm{d}\tau. \qquad (4.3\text{-}9)$$

特别地，在(4.3-9)式中令 $t=t_0$ 得 $\boldsymbol{x}(t_0)=\mathrm{e}^{At_0}\boldsymbol{b}(t_0)$，即有 $\boldsymbol{b}(t_0)=\mathrm{e}^{-At_0}\boldsymbol{x}(t_0)$. 将其代入(4.3-9)式得

$$\boldsymbol{x}(t)=\mathrm{e}^{A(t-t_0)}\boldsymbol{x}(t_0)+\int_{t_0}^{t}\mathrm{e}^{A(t-\tau)}\boldsymbol{g}(\tau)\mathrm{d}\tau. \qquad (4.3\text{-}10)$$

例 5 求解向量矩阵微分方程的初值问题：

$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}=\begin{bmatrix}1 & 2\\ 4 & 3\end{bmatrix}\boldsymbol{x}+\begin{bmatrix}1\\ -1\end{bmatrix},\\ \boldsymbol{x}|_{t=0}=[1,2]^{\mathrm{T}}.\end{cases}$$

解 这里 $\boldsymbol{A}=\begin{bmatrix}1 & 2\\ 4 & 3\end{bmatrix}$，由 4.2 节计算方阵函数的方法可得

$$\mathrm{e}^{At}=\frac{1}{6}\begin{bmatrix}2\mathrm{e}^{5t}+4\mathrm{e}^{-t} & 2\mathrm{e}^{5t}-2\mathrm{e}^{-t}\\ 4\mathrm{e}^{5t}-4\mathrm{e}^{-t} & 4\mathrm{e}^{5t}+2\mathrm{e}^{-t}\end{bmatrix},$$

于是，(4.3-10)式给出

$$\boldsymbol{x}(t)=\mathrm{e}^{At}\begin{bmatrix}1\\ 2\end{bmatrix}+\int_{t_0}^{t}\mathrm{e}^{A(t-\tau)}\begin{bmatrix}1\\ -1\end{bmatrix}\mathrm{d}\tau=\begin{bmatrix}1-\mathrm{e}^{-t}+\mathrm{e}^{5t}\\ -1+\mathrm{e}^{-t}+2\mathrm{e}^{5t}\end{bmatrix}.$$

■

4.3 函数矩阵及其应用

例 6 求解矩阵微分方程的初值问题：

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{A}\boldsymbol{X}(t) + \boldsymbol{X}(t)\boldsymbol{B}, \\ \boldsymbol{X}(t)\big|_{t=0} = \boldsymbol{X}_0, \end{cases} \tag{4.3-11}$$

其中 $\boldsymbol{A}$ 是 m 阶方阵，$\boldsymbol{B}$ 是 n 阶方阵，$\boldsymbol{X}(t)$ 是 $m\times n$ 阶函数矩阵.

解 所给矩阵微分方程等价于下方程

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) - \boldsymbol{A}\boldsymbol{X}(t) - \boldsymbol{X}(t)\boldsymbol{B} = \boldsymbol{O}.$$

由于 $\mathrm{e}^{\boldsymbol{A}t}$ 和 $\mathrm{e}^{\boldsymbol{B}t}$ 都可逆，故上式等价于

$$\mathrm{e}^{-\boldsymbol{A}t}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t)\cdot\mathrm{e}^{-\boldsymbol{B}t} - \mathrm{e}^{-\boldsymbol{A}t}\boldsymbol{A}\boldsymbol{X}(t)\mathrm{e}^{-\boldsymbol{B}t} - \mathrm{e}^{-\boldsymbol{A}t}\boldsymbol{X}(t)\boldsymbol{B}\mathrm{e}^{-\boldsymbol{B}t} = \boldsymbol{O},$$

即

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{-\boldsymbol{A}t}\boldsymbol{X}(t)\mathrm{e}^{-\boldsymbol{B}t}\right] = \boldsymbol{O}.$$

因此，从 0 到 t 积分得 $\mathrm{e}^{-\boldsymbol{A}t}\boldsymbol{X}(t)\mathrm{e}^{-\boldsymbol{B}t} = \boldsymbol{X}_0$，即 $\boldsymbol{X}(t) = \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{X}_0\mathrm{e}^{\boldsymbol{B}t}$. ■

作业（第1部分）

习题 4.2

1. 求下列方阵 $\boldsymbol{A}$ 的函数 $e^{\boldsymbol{A}}$ 和 $e^{\boldsymbol{A}t}$：

$$(3)\ \boldsymbol{A}=\begin{bmatrix} 2 & 1 & 3 \\ 0 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix}.$$

矩阵分解

矩阵论（9）

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黄 炜
博士·副教授
Dr. Wei Huang

5 矩阵分解

矩阵分解是将矩阵表示为几个特殊矩阵的乘积，这对于与矩阵有关的数值计算和理论分析都有重要的意义，并且分解方法和过程本身就提供了某些有效的数值计算方法，我们将在第二部分数值计算方法中应用它们.

事实上，我们在 1.3 节已讨论了矩阵的一种分解——满秩分解，第二章所讨论的方阵相似化简问题实际上也是矩阵分解的内容，只不过局限于相似变换而已.

本章给出矩阵的另几种重要的分解.

5.1 方阵的三角分解

线性代数中利用矩阵的行初等变换把所给线性方程组化为等价的三角形方程组，从而容易求得原方程组的解．例如，对于线性方程组

$$\begin{cases} x_1+2x_2+x_3=0, \\ 2x_1+2x_2+3x_3=0, \\ -x_1-3x_2=2, \end{cases} \tag{5.1-1}$$

求解的过程是，对它的增广矩阵施以行初等变换：

$$\left[\begin{array}{ccc:c} 1 & 2 & 1 & 0 \\ 2 & 2 & 3 & 3 \\ -1 & -3 & 0 & 2 \end{array}\right] \xrightarrow[\text{第三行加上第一行}]{\substack{\text{第二行减去}\\\text{第一行的两倍}}} \left[\begin{array}{ccc:c} 1 & 2 & 1 & 0 \\ 0 & -2 & 1 & 3 \\ 0 & -1 & 1 & 2 \end{array}\right] \xrightarrow{\substack{\text{第三行减去}\\\text{第二行的}\frac{1}{2}\text{倍}}} \left[\begin{array}{ccc:c} 1 & 2 & 1 & 0 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & 1/2 & 1/2 \end{array}\right]$$

得到与方程组(5.1-1)等价的三角形方程组

$$\begin{cases} x_1+2x_2+x_3=0, \\ -2x_2+x_3=3, \\ \dfrac{1}{2}x_3=\dfrac{1}{2}. \end{cases} \tag{5.1-2}$$

于是从最后一个方程解得 $x_3=1$；将 $x_3=1$ 代入第二个方程解得 $x_2=-1$；再将 $x_3=1,x_2=-1$ 代入第一个方程解得 $x_1=1$．这个过程称为**回代**．

5.1 方阵的三角分解

上述对增广矩阵施以行初等变换，实质上是先用矩阵 $\boldsymbol{L}_1$ 然后再用矩阵 $\boldsymbol{L}_2$ 左乘该增广矩阵，这里

$$\boldsymbol{L}_1=\begin{bmatrix}1&0&0\\-2&1&0\\1&0&1\end{bmatrix},\quad \boldsymbol{L}_2=\begin{bmatrix}1&0&0\\0&1&0\\0&-1/2&1\end{bmatrix}.$$

若记方程组(5.1-1)的系数矩阵为 $\boldsymbol{A}$，方程组(5.1-2)的系数矩阵为 $\boldsymbol{R}$，则有

$$\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{A}=\boldsymbol{R}，\text{或者 } \boldsymbol{A}=(\boldsymbol{L}_2\boldsymbol{L}_1)^{-1}\boldsymbol{R}=\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\boldsymbol{R}.$$

容易验证

$$\boldsymbol{L}_1^{-1}=\begin{bmatrix}1&0&0\\2&1&0\\-1&0&1\end{bmatrix},\quad \boldsymbol{L}_2^{-1}=\begin{bmatrix}1&0&0\\0&1&0\\0&1/2&1\end{bmatrix}.$$

令 $\boldsymbol{L}=\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}$，则

$$\boldsymbol{L}=\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}=\begin{bmatrix}1&0&0\\2&1&0\\-1&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&1/2&1\end{bmatrix}=\begin{bmatrix}1&0&0\\2&1&0\\-1&1/2&1\end{bmatrix},$$

且

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{R}. \tag{5.1-3}$$

5.1 方阵的三角分解

主对角线以上(下)的元素全为零的方阵称为下(上)三角矩阵,并称主对角线元素都是 1 的下(上)三角矩阵为**单位下(上)三角矩阵**. 例如上述的 L_1、L_2、L 等都是单位下三角矩阵,而

$$R=\begin{bmatrix}1 & 2 & 1\\0 & -2 & 1\\0 & 0 & 1/2\end{bmatrix}$$

是上三角矩阵. 于是(5.1-3)式表明,方阵 A 可以分解成一个单位下三角矩阵与一个上三角矩阵的乘积.

定义 5.1-1 如果方阵 A 可表示为形如(5.1-3)式的下三角矩阵与上三角矩阵的乘积,则称 A 可直接三角分解,简称为**可三角分解**;而(5.1-3)式称为 A 的一个**三角分解**.

特别地,当 A 的三角分解中的下三角矩阵是单位下三角矩阵,则称它为 **Doolittle 分解**.

显然,一个方阵的三角分解并不唯一. 这是因为,如果 $A=LR$ 是 A 的一个三角分解,取 D 是任一与 A 同阶的可逆对角矩阵,则

$$A=LR=LDD^{-1}R=(LD)(D^{-1}R)=L_1R_1,$$

其中 $L_1=LD$ 是下三角矩阵,$R_1=D^{-1}R$ 是上三角矩阵,故 L_1R_1 也是 A 的一个三角分解. 但我们有下述定理.

5.1 方阵的三角分解

定理 5.1-1 n 阶方阵 $\boldsymbol{A}=[a_{ij}]$ 有唯一 Doolittle 分解的充要条件是 $\boldsymbol{A}$ 的前 $n-1$ 个顺序主子式不为零，即

$$\Delta_k=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}\neq 0 \quad (k=1,2,\cdots,n-1).$$

证 对方阵的阶数 n 用数学归纳法. 当 $n=1$ 时，定理显然成立. 现假设对 $n-1$ 阶方阵定理成立，即 $n-1$ 阶方阵 $\boldsymbol{A}_{n-1}$ 有唯一 Doolittle 分解：

$$\boldsymbol{A}_{n-1}=\boldsymbol{L}_{n-1}\boldsymbol{R}_{n-1},$$

其中 $\boldsymbol{L}_{n-1}$ 是单位下三角矩阵，$\boldsymbol{R}_{n-1}$ 是上三角矩阵.

设 $\boldsymbol{A}$ 的分块形式是

$$\boldsymbol{A}=\left[\begin{array}{c:c} \boldsymbol{A}_{n-1} & \boldsymbol{\gamma} \\ \hdashline \boldsymbol{\beta}^{\mathrm{T}} & a_{nn} \end{array}\right],$$

其中 $n-1$ 维向量 $\boldsymbol{\beta}^{\mathrm{T}}=[a_{n1},a_{n2},\cdots,a_{n,n-1}]$，$\boldsymbol{\gamma}=[a_{1n},a_{2n},\cdots,a_{n-1,n}]^{\mathrm{T}}$ 及数 a_{nn} 都是已知的.

