第一章　线性空间上的线性算子

§1　线性空间

1.1　线性空间的定义及基本性质

一、数环与数域

定义 1-1 设 Z 为非空数集且其中任何两个相同或互异的数之和、差与积仍属于 Z（即数集关于加、减、乘法运算封闭），则说 Z 是一个数环.

根据数环的定义有：

1° 只含一个 0 的数集 $Z=\{0\}$ 显然是个数环，而且是最小的数环；

2° 任何数环 Z 必含有 0．因为若 $a\in Z$，则 $a-a=0\in Z$；

3° 若 $a\in Z$，则 $-a\in Z$．因为 $0-a=-a\in Z$.

定义 1-2 如果 P 是至少含有两个互异数的数环，并且其中任何两个数 a 与 b 之商（$b\neq 0$）仍属于 P（换言之，数集关于四则运算都封闭），则说 P 是一个数域.

根据数域的定义有：

1° 任何数域 P 中必含有 0 与 1，因为 P 中至少有一个数 $a\neq 0$，而 $a/a=1\in P$；

2° 若 $a\neq 0$，则 $1/a=a^{-1}\in P$.

★ 全体整数（包括 0）组成一个数域；

全体有理数组成一个数域，叫做有理数域，记为 Q；

全体实数组成一个数域，叫做实数域，记为 R；
全体复数组成一个数域，叫做复数域，记为 C.

二、线性空间

定义 1-3 设 V 是一个非空集合[★]，P 是一个数域．如果 V 满足如下两个条件：

1．在 V 中定义一个封闭的加法运算，即当 $\vec{x},\vec{y}\in V$ 时，有唯一的和 $\vec{x}+\vec{y}\in V$，并且加法运算满足四条性质：

(1) $\vec{x}+\vec{y}=\vec{y}+\vec{x}$（交换律）；

★ 前提条件，证明时不可缺.

(2) $\vec{x}+(\vec{y}+\vec{z})=(\vec{x}+\vec{y})+\vec{z}$（结合律）；

(3) 存在零元素$\vec{0}\in V$，对于V中任何一个元素$\vec{x}$都有$\vec{x}+\vec{0}=\vec{x}$*；

(4) 存在负元素，即对任一元素$\vec{x}\in V$，存在一元素$\vec{y}\in V$，使$\vec{x}+\vec{y}=\vec{0}$，且称$\vec{y}$为$\vec{x}$的负元素，记为$-\vec{x}$，于是有$\vec{x}+(-\vec{x})=\vec{0}$.

2. 在V中定义一个封闭的数乘运算（数与元素的乘法），即当$\vec{x}\in V$，$\lambda\in P$时，有唯一的$\lambda\vec{x}\in V$，且数乘运算满足四条性质：

(5) $(\lambda+\mu)\vec{x}=\lambda\vec{x}+\mu\vec{x}$（分配律）；

(6) $\lambda(\vec{x}+\vec{y})=\lambda\vec{x}+\lambda\vec{y}$（数因子分配律）；

(7) $\lambda(\mu\vec{x})=(\lambda\mu)\vec{x}$（结合律）；

(8) $1\cdot\vec{x}=\vec{x}$.

其中$\vec{x},\vec{y},\vec{z}$表示V中任意元素；λ,μ是数域P中任意数；1 是数域P中的单位数.

这时，我们说V是数域P上的线性空间.

★ 不是线性空间的例子：

① 次数等于$n(n\geq 1)$的多项式的集合，关于通常的多项式加法与数乘运算是不能构成线性空间的．举例：$f(x)=x^n+x,g(x)=-x^n+1$，则$f(x)+g(x)=x+1$不属于原来集合；

② 平面上全体向量组成的集合，对于通常意义下的向量加法和如下定义的数乘$k\cdot\alpha=\vec{0}$虽然对两种运算都封闭，但不满足运算规律（8）.

例：设数域为R，集合为$V=\{\vec{\alpha}\,|\,\vec{\alpha}=(\xi_1,\xi_2),\xi_i\in R\}$． 对于$\vec{\alpha}=(\xi_1,\xi_2),\vec{\beta}=(\eta_1,\eta_2)$及$k\in R$，指定两种线性运算如下：

（1）加法运算 $\vec{\alpha}+\vec{\beta}=(\xi_1+\eta_1,\xi_2+\eta_2)$ 数乘运算 $k\circ\vec{\alpha}=(k\xi_1,\xi_2)$

（2）加法运算 $\vec{\alpha}\oplus\vec{\beta}=(\xi_1+\eta_1,\xi_2+\eta_2+\xi_1\eta_1)$

数乘运算 $k\odot\vec{\alpha}=(k\xi_1,k\xi_2+\frac{1}{2}k(k-1)\xi_1^2)$

分别判断V是否构成R上的线性空间.

---

* 确定线性空间中零元素的方法：设$\vec{y}\in V$满足$\vec{x}+\vec{y}=\vec{0}(\forall\vec{x}\in\vec{x})$，则$\vec{y}=\vec{0}$.

解：在运算方式（1）之下，考虑$\vec{\alpha}=(1,1)\in V$及$k,l\in R$.

$\because (k+l)\circ\vec{\alpha}=(k+l,1)$，$k\circ\vec{\alpha}+l\circ\vec{\alpha}=(k,1)+(l,1)=(k+l,2)$

$\therefore (k+l)\circ\vec{\alpha}\neq k\circ\vec{\alpha}+l\circ\vec{\alpha}$即元素对数乘运算的分配律不成立，故$V$不能构成$R$上的线性空间.