必要性：若 $\boldsymbol{A}$ 有唯一的 Doolittle 分解，其分块形式为

$$\left[\begin{array}{c:c} \boldsymbol{A}_{n-1} & \boldsymbol{\gamma} \\ \hdashline \boldsymbol{\beta}^{\mathrm{T}} & a_{nn} \end{array}\right]=\left[\begin{array}{c:c} \boldsymbol{L}_{n-1} & \boldsymbol{0} \\ \hdashline \boldsymbol{p}^{\mathrm{T}} & 1 \end{array}\right]\left[\begin{array}{c:c} \boldsymbol{R}_{n-1} & \boldsymbol{q} \\ \hdashline \boldsymbol{0} & r_{nn} \end{array}\right], \qquad (5.1\text{-}4)$$

5.1　方阵的三角分解

充分性：若 $\Delta_{n-1}\neq 0$，则由于 $\det \boldsymbol{L}_{n-1}=1$ 和 $\det \boldsymbol{R}_{n-1}^{\mathrm{T}}=\det \boldsymbol{R}_{n-1}=\Delta_{n-1}\neq 0$，所以由(5.1-5)式的第二式知，给定 $\boldsymbol{\gamma}$ 有唯一的 $\boldsymbol{q}$ 满足线性方程组 $\boldsymbol{L}_{n-1}\boldsymbol{x}=\boldsymbol{\gamma}$；而由(5.1-5)式的第三式知，给定 $\boldsymbol{\beta}$ 有唯一的 $\boldsymbol{p}$ 满足线性方程组 $\boldsymbol{R}_{n-1}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{\beta}$；又由(5.1-5)式的最后一个式子得 $r_{nn}=a_{nn}-\boldsymbol{p}^{\mathrm{T}}\boldsymbol{q}$. 于是存在 $\boldsymbol{A}$ 的形如(5.1-4)式的 Doolittle 分解.

又因为若 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{R}=\tilde{\boldsymbol{L}}\tilde{\boldsymbol{R}}$ 是 $\boldsymbol{A}$ 的两个 Doolittle 分解，则有

$$\boldsymbol{L}_{n-1}\boldsymbol{R}_{n-1}=\boldsymbol{A}_{n-1}=\tilde{\boldsymbol{L}}_{n-1}\tilde{\boldsymbol{R}}_{n-1},$$

且

$$\det \boldsymbol{R}_{n-1}=\det \boldsymbol{A}_{n-1}=\det \tilde{\boldsymbol{R}}_{n-1},$$

故 $\det \tilde{\boldsymbol{R}}_{n-1}\neq 0$，从而 $\tilde{\boldsymbol{R}}_{n-1}$ 可逆. 于是

$$\boldsymbol{L}_{n-1}^{-1}\tilde{\boldsymbol{L}}_{n-1}=\boldsymbol{R}_{n-1}\tilde{\boldsymbol{R}}_{n-1}^{-1}.$$

由于 $\boldsymbol{L}_{n-1}^{-1}\tilde{\boldsymbol{L}}_{n-1}$ 是单位下三角矩阵，而 $\boldsymbol{R}_{n-1}\tilde{\boldsymbol{R}}_{n-1}^{-1}$ 是上三角矩阵，所以

$$\boldsymbol{L}_{n-1}^{-1}\tilde{\boldsymbol{L}}_{n-1}=\boldsymbol{I}=\boldsymbol{R}_{n+1}\tilde{\boldsymbol{R}}_{n-1}^{-1}.$$

因此

$$\tilde{\boldsymbol{L}}_{n-1}=\boldsymbol{L}_{n-1},\quad \tilde{\boldsymbol{R}}_{n-1}=\boldsymbol{R}_{n-1}. \qquad \blacksquare$$

从(5.1-4)式可以看出，若 $\boldsymbol{A}$ 的前 $n-1$ 个顺序主子式不为零，则可用待定系数法确定 $\boldsymbol{A}$ 的三角分解式(5.1-3)中的 $\boldsymbol{L}$ 和 $\boldsymbol{R}$. 事实上，令

5.1 方阵的三角分解

$$L = \begin{bmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ \vdots & \vdots & \ddots & & \\ & & & 1 & \\ l_{n1} & l_{n2} & \cdots & l_{n,n-1} & 1 \end{bmatrix}, \quad R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix},$$

则由

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

可以推出这些元素之间的关系：

$$\begin{cases} r_{1j} = a_{1j} \quad (j = 1, 2, \cdots, m); \quad l_{i1} = \dfrac{a_{i1}}{r_{11}} \quad (i = 2, 3, \cdots, n) \\ r_{kj} = a_{kj} - \sum\limits_{m=1}^{k-1} l_{km} r_{mj} \qquad (k = 2, 3, \cdots, n; j = k, k+1, \cdots, n) \\ l_{ik} = \left(a_{ik} - \sum\limits_{m=1}^{k-1} l_{im} r_{mk}\right) \Big/ r_{kk} \quad (k = 2, 3, \cdots, n-1; i = k+1, \cdots, n). \end{cases} \tag{5.1-6}$$

5.1　方阵的三角分解

因此，从 $\boldsymbol{R}$ 的第 1 行和 $\boldsymbol{L}$ 的第 1 列开始，对于 $k=2,3,\cdots,n$，交替地使用(5.1-6)式的第 2 式和第 3 式，计算共为 n 步，每步先计算 $\boldsymbol{R}$ 的一行，再计算 $\boldsymbol{L}$ 相应的一列，最后得出 $\boldsymbol{L}$ 和 $\boldsymbol{R}$. 其计算过程如图 5.1-1 所示.

在计算机上，算出的 $\boldsymbol{L}$ 和 $\boldsymbol{R}$ 的元素可存放在 $\boldsymbol{A}$ 的相应位置上.

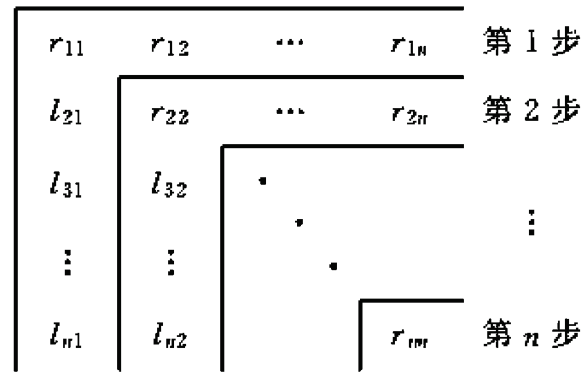

图 5.1-1

如果已知方阵 $\boldsymbol{A}$ 的 Doolittle 分解 $\boldsymbol{A}=\boldsymbol{LR}$，那么线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 等价于 $\boldsymbol{LRx}=\boldsymbol{b}$，从而又等价于两个三角形方程组

$$\begin{cases} \boldsymbol{Ly} = \boldsymbol{b}, \\ \boldsymbol{Rx} = \boldsymbol{y}. \end{cases} \tag{5.1-7}$$

当 $\boldsymbol{A}$ 可逆时，由(5.1-7)式的第 1 个线性方程组依次递推地解得 $y_1,y_2,\cdots,y_n$，而由(5.1-7)式的第 2 个线性方程组用上述的回代过程依次递推地解得 $x_n,x_{n-1},\cdots,x_1$. 因而等价的两个三角形方程组(5.1-7)的求解公式是

$$\begin{cases} y_1 = b_1,\ y_i = b_i - \sum\limits_{j=1}^{i-1} l_{ij} y_j \quad (i=2,3,\cdots,n), \\ x_n = \dfrac{y_n}{r_{nn}},\ x_i = \dfrac{y_i - \sum\limits_{j=k+1}^{n} r_{ij} x_j}{r_{kk}} \quad (i=n-1,\cdots,1). \end{cases} \tag{5.1-8}$$

5.1 方阵的三角分解

用这种方法求解线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 需要假设 $\boldsymbol{A}$ 的前 $n-1$ 个顺序主子式不为零，但这对于解此线性方程组本身并不是必要的. 实际上，只要 $\boldsymbol{A}$ 可逆，它就有唯一解. 因此，我们要讨论只知 $\boldsymbol{A}$ 可逆下如何应用这种方法求解 $\boldsymbol{Ax}=\boldsymbol{b}$ 的问题. 这时可以通过用置换矩阵(即对单位矩阵仅应用第一种行(列)初等变换所得的矩阵)左乘 $\boldsymbol{A}$ 来实现 $\boldsymbol{A}$ 的行交换，使其满足前 $n-1$ 个顺序主子式不为零的条件. 这就是带行交换的三角分解.

定理 5.1-2 若 n 阶方阵可逆，则存在置换矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{PA}$ 的 n 个顺序主子式都不为零，从而 $\boldsymbol{PA}$ 可直接三角分解.

证 由于 $\boldsymbol{A}$ 可逆，故 $\boldsymbol{A}$ 的前 $n-1$ 列必线性无关，从而由 $\boldsymbol{A}$ 的前 $n-1$ 个列向量组成 $n\times(n-1)$ 矩阵 $\boldsymbol{B}$ 的秩是 $n-1$. 由于矩阵 $\boldsymbol{B}$ 的行秩等于 $\boldsymbol{B}$ 的秩，所以必有 $\boldsymbol{B}$ 的某 $n-1$ 行线性无关，从而可以通过行的交换将它们换为前 $n-1$ 行. 也就是说，存在置换矩阵 $\boldsymbol{P}_1$ 使 $\boldsymbol{P}_1\boldsymbol{A}$ 的第 $n-1$ 个顺序主子式不为零. 再对 $\boldsymbol{P}_1\boldsymbol{A}$ 中前 $n-1$ 行与前 $n-1$ 列交叉处的 $(n-1)^2$ 个元素所组成的 $n-1$ 阶矩阵(其行列式即是 $\boldsymbol{P}_1\boldsymbol{A}$ 的第 $n-1$ 个顺序主子式)进行上述过程. 如此继续下去，至多有 $n-1$ 个置换 $\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_{n-1}$，便使 $\boldsymbol{P}_{n-1}\cdots\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{A}$ 的前 $n-1$个顺序主子式都不为零. 显然，置换矩阵的乘积仍为置换矩阵，故得结论. ∎

根据定理 5.1-2，对线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的等价线性方程组 $\boldsymbol{PAx}=\boldsymbol{Pb}$ 可用上述的方法求解，即将 $\boldsymbol{PA}$ 进行三角分解，将它表示为 $\boldsymbol{PA}=\boldsymbol{LR}$，并将线性方程组 $\boldsymbol{PAx}=\boldsymbol{Pb}$ 等价地化为两个三角形方程组

$$\begin{cases}\boldsymbol{Ly}=\boldsymbol{Pb},\\ \boldsymbol{Rx}=\boldsymbol{y}.\end{cases}$$

5.1 方阵的三角分解

例 1 证明 $\boldsymbol{A}=\begin{bmatrix}2 & 2 & -5\\ 1 & -3 & 1\\ 1 & 5 & 2\end{bmatrix}$ 可直接三角分解，并求它的 Doolittle 分解.

解 由于 $\Delta_1=2\neq 0$，$\Delta_2=\begin{vmatrix}2 & 2\\ 1 & -3\end{vmatrix}=-8\neq 0$，所以 $\boldsymbol{A}$ 可直接三角分解. 由

$$\begin{bmatrix}1 & & \\ l_{21} & 1 & \\ l_{31} & l_{32} & 1\end{bmatrix}\begin{bmatrix}r_{11} & r_{12} & r_{13}\\ & r_{22} & r_{23}\\ & & r_{33}\end{bmatrix}=\begin{bmatrix}2 & 2 & -5\\ 1 & -3 & 1\\ 1 & 5 & 2\end{bmatrix}$$

得

$$r_{11}=2,\quad r_{12}=2,\quad r_{13}=-5;\quad l_{21}=\frac{1}{2},\quad l_{31}=\frac{1}{2};$$

$$r_{22}=-3-l_{21}r_{12}=-4,\quad r_{23}=1-l_{21}r_{13}=\frac{7}{2};$$

$$l_{31}r_{12}+l_{32}r_{22}=5,\text{即 } l_{32}=\frac{5-l_{31}r_{12}}{r_{22}}=-1;$$

$$l_{31}r_{13}+l_{32}r_{23}+r_{33}=2,\text{即 } r_{33}=2-l_{31}r_{13}-l_{32}r_{23}=8,$$

因此

$$\boldsymbol{A}=\underset{(L)}{\begin{bmatrix}1 & & \\ 1/2 & 1 & \\ 1/2 & -1 & 1\end{bmatrix}}\underset{(R)}{\begin{bmatrix}2 & 2 & -5\\ & -4 & 7/2\\ & & 8\end{bmatrix}}.$$

∎

5.1 方阵的三角分解

例 2 讨论方阵 $\boldsymbol{A}_1$ 的三角分解，其中

$$\boldsymbol{A}_1=\begin{bmatrix}0&1&-1&2\\1&0&0&1\\2&-1&0&1\\1&3&0&1\end{bmatrix}.$$

解 因 $\Delta_1=0$，故 $\boldsymbol{A}$ 不能直接三角分解. 又因 $|\boldsymbol{A}|=-3\neq0$，故存在置换矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{PA}$ 可直接三角分解. 为了找出 $\boldsymbol{P}$，首先在 $\boldsymbol{A}$ 的第 1 列中选取绝对值最大的元素 2，记下此元素的行序数 $\underline{3}$，并把此行与第 1 行交换得到 $\boldsymbol{A}_2$，并对 $\boldsymbol{A}_2$ 施以行初等变换使 $\boldsymbol{A}_2$ 的第 1 列除第 1 个元素 2 之外的其余元素均为零，即

$$\boldsymbol{A}_2=\begin{bmatrix}2&-1&0&1\\1&0&0&1\\0&1&-1&2\\1&3&0&1\end{bmatrix}\rightarrow\begin{bmatrix}2&-1&0&1\\0&1/2&0&1/2\\0&1&-1&2\\0&7/2&0&1/2\end{bmatrix}\text{记为 }\boldsymbol{A}'_2;$$

再在 $\boldsymbol{A}'_2$ 的第 2 列中选取除第 1 个元素 -1 之外的绝对值最大的元素 $\frac{7}{2}$，记下此元素的原行序数 $\underline{4}$，并把此行与第 2 行交换得到 $\boldsymbol{A}_3$，并对 $\boldsymbol{A}_3$ 施以行初等变换使 $\boldsymbol{A}_3$ 的第 2 列除前 2 个元素 -1、$\frac{7}{2}$ 之外的其余元素均为零，即

5.1 方阵的三角分解

$$A_3=\begin{bmatrix}2 & -1 & 0 & 1\\ 0 & \frac{7}{2} & 0 & \frac{1}{2}\\ 0 & 1 & -1 & 2\\ 0 & \frac{1}{2} & 0 & \frac{1}{2}\end{bmatrix}\to\begin{bmatrix}2 & -1 & 0 & 1\\ 0 & \frac{7}{2} & 0 & \frac{1}{2}\\ 0 & 0 & -1 & \frac{13}{7}\\ 0 & 0 & 0 & \frac{3}{7}\end{bmatrix}\text{记为 }A'_3.$$

由于 A'_3 已是上三角矩阵，故可停止. 否则继续依次进行直至成为上三角矩阵.

对于本例来说，行序数依次为 3，4，1，2. 这就是说，对单位矩阵进行第 3 行与第 1 行交换，第 4 行与第 2 行交换所得的矩阵是所要求的置换矩阵 P，即

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

于是，PA 可直接三角分解，且 $R=A'_3$. 事实上，

$$PA=\begin{bmatrix}2 & -1 & 0 & 1\\ 1 & 3 & 0 & 1\\ 0 & 1 & -1 & 2\\ 1 & 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}1 & & & \\ 1/2 & 1 & & \\ 0 & 2/7 & 1 & \\ 1/2 & 1/7 & 0 & 1\end{bmatrix}\times\begin{bmatrix}2 & -1 & 0 & 1\\ & 7/2 & 0 & 1/2\\ & & -1 & 13/7\\ & & & 3/7\end{bmatrix}.\quad\blacksquare$$

5.1 方阵的三角分解

这种选取置换矩阵的方法叫做**按列选主元**.其特征是 $\boldsymbol{PA}$ 的三角分解中下三角矩阵的元素之绝对值均不大于 1.

当 $\boldsymbol{A}$ 是正定矩阵时，由于它的顺序主子式 $\Delta_k>0(k=1,2,\cdots,n)$，所以 $\boldsymbol{A}$ 有唯一的 Doolittle 分解 $\boldsymbol{A}=\boldsymbol{LR}$，且 $\boldsymbol{R}$ 的主对角线元素 $r_{kk}(k=1,2,\cdots,n)$ 均大于零. 显然，$\boldsymbol{R}$ 又可唯一地表示成一个对角矩阵 $\boldsymbol{D}$ 与一个单位上三角矩阵 $\boldsymbol{R}_1$ 的乘积，即 $\boldsymbol{R}=\boldsymbol{DR}_1$，这里 $\boldsymbol{D}=\mathrm{diag}(r_{11},r_{22},\cdots,r_{nn})$，将 $\boldsymbol{R}=\boldsymbol{DR}_1$ 代入 $\boldsymbol{A}=\boldsymbol{LR}$ 得到

$$\boldsymbol{A}=\boldsymbol{LDR}_1. \tag{5.1-9}$$

又因 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，故由(5.1-9)式得

$$(\boldsymbol{LDR}_1)^{\mathrm{T}}=\boldsymbol{R}_1^{\mathrm{T}}\boldsymbol{DL}^{\mathrm{T}}=\boldsymbol{LDR}_1. \tag{5.1-10}$$

但 $\boldsymbol{L}$、$\boldsymbol{D}$、$\boldsymbol{R}$ 均可逆，所以由(5.1-10)式得

$$\boldsymbol{D}^{-1}\boldsymbol{L}^{-1}\boldsymbol{R}_1^{\mathrm{T}}\boldsymbol{D}=\boldsymbol{R}_1(\boldsymbol{L}^{\mathrm{T}})^{-1}=\boldsymbol{R}_1(\boldsymbol{L}^{-1})^{\mathrm{T}}.$$

因为下(上)三角矩阵的转置是上(下)三角矩阵，所以上式的左端是单位下三角矩阵，而右端是单位上三角矩阵，故有 $\boldsymbol{R}_1(\boldsymbol{L}^{\mathrm{T}})^{-1}=\boldsymbol{I}$，即 $\boldsymbol{R}_1=\boldsymbol{L}^{\mathrm{T}}$. 将此代入(5.1-9)式得

$$\boldsymbol{A}=\boldsymbol{LDL}^{\mathrm{T}}. \tag{5.1-11}$$

也就是说，当 $\boldsymbol{A}$ 正定时，$\boldsymbol{A}$ 有唯一的分解式(5.1-11).

与 $\boldsymbol{A}$ 的三角分解一样，可用待定系数法确定 $\boldsymbol{L}$ 和 $\boldsymbol{D}$. 事实上，由

5.1 方阵的三角分解

$$A = LDL^{\mathrm{T}} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \begin{bmatrix} 1 & l_{21} & \cdots & l_{n1} \\ & 1 & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

不难得到计算 $\boldsymbol{L}$ 和 $\boldsymbol{D}$ 的元素的递推公式：

$$\begin{cases} d_k = a_{kk} - \sum\limits_{m=1}^{k-1} l_{km}^2 d_m \quad (k = 1, 2, \cdots, n), \\ l_{ik} = \left(a_{ik} - \sum\limits_{m=1}^{k-1} l_{im} l_{km} d_m\right) / d_k \quad (i = k+1, \cdots, n). \end{cases} \tag{5.1-12}$$

5.1 方阵的三角分解

例 3 证明 $\boldsymbol{A}=\begin{bmatrix}5 & -2 & 0\\ -2 & 3 & -1\\ 0 & -1 & 1\end{bmatrix}$ 是正定矩阵，并求它的 $\boldsymbol{LDL}^{\mathrm{T}}$ 分解.

解 由于 $\Delta_1=5,\Delta_2=11,\Delta_3=6$ 均大于零，故 $\boldsymbol{A}$ 是正定矩阵，又由

$$\begin{bmatrix}1 & & \\ l_{21} & 1 & \\ l_{31} & l_{32} & 1\end{bmatrix}\begin{bmatrix}d_1 & & \\ & d_2 & \\ & & d_3\end{bmatrix}\begin{bmatrix}1 & l_{21} & l_{31}\\ & 1 & l_{32}\\ & & 1\end{bmatrix}=\begin{bmatrix}5 & -2 & 0\\ -2 & 3 & -1\\ 0 & -1 & 1\end{bmatrix}$$

得

$$d_1=5,\quad l_{21}=-\frac{2}{5},\quad l_{31}=0;$$

$$d_2=3-\left(-\frac{2}{5}\right)^2\times 5=\frac{11}{5},$$

$$l_{32}=(-1-0)\Big/\frac{11}{5}=-\frac{5}{11};$$

$$d_3=1-0\times 5-\left(-\frac{5}{11}\right)^2\times\frac{11}{5}=\frac{6}{11}.$$

因此

$$\boldsymbol{A}=\begin{bmatrix}1 & & \\ -\frac{2}{5} & 1 & \\ 0 & -\frac{11}{5} & 1\end{bmatrix}\begin{bmatrix}5 & & \\ & \frac{11}{5} & \\ & & \frac{6}{11}\end{bmatrix}\begin{bmatrix}1 & -\frac{2}{5} & 0\\ & 1 & -\frac{11}{5}\\ & & 1\end{bmatrix}.$$

∎

5.1 方阵的三角分解

对于正定矩阵 $\boldsymbol{A}$，在其分解式(5.1-11)的基础上再令 $\boldsymbol{D}^{\frac{1}{2}}=\operatorname{diag}(\sqrt{d_1},\sqrt{d_2},\cdots,\sqrt{d_n})$，则

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{L}^{\mathrm{T}}=(\boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}})(\boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}})^{\mathrm{T}}=\boldsymbol{G}\boldsymbol{G}^{\mathrm{T}}, \tag{5.1-13}$$

其中 $\boldsymbol{G}=\boldsymbol{L}\boldsymbol{D}^{\frac{1}{2}}$ 是下三角矩阵，它的主对角线元素均大于零.

定义 5.1-2 设 $\boldsymbol{A}$ 是正定矩阵，则称(5.1-13)式为 $\boldsymbol{A}$ 的 **Cholesky 分解或对称三角分解.**

同样，可用待定系数法确定正定矩阵 $\boldsymbol{A}$ 的 Cholesky 分解中 $\boldsymbol{G}$ 的元素. 令 $\boldsymbol{G}=[g_{ij}]$，则由

5.2 方阵的正交(酉)三角分解

在 1.1 节中，我们讲过用 Schmidt 正交化方法把内积空间的一般基化为标准正交基的公式(1.1-8)．若把它改写一下，用 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 表示 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$，则有

$$\begin{cases}\boldsymbol{\alpha}_1=\boldsymbol{\beta}_1=|\boldsymbol{\beta}_1|\boldsymbol{\varepsilon}_1,\\ \boldsymbol{\alpha}_2=\langle\boldsymbol{\alpha}_2,\boldsymbol{\varepsilon}_1\rangle\boldsymbol{\varepsilon}_1+\boldsymbol{\beta}_2=\langle\boldsymbol{\alpha}_2,\boldsymbol{\varepsilon}_1\rangle\boldsymbol{\varepsilon}_1+|\boldsymbol{\beta}_2|\boldsymbol{\varepsilon}_2,\\ \quad\vdots\\ \boldsymbol{\alpha}_n=\langle\boldsymbol{\alpha}_n,\boldsymbol{\varepsilon}_1\rangle\boldsymbol{\varepsilon}_1+\cdots+\langle\boldsymbol{\alpha}_n,\boldsymbol{\varepsilon}_{n-1}\rangle\boldsymbol{\varepsilon}_{n-1}+\boldsymbol{\beta}_n\\ \quad=\langle\boldsymbol{\alpha}_n,\boldsymbol{\varepsilon}_1\rangle\boldsymbol{\varepsilon}_1+\cdots+\langle\boldsymbol{\alpha}_n,\boldsymbol{\varepsilon}_{n-1}\rangle\boldsymbol{\varepsilon}_{n-1}+|\boldsymbol{\beta}_n|\boldsymbol{\varepsilon}_n.\end{cases}$$