在运算方式（2）之下，显然$\vec{\alpha}\oplus\vec{\beta}\in V,k\odot\vec{\alpha}\in V$，及线性运算封闭．再设$\vec{\gamma}=(t_1,t_2)\in V,l\in R$，则有

①
$$\begin{aligned}&\vec{\alpha}\oplus(\vec{\beta}\oplus\vec{\gamma})=(\xi_1,\xi_2)\oplus(\eta_1+t_1,\eta_2+t_2+\eta_1t_1)\\&=(\xi_1+(\eta_1+t_1),\xi_2+(\eta_2+t_2+\eta_1t_1)+\xi_1(\eta_1+t_1))\\&=((\xi_1+\eta_1)+t_1,(\xi_2+\eta_2+\xi_1\eta_1)+t_2+(\xi_1+\eta_1)t_1)\\&=(\xi_1+\eta_1,\xi_2+\eta_2+\xi_1\eta_1)\oplus(t_1,t_2)=(\vec{\alpha}\oplus\vec{\beta})\oplus\vec{\gamma}\end{aligned};$$

② $\vec{\alpha}\oplus\vec{\beta}=(\xi_1+\eta_1,\xi_2+\eta_2+\xi_1\eta_1)=(\eta_1+\xi_1,\eta_2+\xi_2+\eta_1\xi_1)=\vec{\beta}\oplus\vec{\alpha}$；

③ 对于任意的$\vec{\alpha}\in V$，由$\vec{\alpha}\oplus\vec{\beta}=\vec{\alpha}$可得$(\xi_1+\eta_1,\xi_2+\eta_2+\xi_1\eta_1)=(\xi_1,\xi_2)$，即$\xi_1+\eta_1=\xi_1$，$\xi_2+\eta_2+\xi_1\eta_1=\xi_2$，解之得$\eta_1=0,\eta_2=0$，于是得$\vec{0}=(0,0)$，满足$\vec{\alpha}\oplus\vec{0}=\vec{\alpha}$；

④ 对于任意的$\vec{\alpha}\in V$，由$\vec{\alpha}\oplus\vec{\beta}=\vec{0}$可得$(\xi_1+\eta_1,\xi_2+\eta_2+\xi_1\eta_1)=(0,0)$，即$\xi_1+\eta_1=0$，$\xi_2+\eta_2+\xi_1\eta_1=0$，解之得$\eta_1=-\xi_1,\eta_2=\xi_1^2-\xi_2$，于是$-\vec{\alpha}=(-\xi_1,\xi_1^2-\xi_2)$满足$\vec{\alpha}\oplus(-\vec{\alpha})=\vec{0}$；

⑤
$$\begin{aligned}&\because k\odot(\vec{\alpha}+\vec{\beta})=k\odot(\xi_1+\eta_1,\xi_2+\eta_2+\xi_1\eta_1)\\&=(k(\xi_1+\eta_1),k(\xi_2+\eta_2+\xi_1\eta_1)+\frac{1}{2}k(k-1)(\xi_1+\eta_1)^2)\end{aligned},$$

$$k\odot\vec{\alpha}\oplus k\odot\vec{\beta}=(k\xi_1,k\xi_2+\frac{1}{2}k(k-1)\xi_1^2)\oplus(k\eta_1,k\eta_2+\frac{1}{2}k(k-1)\eta_1^2)$$

$$=(k\xi_1+k\eta_1,k\xi_2+\frac{1}{2}k(k-1)\xi_1^2+k\eta_2+\frac{1}{2}k(k-1)\eta_1^2+(k\xi_1)(k\eta_1))$$

$$=(k(\xi_1+\eta_1),k(\xi_2+\eta_2)+\frac{1}{2}k(k-1)(\xi_1^2+\eta_1^2)+k^2\xi_1\eta_1)$$

$$=(k(\xi_1+\eta_1),k(\xi_2+\eta_2+\xi_1\eta_1)+\frac{1}{2}k(k-1)(\xi_1+\eta_1)^2)$$

$\therefore k\odot(\vec{\alpha}+\vec{\beta})=k\odot\vec{\alpha}\oplus k\odot\vec{\beta}$；

⑥ $\because (k+l)\odot\vec{\alpha}=((k+l)\xi_1,(k+l)\xi_2+\dfrac{1}{2}(k+l)(k+l-1)\xi_1^2)$，

$$k\odot\vec{\alpha}\oplus l\odot\vec{\alpha}=(k\xi_1,k\xi_2+\frac{1}{2}k(k-1)\xi_1^2)\oplus(l\xi_1,l\xi_2+\frac{1}{2}l(l-1)\xi_1^2)$$

$$=(k\xi_1+l\xi_1,k\xi_2+\frac{1}{2}k(k-1)\xi_1^2+l\xi_2+\frac{1}{2}l(l-1)\xi_1^2+(k\xi_1)(l\xi_1))$$

$$=((k+l)\xi_1,(k+l)\xi_2+\frac{1}{2}(k+l)(k+l-1)\xi_1^2)$$

$\therefore (k+l)\odot\vec{\alpha}=k\odot\vec{\alpha}\oplus l\odot\vec{\alpha}$；

⑦ $\because (kl)\odot\vec{\alpha}=((kl)\xi_1,(kl)\xi_2+\dfrac{1}{2}(kl)(kl-1)\xi_1^2)$，

$$k\odot(l\odot\vec{\alpha})=k\odot(l\xi_1,l\xi_2+\frac{1}{2}l(l-1)\xi_1^2)=(kl\xi_1,kl\xi_2+\frac{1}{2}kl(l-1)\xi_1^2+\frac{1}{2}k(k-1)(l\xi_1)^2)$$

$$=((kl)\xi_1,(kl)\xi_2+\frac{1}{2}(kl)(kl-1)\xi_1^2)$$

$\therefore (kl)\odot\vec{\alpha}=k\odot(l\odot\vec{\alpha})$

⑧ $1\odot\vec{\alpha}=(1\times\xi_1,1\times\xi_2+\dfrac{1}{2}\times1\times(1-1)\times\xi_1^2)=(\xi_1,\xi_2)=\vec{\alpha}$

故在运算方式（2）下V构成R上的线性空间*.

三、线性空间的基本性质

性质 1 线性空间的零元素是唯一的.

性质 2 任一元素的负元素是唯一的.