用矩阵来表示，上式可写为 $[\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\cdots\ \boldsymbol{\alpha}_n]=[\boldsymbol{\varepsilon}_1\ \boldsymbol{\varepsilon}_2\cdots\ \boldsymbol{\varepsilon}_n]\boldsymbol{R}$， (5.2-1)

其中

$$\boldsymbol{R}=\begin{bmatrix}|\boldsymbol{\beta}_1| & \langle\boldsymbol{\alpha}_2,\boldsymbol{\varepsilon}_1\rangle & \cdots & \langle\boldsymbol{\alpha}_n,\boldsymbol{\varepsilon}_1\rangle\\ 0 & |\boldsymbol{\beta}_2| & \cdots & \langle\boldsymbol{\alpha}_n,\boldsymbol{\varepsilon}_2\rangle\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \langle\boldsymbol{\alpha}_n,\boldsymbol{\varepsilon}_{n-1}\rangle\\ 0 & 0 & \cdots & |\boldsymbol{\beta}_n|\end{bmatrix}$$

是一个上三角矩阵．

在 $\mathbf{R}^n$ 中，$\boldsymbol{\alpha}_i(i=1,2,\cdots,n)$是 n 维实向量，$[\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\cdots\ \boldsymbol{\alpha}_n]$是 n 阶可逆实矩阵；$\boldsymbol{\varepsilon}_i(i=1,2,\cdots,n)$是 n 维实单位向量，$[\boldsymbol{\varepsilon}_1\ \boldsymbol{\varepsilon}_2\cdots\ \boldsymbol{\varepsilon}_n]$是正交矩阵．而在 $\mathbf{C}^n$ 中，$\boldsymbol{\alpha}_i(i=1,2,\cdots,n)$是 n 维复向量，$[\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\cdots\ \boldsymbol{\alpha}_n]$是 n 阶可逆复矩阵；$\boldsymbol{\varepsilon}_i(i=1,2,\cdots,n)$是 n 维复单位向量，$[\boldsymbol{\varepsilon}_1\ \boldsymbol{\varepsilon}_2\cdots\ \boldsymbol{\varepsilon}_n]$是酉矩阵．因此，(5.2-1)式给出了一个重要结果，即可逆矩阵的正交(酉)三角分解．

5.2 方阵的正交(酉)三角分解

定义 5.2-1 若实(复)可逆矩阵 $\boldsymbol{A}$ 能够分解为正交(酉)矩阵 $\boldsymbol{Q}$ 与实(复)的可逆上三角矩阵 $\boldsymbol{R}$ 的乘积,即

$$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}, \tag{5.2-2}$$

则称 $\boldsymbol{A}$ **可正交(酉)三角分解**,而(5.2-2)式称为 $\boldsymbol{A}$ 的 **QR 分解**.

由于若 $\boldsymbol{D}$ 是一个主对角线元素的绝对值(模)全为 1 的对角矩阵,则

$$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}=(\boldsymbol{Q}\boldsymbol{D})(\boldsymbol{D}^{-1}\boldsymbol{R})=\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{R}},$$

其中 $\tilde{\boldsymbol{Q}}=\boldsymbol{Q}\boldsymbol{D}$,$\tilde{\boldsymbol{R}}=\boldsymbol{D}^{-1}\boldsymbol{R}$,且 $\tilde{\boldsymbol{Q}}$ 是正交(酉)矩阵,$\tilde{\boldsymbol{R}}$ 是上三角矩阵(请读者自证),所以方阵 $\boldsymbol{A}$ 的 QR 分解不唯一.

定理 5.2-1 设 $\boldsymbol{A}$ 是 n 阶实(复)可逆矩阵,则必存在 $\boldsymbol{A}$ 的 QR 分解式(5.2-2),并且除去相差一个主对角线元素的绝对值(模)全为 1 的对角矩阵因子外,(5.2-2)式是唯一的.

证 由于实(复)方阵 $\boldsymbol{A}$ 可逆,故 $\boldsymbol{A}$ 的 n 个列向量构成 $\mathbf{R}^n(\mathbf{C}^n)$ 的一个基,所以可用 Schmidt 正交化方法把这个基化为标准正交基. 因此必存在 $\boldsymbol{A}$ 的 QR 分解式(5.2-2).

再证唯一性. 设 $\boldsymbol{A}$ 有两个分解式,即

$$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}=\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{R}},$$

则有

$$\boldsymbol{Q}=\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{R}}\boldsymbol{R}^{-1}=\tilde{\boldsymbol{Q}}\boldsymbol{G}, \tag{5.2-3}$$

其中 $\boldsymbol{G}=\tilde{\boldsymbol{R}}\boldsymbol{R}^{-1}$ 是上三角矩阵.

5.2 方阵的正交(酉)三角分解

由于 Q、$\tilde{Q}$ 是正交矩阵，故由(5.2-3)式得

$$I=Q^{T}Q=(\tilde{Q}G)^{T}(\tilde{Q}G)=G^{T}\tilde{Q}^{T}\tilde{Q}G=G^{T}G,$$

所以 G 是正交矩阵，又因 G 是上三角的，故 G 只能是主对角元素的绝对值全为 1 的对角矩阵. 因此

$$\tilde{R}=GR,\quad \tilde{Q}=QG^{-1}.$$

这就是说，A 的 QR 分解式除去可能相差一个主对角线元素的绝对值全为 1 的对角矩阵因子外是唯一的.

对于复方阵，同理可证这个定理. ■

可以利用待定系数法确定可逆方阵 A 的 QR 分解式中的方阵 Q 和 R.

以三阶可逆实矩阵 A 为例说明. 设

$$A=\begin{bmatrix}1&2&2\\2&1&2\\1&2&1\end{bmatrix}=[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\boldsymbol{\alpha}_3],$$

$$Q=\begin{bmatrix}q_{11}&q_{12}&q_{13}\\q_{21}&q_{22}&q_{23}\\q_{31}&q_{32}&q_{33}\end{bmatrix}=[\boldsymbol{q}_1\quad\boldsymbol{q}_2\quad\boldsymbol{q}_3],\quad R=\begin{bmatrix}r_{11}&r_{12}&r_{13}\\&r_{22}&r_{23}\\&&r_{33}\end{bmatrix},$$

其中 $\boldsymbol{\alpha}_1=[1,2,1]^{T}$，$\boldsymbol{\alpha}_2=[2,1,2]^{T}$，$\boldsymbol{\alpha}_3=[2,2,1]^{T}$，$\boldsymbol{q}_1=[q_{11},q_{21},q_{31}]^{T}$，$\boldsymbol{q}_2=[q_{12},q_{22},q_{32}]^{T}$，$\boldsymbol{q}_3=[q_{13},q_{23},q_{33}]^{T}$，则由 $A=QR$，即由

5.2 方阵的正交(酉)三角分解

$$[\boldsymbol{q}_1\ \boldsymbol{q}_2\ \boldsymbol{q}_3]\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ & r_{22} & r_{23} \\ & & r_{33} \end{bmatrix} = [\boldsymbol{\alpha}_1 \ \vdots\ \boldsymbol{\alpha}_2 \ \vdots\ \boldsymbol{\alpha}_3]$$

得

$$\begin{cases} r_{11}\boldsymbol{q}_1 = \boldsymbol{\alpha}_1, \\ r_{12}\boldsymbol{q}_1 + r_{22}\boldsymbol{q}_2 = \boldsymbol{\alpha}_2, \\ r_{13}\boldsymbol{q}_1 + r_{23}\boldsymbol{q}_2 + r_{33}\boldsymbol{q}_3 = \boldsymbol{\alpha}_3, \end{cases}$$

即

$$\begin{cases} r_{11}\boldsymbol{q}_1 = \boldsymbol{\alpha}_1, \\ r_{22}\boldsymbol{q}_2 = \boldsymbol{\alpha}_2 - r_{12}\boldsymbol{q}_1, \\ r_{33}\boldsymbol{q}_3 = \boldsymbol{\alpha}_3 - r_{13}\boldsymbol{q}_1 - r_{23}\boldsymbol{q}_2. \end{cases} \tag{5.2-4}$$

由于 Q 是正交矩阵，所以 $\langle \boldsymbol{q}_i, \boldsymbol{q}_j \rangle = \delta_{ij}$.

因此，由(5.2-4)式的第一式得 $|r_{11}| = |\boldsymbol{\alpha}_1| = \sqrt{\langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1 \rangle}$,

即

$$r_{11} = \pm\sqrt{\langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1 \rangle},$$

$$\boldsymbol{q}_1 = \frac{1}{r_{11}}\boldsymbol{\alpha}_1;$$

由(5.2-4)式的第二式得

$$0 = r_{22}\langle \boldsymbol{q}_2, \boldsymbol{q}_1 \rangle = \langle r_{22}\boldsymbol{q}_2, \boldsymbol{q}_1 \rangle = \langle \boldsymbol{\alpha}_2 - r_{12}\boldsymbol{q}_1, \boldsymbol{q}_1 \rangle$$
$$= \langle \boldsymbol{\alpha}_2, \boldsymbol{q}_1 \rangle - r_{12}\langle \boldsymbol{q}_1, \boldsymbol{q}_1 \rangle = \langle \boldsymbol{\alpha}_2, \boldsymbol{q}_1 \rangle - r_{12},$$
$$r_{12} = \langle \boldsymbol{\alpha}_2, \boldsymbol{q}_1 \rangle,$$
$$|r_{22}| = |\boldsymbol{\alpha}_2 - r_{12}\boldsymbol{q}_1| = \sqrt{\langle \boldsymbol{\alpha}_2 - r_{12}\boldsymbol{q}_1, \boldsymbol{\alpha}_2 - r_{12}\boldsymbol{q}_1 \rangle},$$

5.2 方阵的正交(酉)三角分解

即
$$r_{22}=\pm\sqrt{\langle \boldsymbol{\alpha}_2-r_{12}\boldsymbol{q}_1,\boldsymbol{\alpha}_2-r_{12}\boldsymbol{q}_1\rangle},$$
$$\boldsymbol{q}_2=\frac{1}{r_{22}}(\boldsymbol{\alpha}_2-r_{12}\boldsymbol{q}_1).$$

同理,由(5.2-4)式可得 $r_{13}=\langle \boldsymbol{\alpha}_3,\boldsymbol{q}_1\rangle$, $r_{23}=\langle \boldsymbol{\alpha}_3,\boldsymbol{q}_2\rangle$,

$$r_{33}=\pm\sqrt{\langle \boldsymbol{\alpha}_3-r_{13}\boldsymbol{q}_1-r_{23}\boldsymbol{q}_2,\boldsymbol{\alpha}_3-r_{13}\boldsymbol{q}_1-r_{23}\boldsymbol{q}_2\rangle},\quad \boldsymbol{q}_3=\frac{1}{r_{33}}(\boldsymbol{\alpha}_3-r_{13}\boldsymbol{q}_1-r_{23}\boldsymbol{q}_2).$$

譬如说,r_{11}、r_{22}、r_{33}都取正值,则代入具体数字得

$$r_{11}=\sqrt{6},\quad \boldsymbol{q}_1=\frac{1}{\sqrt{6}}[1,2,1]^{\mathrm{T}};$$
$$r_{12}=\sqrt{6},\quad r_{22}=\sqrt{3},\quad \boldsymbol{q}_2=\frac{1}{\sqrt{3}}[1,-1,1]^{\mathrm{T}};$$
$$r_{13}=\frac{7}{\sqrt{6}},\quad r_{23}=\frac{1}{\sqrt{3}},\quad r_{33}=\frac{1}{\sqrt{2}},\quad \boldsymbol{q}_3=\frac{1}{\sqrt{2}}[1,0,-1]^{\mathrm{T}}.$$

因此
$$\boldsymbol{A}=\begin{bmatrix}1&2&2\\2&1&2\\1&2&1\end{bmatrix}=\underset{(\boldsymbol{Q})}{\begin{bmatrix}\frac{\sqrt{6}}{6}&\frac{\sqrt{3}}{3}&\frac{\sqrt{2}}{2}\\\frac{\sqrt{6}}{3}&-\frac{\sqrt{3}}{3}&0\\\frac{\sqrt{6}}{6}&\frac{\sqrt{3}}{3}&-\frac{\sqrt{2}}{2}\end{bmatrix}}\underset{(\boldsymbol{R})}{\begin{bmatrix}\sqrt{6}&\sqrt{6}&\frac{7\sqrt{6}}{6}\\&\sqrt{3}&\frac{\sqrt{3}}{3}\\&&\frac{\sqrt{2}}{2}\end{bmatrix}}.$$

5.2 方阵的正交(酉)三角分解

由于对任何列满秩的 $n\times r$ 矩阵 $\boldsymbol{A}$ 都可以找到 $n\times(n-r)$ 矩阵 $\boldsymbol{F}$，使得由 $\boldsymbol{A}$ 与 $\boldsymbol{F}$ 的列向量构成一个 n 阶可逆矩阵 $[\boldsymbol{A} \mid \boldsymbol{F}]$，从而对于列满秩矩阵 $\boldsymbol{A}$ 有如下分解.