性质 3 设$\lambda,0,-1,1\in P$，$\vec{x},-\vec{x},\vec{0}\in V$，则 1、$0\vec{x}=\vec{0}$；2、$(-1)\vec{x}=-\vec{x}$；3、$\lambda\vec{0}=\vec{0}$；

4、若$\lambda\vec{x}=\vec{0}$，则$\lambda=0$或$\vec{x}=0$.

证明：$\because \vec{x}+0\vec{x}=1\vec{x}+0\vec{x}=(1+0)\vec{x}=\vec{x}$，$\therefore 0\vec{\boldsymbol{x}}=\vec{0}$；

$\because \vec{x}+(-1)\vec{x}=1\vec{x}+(-1)\vec{x}=[1+(-1)]\vec{x}=0\vec{x}=\vec{0}$，$\therefore (-1)\vec{x}=-\vec{x}$；

假设$\lambda\neq0$且$\vec{x}\neq\vec{0}$，那么$\vec{x}=1\vec{x}=(\lambda\dfrac{1}{\lambda})\vec{x}=\dfrac{1}{\lambda}(\lambda\vec{x})=\vec{0}$，与假设矛盾，故$\lambda\neq0$与

$\vec{x}=0$不能同时成立.

---

* 对于一个具体的线性空间，如果指定的线性运算方式不是通常的，那么，相应的零元素和负元素可能与通常的形式不同；另外，线性空间的定义离不开数域，对不同的数域，同一个集合能构成线性空间，也可能构不成线性空间．如：复数集合 C 既是复数域 C 上的线性空间（记为 V_C），又是实数域 R 上的线性空间（记为 V_R），（1，i）在 V_C 中线性相关，在 V_R 中线性无关.

定义 1-4　只含一个元素的线性空间叫做**零空间**，显然，这个元素便是零元.

四、线性空间的基、维数与坐标

★ 根据线性空间的定义，有限个向量组成的集合，总不能满足加法及数乘运算的封闭性，所以除只由一个零向量构成的零空间 $\{\vec{0}\}$ 外，一般线性空间都有无穷多个向量.

如果 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_r}(r\geq 1)$ 为线性空间 V 中一组向量，$k_1,k_2,\cdots,k_r$ 是数域 P 中的数，那么向量 $\vec{x}=k_1\overrightarrow{x_1}+k_2\overrightarrow{x_2}+\cdots+k_r\overrightarrow{x_r}$ 称为向量 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_r}$ 的一个**线性组合**，有时也说向量 $\vec{x}$ 可用向量 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_r}$ **线性表示**. 如果 $k_1,k_2,\cdots,k_r$ 不全为零，且使 $k_1\overrightarrow{x_1}+k_2\overrightarrow{x_2}+\cdots+k_r\overrightarrow{x_r}=\vec{0}$，则称向量组 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_r}$ **线性相关**，否则就称其**线性无关**. 换句话说，只有在 $k_1=k_2=\cdots=k_r=0$ 时才成立，称 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_r}$ 线性无关.

显然，如果 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_r}$ 中有一为零元，则这 r 个元素必然是线性相关的*.

根据定义，R^n 中的两个向量组 $\begin{cases}\varepsilon_1=(1,0,\cdots,0)\\ \varepsilon_2=(0,1,\cdots,0)\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \varepsilon_n=(0,0,\cdots,1)\end{cases}$ 及 $\begin{cases}\varepsilon_1'=(1,1,\cdots,1,1)\\ \varepsilon_2'=(0,1,\cdots,1,1)\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \varepsilon_n'=(0,0,\cdots,0,1)\end{cases}$ 都是线性无关的.

例：讨论 $R^{2\times 2}$ 的矩阵组 $A_1=\begin{bmatrix}a&1\\1&1\end{bmatrix},A_2=\begin{bmatrix}1&a\\1&1\end{bmatrix},A_3=\begin{bmatrix}1&1\\a&1\end{bmatrix},A_4=\begin{bmatrix}1&1\\1&a\end{bmatrix}$ 的线性相关性.

解：设有一组数 $k_1,k_2,k_3,k_4\in R$，使得 $k_1A_1+k_2A_2+k_3A_3+k_4A_4=O$，即得

$\begin{cases}ak_1+k_2+k_3+k_4=0\\ k_1+ak_2+k_3+k_4=0\\ k_1+k_2+ak_3+k_4=0\\ k_1+k_2+k_3+ak_4=0\end{cases}$，该齐次线性方程组的系数行列式为

$$\begin{vmatrix}a&1&1&1\\1&a&1&1\\1&1&a&1\\1&1&1&a\end{vmatrix}=\begin{vmatrix}3+a&1&1&1\\3+a&a&1&1\\3+a&1&a&1\\3+a&1&1&a\end{vmatrix}=(a+3)\begin{vmatrix}1&1&1&1\\1&a&1&1\\1&1&a&1\\1&1&1&a\end{vmatrix}=(a+3)\begin{vmatrix}1&1&1&1\\0&a-1&0&0\\0&0&a-1&0\\0&0&0&a-1\end{vmatrix}$$
$$=(a+3)(a-1)^3$$

---

* 可假设 $\overrightarrow{x_i}=\vec{0}(1\leq i\leq n),\Rightarrow \exists k_i\neq 0$ 满足条件.

根据克拉默法则，当 $a\neq -3$ 且 $a\neq 1$ 时，齐次线性方程组只有零解，从而 A_1,A_2,A_3,A_4 线性无关；当 $a=-3$ 或 $a=1$ 时，齐次线性方程组有非零解，从而 A_1,A_2,A_3,A_4 线性相关.

例： 讨论 $P[t]_3$ 的多项式组 $\begin{aligned} f_1(t)&=a+t+t^2+t^3\\ f_2(t)&=1+at+t^2+t^3\\ f_3(t)&=1+t+2t^2+t^3\end{aligned}$ 的线性相关性.