定理 5.2-2 设 $\boldsymbol{A}$ 是 $n\times r$ 实(复)矩阵，且 $r(\boldsymbol{A})=r$，则存在 n 阶正交(酉)矩阵 $\boldsymbol{Q}$ 和 $n\times r$ 矩阵 $\boldsymbol{R}=\begin{bmatrix}\boldsymbol{R}_r\\ \hdashline \boldsymbol{O}\end{bmatrix}\begin{matrix}\}r\\ \}n-r\end{matrix}$，使 $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}_r$，这里 $\boldsymbol{R}_r$ 是 r 阶可逆的上三角矩阵.

证 由于 $r(\boldsymbol{A})=r$，故 $\boldsymbol{A}$ 的 r 个列向量 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_r$ 线性无关，所以可取到 $n-r$ 个向量，记为 $\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_n$，使 $\{\boldsymbol{A}_1,\cdots,\boldsymbol{A}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_n\}$ 是 $\mathbf{R}^n(\mathbf{C}^n)$ 的一个基. 对这个基施行 Schmidt 正交化方法，得到 $\mathbf{R}^n(\mathbf{C}^n)$ 的一个标准正交基 $\{\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_r,\boldsymbol{\varepsilon}_{r+1},\cdots,\boldsymbol{\varepsilon}_n\}$. 由上述讨论知，

$$\boldsymbol{B}=[\boldsymbol{A}_1 \quad \cdots \quad \boldsymbol{A}_r \mid \boldsymbol{\alpha}_{r+1} \quad \cdots \quad \boldsymbol{\alpha}_n]=[\boldsymbol{\varepsilon}_1 \quad \cdots \quad \boldsymbol{\varepsilon}_r \quad \boldsymbol{\varepsilon}_{r+1} \quad \cdots \quad \boldsymbol{\varepsilon}_n]$$

$$\cdot\left[\begin{array}{cccc:ccc}
|\boldsymbol{\beta}_1| & \langle\boldsymbol{A}_2,\boldsymbol{\varepsilon}_1\rangle & \cdots & \langle\boldsymbol{A}_r,\boldsymbol{\varepsilon}_1\rangle & \langle\boldsymbol{\alpha}_{r+1},\boldsymbol{\varepsilon}_1\rangle & \cdots & \langle\boldsymbol{\alpha}_n,\boldsymbol{\varepsilon}_1\rangle \\
 & |\boldsymbol{\beta}_2| & \cdots & \langle\boldsymbol{A}_r,\boldsymbol{\varepsilon}_2\rangle & \langle\boldsymbol{\alpha}_{r+1},\boldsymbol{\varepsilon}_2\rangle & & \langle\boldsymbol{\alpha}_n,\boldsymbol{\varepsilon}_2\rangle \\
 & & \ddots & \vdots & \vdots & & \vdots \\
 & & & |\boldsymbol{\beta}_r| & \vdots & & \vdots \\
 & & & & |\boldsymbol{\beta}_{r+1}| & & \vdots \\
 & & & & & \ddots & \\
 & & & & & & |\boldsymbol{\beta}_n|
\end{array}\right].$$

因此有

5.2 方阵的正交(酉)三角分解

$$
\begin{aligned}
A &= [A_1 \quad A_2 \quad \cdots \quad A_r] \\
&= [\varepsilon_1 \ \cdots \ \varepsilon_r \ \varepsilon_{r+1} \ \cdots \ \varepsilon_n]
\left[\begin{array}{cccc}
|\beta_1| & \langle A_2, \varepsilon_1\rangle & \cdots & \langle A_r, \varepsilon_1\rangle \\
 & |\beta_2| & \cdots & \langle A_r, \varepsilon_2\rangle \\
 & & & \vdots \\
 & & & |\beta_r| \\
\hdashline
 & & O &
\end{array}\right],
\end{aligned}
\tag{5.2-5}
$$

即 $A=Q\begin{bmatrix} R_r \\ \hdashline O \end{bmatrix}$，且 R_r 可逆. ■

(5.2-5)式也可写成如下紧凑形式：

$$A = [\varepsilon_1 \ \cdots \ \varepsilon_r] R_r = Q_r R_r, \tag{5.2-6}$$

其中 $Q_r = [\varepsilon_1 \quad \cdots \quad \varepsilon_r]$ 是 $n\times r$ 矩阵，满足

$$Q_r^{\mathrm{T}} Q_r = I_r \quad (Q_r^{\mathrm{H}} Q_r = I_r).$$

值得指出的是，当 $n>r$ 时，$Q_r Q_r^{\mathrm{T}} \neq I_n$.

(5.2-5)式和(5.2-6)式这两种形式都称为列满秩矩阵的 QR 分解，在应用中都会用到.

5.2 方阵的正交(酉)三角分解

例 1 求矩阵 $\boldsymbol{A}=\begin{bmatrix}1 & 1/2 & 5\\ 1 & -1/2 & 2\\ -1 & 1/2 & -2\\ 1 & -3/2 & 0\end{bmatrix}$ 的 QR 分解.

解 因 $\boldsymbol{A}_1=[1,1,-1,1]^{\mathrm{T}}$, $\boldsymbol{A}_2=[\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{3}{2}]^{\mathrm{T}}$, $\boldsymbol{A}_3=[5,2,-2,0]^{\mathrm{T}}$, 故 $\boldsymbol{\beta}_1=\boldsymbol{A}_1=[1,1,-1,1]^{\mathrm{T}}$, $|\boldsymbol{\beta}_1|=2$, $\boldsymbol{\varepsilon}_1=\left[\frac{1}{2},\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right]^{\mathrm{T}}$; $\langle\boldsymbol{A}_2,\boldsymbol{\varepsilon}_1\rangle=-1$, $\boldsymbol{\beta}_2=\boldsymbol{A}_2-\langle\boldsymbol{A}_2,\boldsymbol{\varepsilon}_1\rangle\boldsymbol{\varepsilon}_1=[1,0,0,-1]^{\mathrm{T}}$, $|\boldsymbol{\beta}_2|=\sqrt{2}$, $\boldsymbol{\varepsilon}_2=[\frac{1}{\sqrt{2}},0,0,-\frac{1}{\sqrt{2}}]^{\mathrm{T}}$; $\langle\boldsymbol{A}_3,\boldsymbol{\varepsilon}_1\rangle=\frac{9}{2}$, $\langle\boldsymbol{A}_3,\boldsymbol{\varepsilon}_2\rangle=\frac{5}{\sqrt{2}}$, $\boldsymbol{\beta}_3=\boldsymbol{A}_3-\langle\boldsymbol{A}_3,\boldsymbol{\varepsilon}_1\rangle\boldsymbol{\varepsilon}_1-\langle\boldsymbol{A}_3,\boldsymbol{\varepsilon}_2\rangle\boldsymbol{\varepsilon}_2=\left[\frac{1}{4},-\frac{1}{4},\frac{1}{4},\frac{1}{4}\right]^{\mathrm{T}}$, $|\boldsymbol{\beta}_3|=\frac{1}{2}$, $\boldsymbol{\varepsilon}_3=\left[\frac{1}{2},-\frac{1}{2},\frac{1}{2},\frac{1}{2}\right]^{\mathrm{T}}$. 因而, 由 (5.2-6)式得

$$\boldsymbol{A}=\begin{bmatrix}\frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{1}{2}\\ \frac{1}{2} & 0 & -\frac{1}{2}\\ -\frac{1}{2} & 0 & \frac{1}{2}\\ \frac{1}{2} & -\frac{\sqrt{2}}{2} & \frac{1}{2}\end{bmatrix}\begin{bmatrix}2 & -1 & \frac{9}{2}\\ & \sqrt{2} & \frac{5\sqrt{2}}{2}\\ & & \frac{1}{2}\end{bmatrix}.$$

■

5.2 方阵的正交(酉)三角分解

下面介绍 QR 分解的 Householder 方法，它是用一系列 Householder 变换将方阵化为上三角矩阵.

在 2.4 节中曾讲过，Householder 变换是一种正交线性变换，对于任给的两个非零向量 $\boldsymbol{x}$，$\boldsymbol{y}$，只要它们的长度相等，就一定存在 Householder 变换 $\mathscr{H}$ 使 $\boldsymbol{x}$ 变为 $\boldsymbol{y}$，其变换矩阵是 Householder 矩阵 $\boldsymbol{H}$：

$$\boldsymbol{H}=\boldsymbol{I}-\frac{2(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}}{|\boldsymbol{x}-\boldsymbol{y}|^{2}}=\boldsymbol{I}-2\,\frac{(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}}{(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{y})}. \tag{5.2-7}$$

由于 Householder 矩阵是对称正交矩阵，所以在数值计算中一般都用 Householder 变换对 $\boldsymbol{A}$ 进行 QR 分解.

5.2 方阵的正交(酉)三角分解

例 2 求将向量 $\boldsymbol{x}=[0,3,0,4]^{\mathrm{T}}$变换为 $\boldsymbol{y}=[5,0,0,0]^{\mathrm{T}}$的 Householder 矩阵.

解 由于$|\boldsymbol{x}|=5=|\boldsymbol{y}|$,所以这样的 Householder 矩阵存在. 将 $\boldsymbol{x}-\boldsymbol{y}=[-5,3,0,4]^{\mathrm{T}}$代入(5.2-7)式,得

$$\boldsymbol{H}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}-2\times\frac{1}{50}\begin{bmatrix}-5\\3\\0\\4\end{bmatrix}[-5\ 3\ 0\ 4]$$

$$=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}-\frac{1}{25}\begin{bmatrix}25&-15&0&-20\\-15&9&0&12\\0&0&0&0\\-20&12&0&16\end{bmatrix}=\begin{bmatrix}0&\frac{3}{5}&0&\frac{4}{5}\\\frac{3}{5}&\frac{16}{25}&0&-\frac{12}{15}\\0&0&1&0\\\frac{4}{5}&-\frac{12}{25}&0&\frac{9}{25}\end{bmatrix}.\quad\blacksquare$$

定理 5.2-3 任何实可逆矩阵都可用一系列 Householder 变换化为上三角矩阵.

证 对方阵的阶数 n 用数学归纳法.

当 $n=1$ 时,定理显然成立. 现假设对 $n-1$ 阶实可逆矩阵定理成立,要证对 n 阶实可逆矩阵定理也成立.