解：取 $P[t]_3$ 的简单基 $1,t,t^2,t^3$，$f_i(t)(i=1,2,3)$ 在该基下的坐标依次为 $\overrightarrow{\beta_1}=\begin{pmatrix}a\\1\\1\\1\end{pmatrix}$，$\overrightarrow{\beta_2}=\begin{pmatrix}1\\a\\1\\1\end{pmatrix}$，

$\overrightarrow{\beta_3}=\begin{pmatrix}1\\1\\2\\1\end{pmatrix}$，$\because A=\begin{pmatrix}a&1&1\\1&a&1\\1&1&2\\1&1&1\end{pmatrix}\sim\begin{pmatrix}1&1&1\\1&a&1\\1&1&2\\a&1&1\end{pmatrix}\sim\begin{pmatrix}1&1&1\\0&a-1&0\\0&0&1\\a-1&0&0\end{pmatrix}\sim\begin{pmatrix}1&1&0\\0&0&1\\0&a-1&0\\a-1&0&0\end{pmatrix}$，

$\therefore$ 当 $a\neq 1$ 时 $A=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\0&0&0\end{pmatrix}$，$R(A)=3$，从而 $f_1(t),f_2(t),f_3(t)$ 线性无关；当 $a=1$ 时 $R(A)=2$，从而 $f_1(t),f_2(t),f_3(t)$ 线性相关.

例： 设 V 是数域 R 上所有实函数构成的线性空间，讨论 V 中元素组 t,e^t,e^{2t} 的线性相关性.

解：设一组数 $k_1,k_2,k_3\in R$ 使得 $k_1t+k_2e^t+k_3e^{2t}=0$，该式两端对 t 求一阶和二阶导数，

联立后得到 $\begin{cases}tk_1+e^tk_2+e^{2t}k_3=0\\k_1+e^tk_2+2e^{2t}k_3=0\\e^tk_2+4e^{2t}k_3=0\end{cases}$，该齐次线性方程组的系数行列式为

$\begin{vmatrix}t&e^t&e^{2t}\\1&e^t&2e^{2t}\\0&e^t&4e^{2t}\end{vmatrix}=\begin{vmatrix}t&e^t&e^{2t}\\1-t&0&e^{2t}\\-t&0&3e^{2t}\end{vmatrix}=e^{3t}(2t-3)$，当 $t=\dfrac{3}{2}$ 时，方程组有非零解，即 t,e^t,e^{2t} 线性无关，当 $t\neq\dfrac{3}{2}$ 时，方程组只有零解，即 t,e^t,e^{2t} 线性相关.

命题 1-1 $r\geq 2$ 时，V 中的向量组 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_r}$ 线性相关的充要条件是其中至少有一个向量可由向量组中其余向量线性表示；而线性无关的充要条件则是其中每一个向量都不能由向

量组中其余向量线性表示.

命题 1-2 若V中向量组的子向量线性相关，则该向量组也线性相关.

命题 1-3 若V中某向量组线性无关，则其任一子向量组也线性无关.

定义 1-5 设V是数域P上的线性空间，$\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$（$n\geq 1$）是属于$V$的任意$n$个向量，如果它满足：

（1）$\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$线性无关；

（2）V中任一向量$\vec{x}$均可由$\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$来线性表示，则称$\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$是$V$的一组**基**（或**基底**），并称$\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$为基向量.

线性空间V的基向量所含向量的个数n，称为线性空间V的**维数**，记为$\dim V=n$，并称V为n**维线性空间**，可简记为V^n. ★

维数实际上就是V中线性无关向量组中向量的最大个数；而基只不过是V中的最大线性无关组而已.

一个线性空间的基不是唯一的，线性空间里不同基所含向量的个数是相等的，即线性空间的维数是确定的.

例： 求实数域上全体n阶实对称矩阵构成的线性空间$S^{n\times n}$的基与维数.

解：选取$S^{n\times n}$中的一组矩阵：$F_{ii}=E_{ii}(i=1,2,\cdots,n)$，$F_{ij}=E_{ij}+E_{ji}(i<j,i,j=1,2,\cdots,n)$

设$A=(a_{ij})_{n\times n}\in S^{n\times n}$，则有$a_{ij}=a_{ji}$，从而$A=\sum\limits_{i\leq j}a_{ij}F_{ij}$. 有一组数

$k_{ij}(i\leq j,i,j=1,2,\cdots,n)$使得$\sum\limits_{i\leq j}k_{ij}F_{ij}=O$，即有$\begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1n} \\ k_{12} & k_{22} & \cdots & k_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ k_{1n} & k_{2n} & \cdots & k_{nn} \end{pmatrix}=O$，则必有

$k_{ij}=0(i\leq j,i,j=1,2,\cdots,n)$，从而矩阵组$F_{11},F_{12},\cdots,F_{1n},F_{22},F_{23},\cdots,F_{nn}$线性无关. 由定义可知，该矩阵组是$S^{n\times n}$的一个基，且

$$\dim S^{n\times n}=n+(n-1)+\cdots+2+1=\frac{n(n+1)}{2}.$$

例： 全体$m\times n$阶实矩阵的集合构成一个实线性空间（对于矩阵加法和数对矩阵的数乘运算），求其维数.

解：一个直接的方法就是找一个最大线性无关组，其元素尽可能简单.

令E_{ij}为这样的一个$m\times n$阶矩阵，其(i,j)位元素为1，其余元素为零.

★ 通常求维数的方法：若$\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$线性无关，$\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$即为一组基，维数 n；若若$\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$线性相关，则求它的一个最大无关组.

显然，这样的矩阵共有 mn 个，构成一个具有 mn 个元素的线性无关元素组 $\{E_{11},E_{12},\cdots E_{1n};E_{21},E_{22},\cdots E_{2n};\cdots;E_{m1},E_{m2},\cdots E_{mn}\}$. 另一方面，还需说明元素个数最大. 对于任意的 $A=(a_{ij})_{m\times n}$，都可由以上元素组线性表示，$A=\sum_{i,j}a_{ij}E_{ij}$

即 $\{E_{ij}\,|\,i=1,2,\cdots,m;\ j=1,2,\cdots,n\}$ 构成了最大线性无关元素组，所以该空间的维数为 mn .