5.2 方阵的正交(酉)三角分解

设 $A=[\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\cdots\ \boldsymbol{\alpha}_n]$是 n 阶实可逆矩阵，其中 $\boldsymbol{\alpha}_j=[a_{1j},a_{2j},\cdots,a_{nj}]^{\mathrm{T}}\quad(j=1,2,\cdots,n)$ 是 A 的第 j 个列向量. 由于 $\boldsymbol{\alpha}_1\neq\mathbf{0}$，所以存在 Householder 矩阵 H_1 使

$$H_1A=\left[\begin{array}{c:c}\sigma\,|\boldsymbol{\alpha}_1| & *\ \cdots\ * \\ \hdashline \mathbf{0} & A_{n-1}\end{array}\right],$$

其中 A_{n-1}是 $n-1$ 阶实可逆矩阵，σ 是取 1 或 -1 的符号函数sign(·). 因

$$\boldsymbol{\alpha}_1-\sigma\,|\boldsymbol{\alpha}_1|\,\boldsymbol{e}_1=[a_{11}-\sigma\,|\boldsymbol{\alpha}_1|,\underbrace{a_{21},\cdots,a_{n1}}_{n-1}]^{\mathrm{T}},$$

其中 $\boldsymbol{e}_1=[1,\underbrace{0,\cdots,0}_{n-1}]^{\mathrm{T}}$，故 $\sigma=-\mathrm{sign}(a_{11})$能使 n 维向量$\boldsymbol{\alpha}_1-\sigma|\boldsymbol{\alpha}_1|\boldsymbol{e}_1$的模最大. 因此，按数值计算中应注意的一些原则(见第二部分数值计算方法)，σ 应取$-\mathrm{sign}(a_{11})$，即当 $a_{11}\geqslant0$ 时 $\sigma=-1$，而当 $a_{11}<0$ 时 $\sigma=1$.

根据归纳假设，对于 $n-1$ 阶矩阵 A_{n-1}，至多经 $n-2$ 次 Householder 变换使 A_{n-1}化为上三角矩阵，即存在至多 $n-2$ 个 Householder 矩阵的乘积，记其为 Q_{n-1}，使 $Q_{n-1}A_{n-1}=R_{n-1}$，这里 R_{n-1}是可逆的上三角实矩阵. 令 n 阶方阵 Q 为

则有

$$QH_1A=\left[\begin{array}{c:c}\sigma\,|\boldsymbol{\alpha}_1| & *\ \cdots\ * \\ \hdashline \mathbf{0} & R_{n-1}\end{array}\right],$$

且因 Q_{n-1}是若干个正交矩阵的乘积，故 Q_{n-1}是正交矩阵，从而 Q 也是正交矩阵，因此 QH_1是正交矩阵. ∎

5.2 方阵的正交(酉)三角分解

例 3 用一系列 Householder 变换求方阵

$$\boldsymbol{A}=\begin{bmatrix}2&2&1\\1&2&2\\2&1&2\end{bmatrix}$$

的 QR 分解.

解 由于 $\boldsymbol{\alpha}_1=[2,1,2]^{\mathrm{T}}$，且 $a_{11}=2>0$，$|\boldsymbol{\alpha}_1|=3$，故可将 $\boldsymbol{x}=[2,1,2]^{\mathrm{T}}$，$\boldsymbol{y}=[-3,0,0]^{\mathrm{T}}$代入(5.2-7)式得到 Householder 矩阵 $\boldsymbol{H}_1$. 由于 $\boldsymbol{x}-\boldsymbol{y}=[5,1,2]^{\mathrm{T}}$，故有

$$\boldsymbol{H}_1=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}-\frac{2}{30}\begin{bmatrix}5\\1\\2\end{bmatrix}\begin{bmatrix}5&1&2\end{bmatrix}$$

$$=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}-\frac{1}{15}\begin{bmatrix}25&5&10\\5&1&2\\10&2&4\end{bmatrix}=\frac{1}{15}\begin{bmatrix}-10&-5&-10\\-5&14&-2\\-10&-2&11\end{bmatrix},$$

从而

$$\boldsymbol{H}_1\boldsymbol{A}=\frac{1}{15}\begin{bmatrix}-10&-5&-10\\-5&14&-2\\-10&-2&11\end{bmatrix}\begin{bmatrix}2&2&1\\1&2&2\\2&1&2\end{bmatrix}=\begin{bmatrix}-3&-\frac{8}{3}&-\frac{8}{3}\\0&\frac{16}{15}&\frac{19}{15}\\0&-\frac{13}{15}&\frac{8}{15}\end{bmatrix}.$$

5.2 方阵的正交(酉)三角分解

再令 $\boldsymbol{x}=\left[-\frac{8}{3},\frac{16}{15},-\frac{13}{15}\right]^{\mathrm{T}}$，$\boldsymbol{y}=\left[-\frac{8}{3},a,0\right]^{\mathrm{T}}$，则由 $|\boldsymbol{x}|=|\boldsymbol{y}|$ 得 $|a|=\sqrt{\left(\frac{16}{15}\right)^2+\left(\frac{13}{15}\right)^2}=\frac{\sqrt{17}}{3}$，又因 $\frac{16}{15}>0$，故取 $a=-\frac{\sqrt{17}}{3}$. 从而 $\boldsymbol{x}-\boldsymbol{y}=\left[0,\frac{16+5\sqrt{17}}{15},\frac{-13}{15}\right]^{\mathrm{T}}$. 将其代入(5.2-7)式得

$$\boldsymbol{H}_2=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}-\frac{2\times(15)^2}{(16+5\sqrt{17})^2+(-13)^2}\times\begin{bmatrix}0\\\frac{16+5\sqrt{17}}{15}\\-\frac{13}{15}\end{bmatrix}\begin{bmatrix}0&\frac{16+5\sqrt{17}}{15}&-\frac{13}{15}\end{bmatrix}$$

$$=\left[\begin{array}{c:cc}1&0&0\\\hdashline 0&&\\&\boldsymbol{Q}_2&\\0&&\end{array}\right],$$

其中

$$\boldsymbol{Q}_2=\frac{1}{(16+5\sqrt{17})^2+169}\times\begin{bmatrix}169-(16+5\sqrt{17})^2&26\times(16+5\sqrt{17})\\26\times(16+5\sqrt{17})&(16+5\sqrt{17})^2-169\end{bmatrix}.$$

（实际上，$\boldsymbol{Q}_2$ 是将向量 $\left[\frac{16}{15},-\frac{13}{15}\right]^{\mathrm{T}}$ 变为向量 $\left[-\frac{\sqrt{17}}{3},0\right]^{\mathrm{T}}$ 的 Householder 矩阵.）

5.2 方阵的正交(酉)三角分解

从而

$$H_2H_1A=\left[\begin{array}{c|cc}1 & 0 & 0\\ \hline 0 & & \\ & Q_2 & \\ 0 & & \end{array}\right]\left[\begin{array}{c|cc}-3 & -\frac{8}{3} & -\frac{8}{3}\\ \hline & \frac{16}{15} & \frac{19}{15}\\ & -\frac{13}{15} & \frac{8}{15}\end{array}\right]=\left[\begin{array}{c|cc}-3 & -\frac{8}{3} & -\frac{8}{3}\\ \hline & -\frac{\sqrt{17}}{3} & -\frac{8\sqrt{17}}{51}\\ & & \frac{5\sqrt{17}}{17}\end{array}\right].$$

因此 A 的 QR 分解是

$$A=(H_2H_1)^{-1}\begin{bmatrix}-3 & -\frac{8}{3} & -\frac{8}{3}\\ & -\frac{\sqrt{17}}{3} & -\frac{8\sqrt{17}}{51}\\ & & \frac{5\sqrt{17}}{17}\end{bmatrix}=H_1H_2\begin{bmatrix}-3 & -\frac{8}{3} & -\frac{8}{3}\\ & -\frac{\sqrt{17}}{3} & -\frac{8\sqrt{17}}{51}\\ & & \frac{5\sqrt{17}}{17}\end{bmatrix}$$

$$=\begin{bmatrix}-\frac{2}{3} & -\frac{2\sqrt{17}}{51} & -\frac{3\sqrt{17}}{17}\\ -\frac{1}{3} & -\frac{10\sqrt{17}}{51} & \frac{2\sqrt{17}}{17}\\ -\frac{2}{3} & \frac{7\sqrt{17}}{51} & \frac{2\sqrt{17}}{17}\end{bmatrix}\begin{bmatrix}-3 & -\frac{8}{3} & -\frac{8}{3}\\ & -\frac{\sqrt{17}}{3} & -\frac{8\sqrt{17}}{51}\\ & & \frac{5\sqrt{17}}{17}\end{bmatrix}. \qquad \blacksquare$$

5.2　方阵的正交(酉)三角分解

我们将在第二部分数值计算方法中说明求方阵特征值的 QR 算法，以及 QR 分解在解最小二乘问题中的应用．这里就解线性方程组说明 QR 分解的应用．

如果已知可逆矩阵 $\boldsymbol{A}$ 的 QR 分解，那么线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 等价于下述方程组：

$$\begin{cases} \boldsymbol{Qy} = \boldsymbol{b}, \\ \boldsymbol{Rx} = \boldsymbol{y}. \end{cases} \tag{5.2-8}$$

显然，(5.2-8)式中第一个方程组的解是

$$\boldsymbol{y}=\boldsymbol{Q}^{-1}\boldsymbol{b}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b},$$

而第二个方程组与前相同，可用回代过程解之，得到

$$\boldsymbol{x}=\boldsymbol{R}^{-1}\boldsymbol{y}=\boldsymbol{R}^{-1}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}.$$

5.3 矩阵的奇异值分解

已经在 2.3 节介绍过，一个实(复)方阵 $\boldsymbol{A}$ 可正交(酉)相似对角化的充要条件是 $\boldsymbol{A}$ 为正规矩阵，即 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$(今后，为了一般起见，我们不分是实矩阵 $\boldsymbol{A}$ 的转置还是复矩阵 $\boldsymbol{A}$ 的共轭转置，都用 $\boldsymbol{A}^{\mathrm{H}}$ 表示之)，因此，任一实(复)方阵 $\boldsymbol{A}$ 未必能正交(酉)相似于对角矩阵. 然而，我们将证明存在两个正交(酉)矩阵 $\boldsymbol{U}$ 和 $\boldsymbol{V}$，使 $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{V}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}$ 为对角矩阵.

更一般地，对于秩为 r 的任何 $m\times n$ 实(复)矩阵，一定存在 m 阶正交(酉)矩阵 $\boldsymbol{U}$ 和 n 阶正交(酉)矩阵 $\boldsymbol{V}$，使得

$$\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{V}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}=\begin{array}{c}\left[\begin{array}{c:c}\boldsymbol{\Sigma}_r & \boldsymbol{O}\\ \hdashline \underbrace{\boldsymbol{O}}_{r} & \underbrace{\boldsymbol{O}}_{n-r}\end{array}\right]\begin{array}{l}\}r\\ \}n-r\end{array}\end{array}, \tag{5.3-1}$$

其中 $\boldsymbol{\Sigma}_r$ 是 r 阶可逆的对角矩阵. (5.3-1)式可以改写为

$$\boldsymbol{A}=\boldsymbol{U}\left[\begin{array}{c:c}\boldsymbol{\Sigma}_r & \boldsymbol{O}\\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right]\boldsymbol{V}^{-1}=\boldsymbol{U}\left[\begin{array}{c:c}\boldsymbol{\Sigma}_r & \boldsymbol{O}\\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right]\boldsymbol{V}^{\mathrm{H}}, \tag{5.3-1$'$}$$

(5.3-1′)式称为 $\boldsymbol{A}$ 的奇异值分解，它在线性系统理论、最小二乘问题、广义逆矩阵、实验数据处理等方面都有重要应用.

首先介绍一下方阵 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的一些性质.

5.3 矩阵的奇异值分解

定理 5.3-1 设 $\boldsymbol{A}$ 是秩为 r 的 $m\times n$ 复矩阵. 则

(1) $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 分别是 n 阶与 m 阶半正定 Hermite 矩阵；

(2) $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 有相同的正特征值，且 $r(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=r(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})=r$.