定理 1-1 设 $\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$ 是 V^n 的一组基，对于任何向量 $\vec{x}\in V^n$，则它可唯一地用 $\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$ 线性表示.

定义 1-6 设 $\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$ 是线性空间 V^n 的一组基，对于任一向量 $\vec{x}\in V^n$，总有且仅有一组有序数 $x_1,x_2,\cdots,x_n$ 使 $\vec{x}=x_1\vec{x_1}+x_2\vec{x_2}+\cdots x_n\vec{x_n}$，$x_1,x_2,\cdots,x_n$ 这组有序数就称为向量 $\vec{x}$ 在基 $\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$ 下的**坐标**，并记作 $\overrightarrow{X}=(x_1,x_2,\cdots,x_n)$ 或 $\overrightarrow{X}=(x_1,x_2,\cdots,x_n)^T$.

同一向量 $\vec{x}$ 在不同的基（或称**坐标系**）下的坐标往往不同. 例如：在线性空间 $P[x]_n$ 中，多项式 $f(x)=a_0+a_1x+a_2x^2+\cdots a_nx^n$ 在基 $1,x,x^2,\cdots,x^n$ 下的坐标就是它的系数构成的行向量 $(a_0,a_1,\cdots,a_n)$. 在另一组基 $1,(x-a),(x-a)^2,\cdots,(x-a)^n$ 下的坐标为

$$(f(a),f^{'}(a),\frac{f^{''}(a)}{2!},\cdots,\frac{f^{(n)}(a)}{n!})$$

（$\because f(x)=f(a)+f^{'}(a)(x-a)+\dfrac{f^{''}(a)(x-a)^2}{2!}+\cdots+\dfrac{f^{(n)}(a)}{n!}(x-a)^n$）

五、基变换与坐标变换

设 $\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$ 及 $\vec{e_1'},\vec{e_2'},\cdots,\vec{e_n'}$ 是 V^n 中的两组基，且 $(\vec{e_1'},\vec{e_2'},\cdots,\vec{e_n'})=(\vec{e_1},\vec{e_2},\cdots,\vec{e_n})C$ *，其中矩阵 $C=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$，称为由旧基 $\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$ 变到新基 $\vec{e_1'},\vec{e_2'},\cdots,\vec{e_n'}$ 的**过渡矩阵**. 由于 $\vec{e_1'},\vec{e_2'},\cdots,\vec{e_n'}$ 线性无关，故过渡矩阵 C 可逆，即 C^{-1} 存在.

设 $\vec{x}\in V^n$，且 $\vec{x}$ 在两组基下的坐标分别为 $\vec{\alpha}=(x_1,x_2,\cdots,x_n)^T$ 及 $\vec{\beta}=(x_1',x_2',\cdots,x_n')^T$，

* 注意矩阵 C 的位置以及哪个是旧基，哪个是新基.

$\vec{x}=x_1\vec{e_1}+x_2\vec{e_2}+\cdots+x_n\vec{e_n}=x_1'\vec{e_1'}+x_2'\vec{e_2'}+\cdots+x_n'\vec{e_n'}$，写成矩阵形式即为

$\vec{x}=(\vec{e_1},\vec{e_2},\cdots,\vec{e_n})\begin{bmatrix}x_1\\x_2\\\cdots\\x_n\end{bmatrix}=(\vec{e_1'},\vec{e_2'},\cdots,\vec{e_n'})\begin{bmatrix}x_1'\\x_2'\\\cdots\\x_n'\end{bmatrix}=(\vec{e_1},\vec{e_2},\cdots,\vec{e_n})C\begin{bmatrix}x_1'\\x_2'\\\cdots\\x_n'\end{bmatrix}$，由于 $\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$ 线性无关

且为基，故有 $\begin{bmatrix}x_1\\x_2\\\cdots\\x_n\end{bmatrix}=C\begin{bmatrix}x_1'\\x_2'\\\cdots\\x_n'\end{bmatrix}$，由 C 可逆，故 $\begin{bmatrix}x_1'\\x_2'\\\cdots\\x_n'\end{bmatrix}=C^{-1}\begin{bmatrix}x_1\\x_2\\\cdots\\x_n\end{bmatrix}$，即 $\vec{\alpha}=C\vec{\beta}$*，或 $\vec{\beta}=C^{-1}\vec{\alpha}$.

试证明存在 $\vec{0}\neq\vec{x}\in V^n$，使得 $\vec{x}$ 在两组基下有相同坐标的充要条件是 1 为 C 的一个特征值.

证： 必要性．若 $\vec{\alpha}=\vec{\beta}$，则由坐标变换公式 $\vec{\alpha}=C\vec{\beta}$ 可得 $C\vec{\beta}=1\vec{\beta}$，因为 $\vec{\beta}\neq 0$，所以 1 为 C 的一个特征值.

充分性．若 1 为 C 的一个特征值，取 C 的对应于特征 1 的特征向量作为 $\vec{\beta}$，则有 $C\vec{\beta}=1\vec{\beta}$．那么 $\vec{x}=(\vec{e_1'},\vec{e_2'},\cdots,\vec{e_n'})\vec{\beta}\neq 0$，根据坐标变换公式，$\vec{\alpha}=C\vec{\beta}=\vec{\beta}$.

例：设 $P[t]_3$ 的两个基为：

（Ⅰ）$f_1(t)=1,f_2(t)=1+t,f_3(t)=1+t+t^2,f_4(t)=1+t+t^2+t^3$

（Ⅱ）$g_1(t)=1+t^2+t^3,g_2(t)=t+t^2+t^3,g_3(t)=1+t+t^2,g_4(t)=1+t+t^3$

（1）求由基（Ⅰ）改变为基（Ⅱ）的过渡矩阵；

（2）求 $P[t]_3$ 中在基（Ⅰ）和基（Ⅱ）下有相同坐标的全体多项式.