证 (1) 由于 $(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{A}^{\mathrm{H}})^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$，故 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 是 n 阶 Hermite 矩阵. 又因对任一 n 维复向量 $\boldsymbol{x}$，有

$$\boldsymbol{x}^{\mathrm{H}}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})\boldsymbol{x}=(\boldsymbol{A}\boldsymbol{x})^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{x})=\overline{(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}\overline{(\boldsymbol{A}\boldsymbol{x})}}=\overline{\langle\boldsymbol{A}\boldsymbol{x},\boldsymbol{A}\boldsymbol{x}\rangle}=\langle\boldsymbol{A}\boldsymbol{x},\boldsymbol{A}\boldsymbol{x}\rangle\geqslant 0,$$

故 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 是 n 阶半正定 Hermite 矩阵.

同理可证 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 是 m 阶半正定 Hermite 矩阵.

(2) 不妨设 $m\geqslant n$，则由 1.5 节中的定理 1.5-4，得

$$\det(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}-\lambda\boldsymbol{I}_m)=(-\lambda)^{m-n}\det(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}-\lambda\boldsymbol{I}_n),$$

因此 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 有相同的非零特征值. 又因它们都是半正定 Hermite 矩阵，故其特征值都是非负的实数，从而它们有相同的正特征值.

由于 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 和 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 都是 Hermite 矩阵，所以它们分别酉相似于 n 阶和 m 阶对角矩阵，且主对角线元素是其特征值，故由它们有相同的非零特征值得知 $r(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=r(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})$.

又因为由 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 可得 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$. 反之若 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，即 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，所以方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 等价于方程组 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，从而 $r(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=r(\boldsymbol{A})$. 因此有

$$r(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=r(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})=r. \qquad \blacksquare$$

5.3 矩阵的奇异值分解

定义 5.3-1 设 $\boldsymbol{A}$ 是秩为 $r>0$ 的 $m\times n$ 矩阵，则半正定 Hermite 矩阵 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的 n 个特征值 $\lambda_i\geqslant 0(i=1,2,\cdots,n)$ 的算术平方根 $\sigma_i=\sqrt{\lambda_i}$ 称为 $\boldsymbol{A}$ 的**奇异值**.

例如，对于 $\boldsymbol{A}=\begin{bmatrix}1&1\\1&-1\\2&3\end{bmatrix}$，则 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\begin{bmatrix}1&1&2\\1&-1&3\end{bmatrix}\begin{bmatrix}1&1\\1&-1\\2&3\end{bmatrix}=\begin{bmatrix}6&6\\6&11\end{bmatrix}$，

$$\det(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}-\boldsymbol{I})=\begin{vmatrix}6-\lambda&6\\6&11-\lambda\end{vmatrix}=(15-\lambda)(2-\lambda),$$

因而 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的特征值是 $\lambda_1=15,\lambda_2=2$. 于是 $\boldsymbol{A}$ 的奇异值是 $\sigma_1=\sqrt{\lambda_1}=\sqrt{15},\sigma_2=\sqrt{\lambda_2}=\sqrt{2}$.

例 1 设 $\boldsymbol{A}$、$\boldsymbol{B}$ 都是 $m\times n$ 矩阵，且有 m 阶酉(正交)矩阵 $\boldsymbol{U}$ 和 n 阶酉(正交)矩阵 $\boldsymbol{V}$，使得

$$\boldsymbol{B}=\boldsymbol{U}\boldsymbol{A}\boldsymbol{V},$$

(这时称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉(正交)等价.)证明 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的奇异值.

证 由 $\boldsymbol{B}=\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}$ 得

$$\boldsymbol{B}^{\mathrm{H}}=(\boldsymbol{U}\boldsymbol{A}\boldsymbol{V})^{\mathrm{H}}=\boldsymbol{V}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}},$$

从而有

$$\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}=\boldsymbol{V}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\cdot\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}=\boldsymbol{V}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V},$$

即 $\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}$ 与 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 酉相似. 于是，$\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}$ 与 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 有相同的特征值，因而它们也有相同的奇异值. ∎

5.3 矩阵的奇异值分解

定理 5.3-2 设 $\boldsymbol{A}$ 为 n 阶可逆矩阵，则存在 n 阶酉矩阵 $\boldsymbol{Q}_1$、$\boldsymbol{Q}_2$，使得

$$\boldsymbol{Q}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q}_2 = \mathrm{diag}(\sigma_1, \cdots, \sigma_n), \tag{5.3-2}$$

其中 $\sigma_i > 0 (i = 1, 2, \cdots, n)$ 是 $\boldsymbol{A}$ 的奇异值.

证 因 $\boldsymbol{A}$ 可逆，故 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 为正定 Hermite 矩阵，因而存在酉矩阵 $\boldsymbol{Q}_2$ 使

$$\boldsymbol{Q}_2^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q}_2 = \mathrm{diag}(\lambda_1, \cdots, \lambda_n), \tag{5.3-3}$$

其中 $\lambda_i (i = 1, \cdots, n)$ 是 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的特征值，且 $\lambda_i > 0$.

令 n 阶可逆矩阵 $\boldsymbol{\Sigma}$ 为

$$\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \cdots, \sigma_n),$$

其中 $\sigma_i = \sqrt{\lambda_i} > 0 (i = 1, \cdots, n)$ 是 $\boldsymbol{A}$ 的奇异值，则有 $\boldsymbol{\Sigma}^{\mathrm{H}}\boldsymbol{\Sigma} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$，从而由(5.3-3)式得

$$\boldsymbol{Q}_2^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q}_2 = \boldsymbol{\Sigma}^{\mathrm{H}}\boldsymbol{\Sigma}. \tag{5.3-4}$$

令

$$\boldsymbol{Q}_1 = \boldsymbol{A}\boldsymbol{Q}_2\boldsymbol{\Sigma}^{-1}.$$

则由(5.3-4)式知

$$\boldsymbol{Q}_1^{\mathrm{H}}\boldsymbol{Q}_1 = (\boldsymbol{A}\boldsymbol{Q}_2\boldsymbol{\Sigma}^{-1})^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{Q}_2\boldsymbol{\Sigma}^{-1}) = (\boldsymbol{\Sigma}^{-1})^{\mathrm{H}}\boldsymbol{Q}_2^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q}_2\boldsymbol{\Sigma}^{-1} = (\boldsymbol{\Sigma}^{\mathrm{H}})^{-1}\boldsymbol{Q}_2^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q}_2\boldsymbol{\Sigma}^{-1} = \boldsymbol{I}_n,$$

即 $\boldsymbol{Q}_1$ 是酉矩阵，并且

$$\boldsymbol{Q}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q}_2 = (\boldsymbol{A}\boldsymbol{Q}_2\boldsymbol{\Sigma}^{-1})^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q}_2 = (\boldsymbol{\Sigma}^{\mathrm{H}})^{-1}\boldsymbol{Q}_2^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q}_2 = \boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \cdots, \sigma_n). \quad \blacksquare$$

我们称(5.3-2)式为可逆矩阵 $\boldsymbol{A}$ 的酉对角分解.

5.3 矩阵的奇异值分解

例 2 求可逆矩阵 $\boldsymbol{A}=\begin{bmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 1 & -\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}}\end{bmatrix}$的正交对角分解.

解 由于

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{bmatrix}\frac{3}{2} & 0 & -\frac{1}{2}\\ 0 & 2 & 0\\ -\frac{1}{2} & 0 & \frac{3}{2}\end{bmatrix}$$

的特征值是 $\lambda_1=\lambda_2=2,\lambda_3=1$，故 $\boldsymbol{A}$ 的奇异值是 $\sigma_1=\sigma_2=\sqrt{2},\sigma_3=1$.

$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的属于二重特征值 2 的两个相互正交的单位特征向量可取为

$$\boldsymbol{q}_1=\left[-\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right]^{\mathrm{T}},\quad \boldsymbol{q}_2=[0,1,0]^{\mathrm{T}}.$$

$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的属于特征值 1 的单位特征向量可取为

$$\boldsymbol{q}_3=\left[\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right]^{\mathrm{T}}.$$

5.3 矩阵的奇异值分解

令
$$Q_2=[\boldsymbol{q}_1\ \boldsymbol{q}_2\ \boldsymbol{q}_3]=\begin{bmatrix}-\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ 0 & 1 & 0\\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{bmatrix},$$

则 Q_2 是正交矩阵，且有
$$Q_2^{\mathrm{T}}A^{\mathrm{T}}AQ_2=\begin{bmatrix}2 & & \\ & 2 & \\ & & 1\end{bmatrix}=\begin{bmatrix}\sqrt{2} & & \\ & \sqrt{2} & \\ & & 1\end{bmatrix}\begin{bmatrix}\sqrt{2} & & \\ & \sqrt{2} & \\ & & 1\end{bmatrix}.$$

再令
$$Q_1=AQ_2\begin{bmatrix}\sqrt{2} & & \\ & \sqrt{2} & \\ & & 1\end{bmatrix}^{-1}=\begin{bmatrix}0 & 0 & 1\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\end{bmatrix},$$

则 Q_1 是正交矩阵，且容易验证
$$Q_1^{\mathrm{T}}AQ_2=\mathrm{diag}(\sqrt{2},\sqrt{2},1).\qquad\blacksquare$$

对于一般的 $m\times n$ 矩阵，我们有下述定理：

5.3 矩阵的奇异值分解

定理 5.3-3 设 $\boldsymbol{A}$ 是秩为 r 的 $m\times n$ 矩阵，则存在 m 阶酉矩阵 $\boldsymbol{U}$ 和 n 阶酉矩阵 $\boldsymbol{V}$ 使

$$\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{V}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}=\left[\begin{array}{c|c}\boldsymbol{\Sigma}_r & \boldsymbol{O}\\ \hline \boldsymbol{O} & \boldsymbol{O}\end{array}\right],$$

其中 $\boldsymbol{\Sigma}_r=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_r)$，且 $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_r>0$ 是 $\boldsymbol{A}$ 的正奇异值.

证 由于 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 是半正定 Hermite 矩阵，其秩为 r，故 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的特征值是非负实数，从而可设它的 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 按大小顺序为

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r>0,\quad \lambda_{r+1}=\cdots=\lambda_n=0,$$

对应的相互正交的单位特征向量依次为 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$，则

$$\boldsymbol{V}=[\boldsymbol{v}_1\ \boldsymbol{v}_2\ \cdots\ \boldsymbol{v}_r \mid \boldsymbol{v}_{r+1}\cdots\ \boldsymbol{v}_n]=[\underbrace{\boldsymbol{V}_r}_{n\times r} \mid \underbrace{\boldsymbol{V}_{n-r}}_{n\times(n-r)}]$$

是 n 阶酉矩阵，且有

$$\boldsymbol{V}^{-1}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}=\boldsymbol{V}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}=\left[\begin{array}{c}\boldsymbol{V}_r^{\mathrm{H}}\\ \hline \boldsymbol{V}_{n-r}^{\mathrm{H}}\end{array}\right]\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}[\boldsymbol{V}_r \mid \boldsymbol{V}_{n-r}]$$

$$=\left[\begin{array}{c|c}\boldsymbol{V}_r^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}_r & \boldsymbol{V}_r^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}_{n-r}\\ \hline \boldsymbol{V}_{n-r}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}_r & \boldsymbol{V}_{n-r}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}_{n-r}\end{array}\right]=\begin{array}{c}\left[\begin{array}{c|c}\begin{matrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2\end{matrix} & \boldsymbol{O}\\ \hline \underbrace{\boldsymbol{O}}_{r} & \underbrace{\boldsymbol{O}}_{n-r}\end{array}\right]\begin{array}{l}\}\,r\\ \}\,n-r\end{array}\end{array}\qquad(5.3\text{-}5)$$

5.3 矩阵的奇异值分解

其中 $V_r=[v_1\ v_2 \cdots\ v_r]$ 是 $n\times r$ 矩阵，$V_{n-r}=[v_{r+1} \cdots v_n]$ 是 $n\times(n-r)$ 矩阵，$\sigma_i=\sqrt{\lambda_i}\,(i=1,2,\cdots,r)$ 是 A 的正奇异值.

由(5.3-5)式得

$$\begin{cases} V_r^{\mathrm{H}}A^{\mathrm{H}}AV_r=\Sigma_r^2, \\ V_r^{\mathrm{H}}A^{\mathrm{H}}AV_{n-r}=O_{r\times(n-r)}, \\ V_{n-r}^{\mathrm{H}}A^{\mathrm{H}}AV_r=O_{(n-r)\times r}, \\ V_{n-r}^{\mathrm{H}}A^{\mathrm{H}}AV_{n-r}=O_{(n-r)\times(n-r)}, \end{cases} \tag{5.3-6}$$

式中 $\Sigma_r=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_r)$ 是 r 阶可逆对角矩阵.

于是，由(5.3-6)式的第一式知

$$\Sigma_r^{-1}V_r^{\mathrm{H}}A^{\mathrm{H}}AV_r\Sigma_r^{-1}=(\Sigma_r^{-1})^{\mathrm{H}}V_r^{\mathrm{H}}A^{\mathrm{H}}AV_r\Sigma_r^{-1}=I_r.$$

令 $m\times r$ 矩阵 U_r 为

$$U_r=AV_r\Sigma_r^{-1}, \tag{5.3-7}$$

则有 $U_r^{\mathrm{H}}U_r=I_r$，故 U_r 的 r 个列向量是两两正交的单位向量. 因而存在 $m\times(m-r)$ 矩阵 U_{m-r} 使

$$U=[U_r \,\vdots\, U_{m-r}] \tag{5.3-8}$$

是 m 阶酉矩阵，且有

$$U^{-1}AV=U^{\mathrm{H}}AV=\begin{bmatrix} U_r^{\mathrm{H}} \\ \hdashline U_{m-r}^{\mathrm{H}} \end{bmatrix}A[V_r \,\vdots\, V_{n-r}]=\begin{array}{c} \left[\begin{array}{c:c} U_r^{\mathrm{H}}AV_r & U_r^{\mathrm{H}}AV_{n-r} \\ \hdashline \underbrace{U_{m-r}^{\mathrm{H}}AV_r}_{r} & \underbrace{U_{m-r}^{\mathrm{H}}AV_{n-r}}_{n-r} \end{array}\right] \begin{array}{l} \}r \\ \}m-r \end{array} \end{array}. \tag{5.3-9}$$

5.3　矩阵的奇异值分解

但由(5.3-6)式的第四式得$(\boldsymbol{AV}_{n-r})^{\mathrm{H}}(\boldsymbol{AV}_{n-r})=\boldsymbol{O}_{(n-r)\times(n-r)}$，故$\boldsymbol{AV}_{n-r}=\boldsymbol{O}_{m\times(n-r)}$（为什么？），而由(5.3-7)式得

$$\boldsymbol{U}_r^{\mathrm{H}}\boldsymbol{AV}_r=\boldsymbol{U}_r^{\mathrm{H}}\boldsymbol{U}_r\boldsymbol{\Sigma}_r=\boldsymbol{\Sigma}_r.$$

又因$\boldsymbol{U}$是酉矩阵，它的m个列向量是两两正交的单位向量，故有$\boldsymbol{U}_{m-r}^{\mathrm{H}}\boldsymbol{U}_r=\boldsymbol{O}_{(m-r)\times r}$，从而

$$\boldsymbol{U}_{m-r}^{\mathrm{H}}\boldsymbol{AV}_r=\boldsymbol{U}_{m-r}^{\mathrm{H}}\boldsymbol{U}_r\boldsymbol{\Sigma}_r=\boldsymbol{O}_{(m-r)\times r}.$$

因此，(5.3-9)式给出

$$\boldsymbol{U}^{-1}\boldsymbol{AV}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{AV}=\left[\begin{array}{c:c}\boldsymbol{\Sigma}_r & \boldsymbol{O}\\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right].$$

■

例 3　求矩阵$\boldsymbol{A}$的奇异值分解，其中

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

解　$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\begin{bmatrix}2 & 0 & -2\\ 0 & 0 & 0\\ -2 & 0 & 2\end{bmatrix}$的特征值为$\lambda_1=4,\lambda_2=\lambda_3=0$，故$\boldsymbol{A}$的正奇异值是$\sigma_1=\sqrt{4}=2$.

$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$的属于$\lambda_1=4$的单位特征向量$\boldsymbol{v}_1=\left[\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}}\right]^{\mathrm{T}}$，属于$\lambda_2=\lambda_3=0$的两个正交的单位特征向量可取为$\boldsymbol{v}_2=[0\ 1\ 0]^{\mathrm{T}}$和$\boldsymbol{v}_3=\left[\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right]^{\mathrm{T}}$. 于是

5.3 矩阵的奇异值分解

$$\boldsymbol{V}=[\boldsymbol{v}_1 \mid \boldsymbol{v}_2 \quad \boldsymbol{v}_3]=[\boldsymbol{V}_1 \mid \boldsymbol{V}_2]=\left[\begin{array}{c:cc} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{array}\right],$$

$$\boldsymbol{U}_1=\boldsymbol{A}\boldsymbol{V}_1\boldsymbol{\Sigma}^{-1}=\frac{1}{2}\boldsymbol{A}\boldsymbol{V}_1=\left[\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right]^{\mathrm{T}}.$$

取 $\boldsymbol{U}_2=\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$，则 $\boldsymbol{U}=[\boldsymbol{U}_1 \mid \boldsymbol{U}_2]=\left[\begin{array}{c:c} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right]$为正交矩阵．容易验证

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}=\left[\begin{array}{cc} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \hdashline -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right]\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}\left[\begin{array}{c:cc} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{array}\right]=\left[\begin{array}{c:cc} 2 & 0 & 0 \\ \hdashline 0 & 0 & 0 \end{array}\right].$$

∎

5.3 矩阵的奇异值分解

从定理 5.3-3 的证明过程中可以看出，$\boldsymbol{V}$ 的 n 个列向量是 n 阶半正定 Hermite 矩阵 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的两两正交的单位特征向量，亦称它们为 $\boldsymbol{A}$ 的**右奇异向量**，且 $\boldsymbol{V}$ 的前 r 个列向量分别是 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的属于 r 个非零特征值的两两正交的单位特征向量. 又由 $\boldsymbol{U}_r=\boldsymbol{A}\boldsymbol{V}_r\boldsymbol{\Sigma}_r^{-1}$ 得到

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}_r=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}_r\boldsymbol{\Sigma}_r^{-1}=\boldsymbol{A}\boldsymbol{V}_r\boldsymbol{\Sigma}_r^{2}\boldsymbol{\Sigma}_r^{-1}=\boldsymbol{A}\boldsymbol{V}_r\boldsymbol{\Sigma}_r=\boldsymbol{U}_r\boldsymbol{\Sigma}_r^{2}. \tag{5.3-10}$$

(5.3-10)式表明 $\boldsymbol{U}_r$ 的 r 个列向量恰分别是 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的属于其特征值 $\sigma_1^2=\lambda_1,\cdots,\sigma_r^2=\lambda_r$ 的两两正交的单位特征向量. 这也说明了 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 有相同的非零特征值，且属于特征值 $\lambda_i>0(i=1,2,\cdots,r)$ 的单位特征向量 $\boldsymbol{u}_i$ 与 $\boldsymbol{v}_i$ 之间有如下关系

$$\boldsymbol{u}_i=\sigma_i^{-1}\boldsymbol{A}\boldsymbol{v}_i \quad (i=1,2,\cdots,r).$$

由于从(5.3-10)式可得 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{u}_i=\sigma_i^2\boldsymbol{u}_i$，从而

$$\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{u}_i)^{\mathrm{H}}=(\sigma_i^2\boldsymbol{u}_i)^{\mathrm{H}}=\sigma_i^2\boldsymbol{u}_i^{\mathrm{H}}(i=1,2,\cdots,r),$$

故亦称 $\boldsymbol{u}_i(i=1,2,\cdots,r)$ 为 $\boldsymbol{A}$ 的**左奇异向量**.

不难验证如下的结果：

(1) Hermite 矩阵(即 $\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$)的奇异值 σ_i 为其特征值 λ_i 的绝对值，即 $\sigma_i=|\lambda_i|(i=1,2,\cdots,n)$，其中 λ_i 是 n 阶方阵 $\boldsymbol{A}$ 的特征值；

(2) $\|\boldsymbol{A}\|_2=\sigma_1$，这里 σ_1 是 $\boldsymbol{A}$ 的最大奇异值；

(3) $\boldsymbol{A}$ 的 Frobenius 范数 $\|\boldsymbol{A}\|_F=\sqrt{\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})}=\sqrt{\sum_{i=1}^{n}\sigma_i^2}$，其中 $\sigma_1,\cdots,\sigma_n$ 是 $m\times n$ 矩阵 $\boldsymbol{A}$ 的奇异值.

5.3 矩阵的奇异值分解

例 4 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{B}$ 是 $n\times s$ 矩阵，记 $\sigma_1(\boldsymbol{X})$ 为任一矩阵 $\boldsymbol{X}$ 的最大奇异值，证明

$$\sigma_1(\boldsymbol{AB})\leqslant\sigma_1(\boldsymbol{A})\cdot\sigma_1(\boldsymbol{B}). \tag{5.3-11}$$

证 因 $\|\boldsymbol{AB}\|_2\leqslant\|\boldsymbol{A}\|_2\cdot\|\boldsymbol{B}\|_2$，而 $\|\boldsymbol{A}\|_2=\sigma_1(\boldsymbol{A})$，$\|\boldsymbol{B}\|_2=\sigma_1(\boldsymbol{B})$，$\|\boldsymbol{AB}\|_2=\sigma_1(\boldsymbol{AB})$，故有(5.3-11)式. ■

类似地，若 $\boldsymbol{A}$、$\boldsymbol{B}$ 都是 $m\times n$ 矩阵，则有

$$\sigma_1(\boldsymbol{A}+\boldsymbol{B})\leqslant\sigma_1(\boldsymbol{A})+\sigma_1(\boldsymbol{B}). \tag{5.3-12}$$

请读者自证(5.3-12)式.

将 m 阶酉矩阵 $\boldsymbol{U}$ 和 n 阶酉矩阵 $\boldsymbol{V}$ 分块为

$$\boldsymbol{U}=[\underbrace{\boldsymbol{U}_r}_{r}\ \vdots\ \underbrace{\boldsymbol{U}_{m-r}}_{m-r}],\boldsymbol{V}=[\underbrace{\boldsymbol{V}_r}_{r}\ \vdots\ \underbrace{\boldsymbol{V}_{n-r}}_{n-r}],$$

则由(5.3-1′)式可得

$$\boldsymbol{A}=\boldsymbol{U}_r\boldsymbol{\Sigma}\boldsymbol{V}_r^{\mathrm{H}},\quad \boldsymbol{A}=\sum_{i=1}^{r}\sigma_i\boldsymbol{u}_i\boldsymbol{v}_i^{\mathrm{H}}. \tag{5.3-13}$$

我们称(5.3-13)式为 $\boldsymbol{A}$ 的奇异值分解的紧凑形式，它们在线性多变量控制系统中有重要应用[①].

作业（第2部分）

习题 5.1

2. 证明 $A=\begin{bmatrix} 4 & -1 & 1 \\ -1 & 4.25 & 2.75 \\ 1 & 2.75 & 3.5 \end{bmatrix}$ 是正定矩阵，并求它的 LDL^{T} 分解和 Cholesky 分解.

习题 5.2

3. 设

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

证明方程组 $Ax=b$ 有解，并用 QR 分解方法解方程组 $Ax=b$.

习题 5.3

2. 求矩阵 $A=\begin{bmatrix} 2 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}$ 的奇异值及奇异值分解.

謝謝觀看！

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黄 煒
博士，副教授
Dr. Wei Huang