解：（1）取 $P[t]_3$ 的简单基 $1,t,t^2,t^3$，则由简单基到基（Ⅰ）和基（Ⅱ）的过渡矩阵分别为

$C_1=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{pmatrix},C_2=\begin{pmatrix}1&0&1&1\\0&1&1&1\\1&1&1&0\\1&1&0&1\end{pmatrix}$，则有 $(f_1,f_2,f_3,f_4)=(1,t,t^2,t^3)C_1$，且

$(g_1,g_2,g_3,g_4)=(1,t,t^2,t^3)C_2=(f_1,f_2,f_3,f_4)C_1^{-1}C_2$，那么由基（Ⅰ）改变为基（Ⅱ）的过渡矩阵则为 $C=C_1^{-1}C_2$，显然 $|C_1|\neq 0$，所以 C_1^{-1} 存在，下面求 C_1^{-1}.

* $\vec{\alpha},\vec{\beta}$ 均为列向量，注意三者的位置.

$$\left(C_1|E\right)=\begin{pmatrix}1&1&1&1&1&0&0&0\\0&1&1&1&0&1&0&0\\0&0&1&1&0&0&1&0\\0&0&0&1&0&0&0&1\end{pmatrix}\sim\begin{pmatrix}1&0&0&0&1&-1&0&0\\0&1&0&0&0&1&-1&0\\0&0&1&0&0&0&1&-1\\0&0&0&1&0&0&0&1\end{pmatrix}$$，所以有

$$C_1^{-1}=\begin{pmatrix}1&-1&0&0\\0&1&-1&0\\0&0&1&-1\\0&0&0&1\end{pmatrix}$$，那么 $C=C_1^{-1}C_2=\begin{pmatrix}1&-1&0&0\\0&1&-1&0\\0&0&1&-1\\0&0&0&1\end{pmatrix}\begin{pmatrix}1&0&1&1\\0&1&1&1\\1&1&1&0\\1&1&0&1\end{pmatrix}$

$$=\begin{pmatrix}1&-1&0&0\\-1&0&0&1\\0&0&1&-1\\1&1&0&1\end{pmatrix}.$$

（2）设 $f(t)\in P[t]_3$ 在基（Ⅰ）和基（Ⅱ）下的坐标分别为 $\vec{\alpha}=(\xi_1,\xi_2,\xi_3,\xi_4)^T$，$\vec{\beta}=(\eta_1,\eta_2,\eta_3,\eta_4)^T$，根据坐标变换公式可得 $\vec{\alpha}=C\vec{\beta}$，现要求 $\vec{\alpha}=\vec{\beta}$，即有 $(E-C)\vec{\beta}=\vec{0}$，$E-C=\begin{pmatrix}0&1&0&0\\1&1&0&-1\\0&0&0&1\\-1&-1&0&0\end{pmatrix}\sim\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&0&0\end{pmatrix}$，齐次线性方程组的通解为 $\vec{\beta}=k(0,0,1,0)^T(k\in R)$，于是在基（Ⅰ）和基（Ⅱ）下有相同坐标的全体多项式为 $f(t)=(g_1(t),g_2(t),g_3(t),g_4(t))\beta=kg_3(t)=k+kt+kt^2(k\in R)$.

1.2 线性子空间

一、子空间的概念

定义 1-7 设 V_1 是数域 P 上线性空间 V 的一个子集合，且这个子集合对 V 已有的加法及数乘运算也构成线性空间，则称 V_1 为 V 的**线性子空间**，简称**子空间**，记为 $V_1\subseteq V$，当 $V_1\neq V$ 时，记为 $V_1\subset V$.

定理 1-2 设 V_1 是线性空间 V 的一个非空子集合，则 V_1 是 V 的一个字空间的充要条件为：

（1）如果 $\vec{x},\vec{y}\in V_1$，则 $\vec{x}+\vec{y}\in V_1$；（2）如果 $\vec{x}\in V_1$，$k\in P$，则 $k\vec{x_1}\in V_1$.

容易看出，每个线性空间至少有两个子空间，一个是它自身，另一个是仅由零向量所构成的

子集合，称后者为零子空间．这两个子空间通常称为平凡子空间，而其他的子空间称为非平凡子空间（或真子空间）．

任何一个线性子空间的维数不能超过整个空间的维数，即有 $\dim V_1 \le \dim V$ ．

设 $\overrightarrow{x_1}, \overrightarrow{x_2}, \cdots, \overrightarrow{x_m}$ 是线性空间 V 中一组向量，由这组向量所有可能的线性组合的集合 $V_1 = \{k_1\overrightarrow{x_1} + k_2\overrightarrow{x_2} + \cdots k_m\overrightarrow{x_m}\}$ 是非空的，验证 V_1 是 V 的一个字空间，这个子空间叫做 $\overrightarrow{x_1}, \overrightarrow{x_2}, \cdots, \overrightarrow{x_m}$ 生成的子空间，记为 $Span(\overrightarrow{x_1}, \overrightarrow{x_2}, \cdots, \overrightarrow{x_m}) = \{k_1\overrightarrow{x_1} + k_2\overrightarrow{x_2} + \cdots k_m\overrightarrow{x_m}\}$ ．

常用的子空间有：

$R^n = Span(\overrightarrow{\varepsilon_1}, \overrightarrow{\varepsilon_2}, \cdots, \overrightarrow{\varepsilon_n}), \overrightarrow{\varepsilon_i} = (0, 0, \cdots, 0, 1, 0 \cdots, 0)^T$ （第 i 位是 1）， $\dim R^n = n$ ；

$R^{m\times n} = Span(E_{11}, E_{12}, \cdots, E_{1n}, E_{21}, E_{22}, \cdots, E_{mn})$ ， $\dim R^{m\times n} = mn$ ；

$P[x]_n = Span(1, x, x^2, \cdots, x^n)$ ， $\dim P[x]_n = n+1$ ．

例： 给定矩阵 $P \in R^{n\times n}$ ，判断 $R^{n\times n}$ 的子集 $V_1 = \{A \mid AP = PA, A \in R^{n\times n}\}$ 是否构成子空间．

解：因为 $O_{n\times n} \in V_1$ ，所以 V_1 非空．设 $A, B \in V_1$ ，则有 $AP = PA, BP = PB$ ．

$\because (A+B)P = AP + BP = PA + PB = P(A+B);$

$(kA)P = k(AP) = k(PA) = P(kA), (k \in R)$

$\therefore A+B \in V_1, kA \in V_1$ ，故 V_1 是 $R^{n\times n}$ 上的子空间．

★ 两个行列式为零的同阶方阵之和的行列式不一定为零．两个同阶幂等矩阵之和不一定是幂等矩阵．即 $V_1 = \{A \mid \det A = 0, A \in R^{2\times 2}\}, V_2 = \{A \mid A^2 = A, A \in R^{2\times 2}\}$ 不能构成子空间．

例： 已知 $P = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$ 及 $R^{2\times 2}$ 的子空间 $V_1 = \{A \mid AP = PA, A \in R^{n\times n}\}$

（1）求 V_1 的基与维数；（2）写出 V_1 中矩阵的一般形式．

解：（1）设 $A = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \in V_1$ ，则由 $AP = PA$ 得到 $\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$ ，

可得齐次线性方程组 $\begin{cases} 3x_1 + x_2 - 3x_4 = 0 \\ x_3 = 0 \end{cases}$ ，显然，系数矩阵的秩为 2，那么解空间的维数为

4-2=2，即有 $\dim V_1 = 2$，它的基础解系为 $(1,-3,0,0)^T,(1,0,0,1)^T$，从而得 V_1 的基为

$A_1 = \begin{pmatrix} 1 & -3 \\ 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

（2）V_1 中矩阵的一般形式为 $A = k_1A_1 + k_2A_2 = \begin{pmatrix} k_1+k_2 & -3k_1 \\ 0 & k_2 \end{pmatrix}(k_1,k_2 \in R)$.

二、子空间的交和

定义 1-8 设 V_1 和 V_2 是 n 维线性空间 V 的两个子空间，由同时属于这两个子空间中的向量构成的子集合，叫做 V_1 与 V_2 的**交**，记作 $V_1 \cap V_2$.

定理 1-3 设 V_1、V_2 是数域 P 上的线性空间 V 的两个子空间，则它们的交 $V_1 \cap V_2$ 也是 V 的子空间.

定义 1-9 设 V_1、V_2 是数域 P 上的线性空间 V 的两个子空间，且 $\vec{x} \in V_1, \vec{y} \in V_2$，由所有 $\vec{x}+\vec{y}$ 这样的向量构成的集合叫做 V_1 与 V_2 的**和**，或**和空间**，记作 $V_1 + V_2$.

定理 1-4 如果 V_1、V_2 是数域 P 上的线性空间 V 的两个子空间，则它们的和 V_1+V_2 也是 V 的子空间.

定理 1-5 设 V_1 是数域 P 上 n 维线性空间 V 的一个 m 维子空间，$\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\cdots,\overrightarrow{\alpha_m}$ 是 V_1 的一组基，那么向量必定可扩充为整个空间的基，也就是说，在 V 中必定能够找到 $n-m$ 个向量 $\overrightarrow{\alpha_{m+1}},\overrightarrow{\alpha_{m+2}},\cdots,\overrightarrow{\alpha_n}$ 使得 $\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\cdots,\overrightarrow{\alpha_n}$ 是 V 的一组基.

定理 1-6 （维数公式）设 V_1、V_2 是数域 P 上的线性空间 V 的两个子空间，则

$$\dim V_1 + \dim V_2 = \dim(V_1+V_2) + \dim(V_1 \cap V_2)$$

从公式可以看出，两个子空间和的维数往往要比维数的和来得小.

定义 1-10 如果 V_1+V_2 中的任一向量只能唯一地表示为子空间 V_1 的一个向量与子空间 V_2 的一个向量的和，则称 V_1+V_2 为**直和**（或**直接和**），记为 $V_1 \oplus V_2$.

定理 1-6 和 V_1+V_2 是直和的充要条件为：V_1 与 V_2 之交 $V_1 \cap V_2$ 为零子空间，即 $V_1 \cap V_2 = \{\vec{0}\}$

推论 1 和 V_1+V_2 是直和的充要条件为 $\dim(V_1+V_2) = \dim V_1 + \dim V_2$.

推论 2 设 V_1+V_2 是直和，若 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_{n_1}}$ 是 V_1 的基，$\overrightarrow{y_1},\overrightarrow{y_2},\cdots,\overrightarrow{y_{n_2}}$ 是 V_2 的基，则

$\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_{n_1}};\overrightarrow{y_1},\overrightarrow{y_2},\cdots,\overrightarrow{y_{n_2}}$ 为 $V_1 \oplus V_2$ 的基.

定理 1-8 设 V_1 是 n 维线性空间 V 的一个子空间，则一定存在 V 的一个子空间 V_2，使 $V = V_1 \oplus V_2$.

例: 设 $R^{2\times2}$ 的两个子空间为 $V_1 = \{A \mid A = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}, x_1 - x_2 + x_3 - x_4 = 0\}$；

$V_2 = Span(B_1, B_2)$，$B_1 = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}, B_2 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$.

（1）将 $V_1 \oplus V_2$ 表示为生成子空间；（2）求 $V_1 \oplus V_2$ 基和维数；（3）求 $V_1 \cap V_2$ 基和维数.

解:（1）先将 V_1 表示成生成子空间. $\because$ 齐次线性方程组 $x_1 - x_2 + x_3 - x_4 = 0$ 的基础解系为

$\overrightarrow{\alpha_1} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \overrightarrow{\alpha_2} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \overrightarrow{\alpha_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$，$\therefore V_1$ 的一个基为 $A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$，于是 $V_1 = Span(A_1, A_2, A_3)$，$\therefore V_1 \oplus V_2 = Span(A_1, A_2, A_3, B_1, B_2)$.

（2）矩阵组 A_1, A_2, A_3, B_1, B_2 在 $R^{2\times2}$ 的简单基 $E_{11}, E_{12}, E_{21}, E_{22}$ 下的坐标依次为：

$\overrightarrow{\alpha_1}, \overrightarrow{\alpha_2}, \overrightarrow{\alpha_3}, \overrightarrow{\beta_1} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 3 \end{pmatrix}, \overrightarrow{\beta_2} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$，$(\overrightarrow{\alpha_1}, \overrightarrow{\alpha_2}, \overrightarrow{\alpha_3}, \overrightarrow{\beta_1}, \overrightarrow{\beta_2}) = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 3 & 1 \end{pmatrix}$

$\sim \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & -2 \\ 0 & 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$，显然 $R(\overrightarrow{\alpha_1}, \overrightarrow{\alpha_2}, \overrightarrow{\alpha_3}, \overrightarrow{\beta_1}, \overrightarrow{\beta_2}) = 4$，所以该向量组的最大无关组含有 4 个向量，而四个非零行的非零首元在 1，2，3，5 列，所以向量组的一个最大无关组为 $\overrightarrow{\alpha_1}, \overrightarrow{\alpha_2}, \overrightarrow{\alpha_3}, \overrightarrow{\beta_2}$，从而矩阵组 A_1, A_2, A_3, B_1, B_2 的一个最大无关组为 A_1, A_2, A_3, B_2，它们构成 $V_1 \oplus V_2$ 的一个基，且 $\dim(V_1 \oplus V_2) = 4$.

（3）设 $A \in V_1 \cap V_2$，则有数组 k_1, k_2, k_3 与数组 l_1, l_2 使得

$A = k_1A_1 + k_2A_2 + k_3A_3 = l_1B_1 + l_2B_2$，即 $k_1A_1 + k_2A_2 + k_3A_3 - l_1B_1 - l_2B_2 = O$，于是可得

齐次线性方程组$\begin{cases} k_1-l_1-l_2=0 \\ k_1+k_2+l_2=0 \\ k_2+k_3-2l_1=0 \\ k_3-3l_1-l_2=0 \end{cases}$，系数矩阵$\begin{pmatrix} 1 & 0 & 0 & -1 & -1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & -2 & 0 \\ 0 & 0 & 1 & -3 & -1 \end{pmatrix}$

$\sim\begin{pmatrix} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$，齐次线性方程组的通解为$(k_1,k_2,k_3,l_1,l_2)=k(1,-1,3,1,0)^T$

$(k\in R)$，于是可得$A=l_1B_1+l_2B_2=k(1\times B_1+0\times B_2)=k\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$，故$V_1\cap V_2$的一个基为

$\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$，且$\dim(V_1\cap V_2)=1$.

例： 假定$\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3}$是$R^3$的一组基，试求由$\overrightarrow{x_1'}=\overrightarrow{x_1}-2\overrightarrow{x_2}+3\overrightarrow{x_3},\overrightarrow{x_2'}=2\overrightarrow{x_1}+3\overrightarrow{x_2}+2\overrightarrow{x_3}$,

$\overrightarrow{x_3'}=4\overrightarrow{x_1}+13\overrightarrow{x_2}$生成的子空间$Span(\overrightarrow{x_1'},\overrightarrow{x_2'},\overrightarrow{x_3'})$的基.

解：$Span(\overrightarrow{x_1'},\overrightarrow{x_2'},\overrightarrow{x_3'})$的基为向量组$\overrightarrow{x_1'},\overrightarrow{x_2'},\overrightarrow{x_3'}$的一个最大无关组，在基$\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3}$下的坐标为

$\overrightarrow{\alpha_1}=\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix},\overrightarrow{\alpha_2}=\begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix},\overrightarrow{\alpha_3}=\begin{pmatrix} 4 \\ 13 \\ 0 \end{pmatrix},\begin{pmatrix} 1 & 2 & 4 \\ -2 & 3 & 13 \\ 3 & 2 & 0 \end{pmatrix}\sim\begin{pmatrix} 1 & 2 & 4 \\ 0 & 7 & 21 \\ 0 & -4 & -12 \end{pmatrix}\sim\begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$，可见$\overrightarrow{\alpha_1},\overrightarrow{\alpha_2}$

为一个最大无关组，$\therefore Span(\overrightarrow{x_1'},\overrightarrow{x_2'},\overrightarrow{x_3'})$的一个基为$\overrightarrow{x_1'},\overrightarrow{x_2'}$.

例： 试证明所有二阶矩阵之集合形成的实线性空间$R^{2\times2}$是所有二阶实对称矩阵之集合形成的子空间与所有二阶反对称矩阵之集合形成的子空间的直和.

证：设$R^{2\times2}$，$V_1=\{A=(a_{ij})_{2\times2}\,|\,a_{ij}=a_{ji},a_{ij}\in R\}$，$V_2=\{B=(b_{ij})_{2\times2}\,|\,b_{ij}=-b_{ji},b_{ij}\in R\}$，

易验证V_1,V_2均为V的子空间，$\forall C\in V$，有$C=\frac{1}{2}(C+C^T)+\frac{1}{2}(C-C^T)$，且

$\frac{1}{2}(C+C^T)\in V_1$，$\frac{1}{2}(C-C^T)\in V_2$ $\therefore V=V_1+V_2$，设$D=(d_{ij})_{2\times2}\in V_1\cap V_2$，那么$D\in V_1$

且$D\in V_2$ $\Rightarrow d_{ij}=d_{ji}$且$d_{ij}=-d_{ji}\Rightarrow d_{ij}=0$，$\therefore D=O$，所以$V_1\cap V_2=\{O\}$，即

$V=V_1\oplus V_2$.

§2　线性算子及其矩阵

2.1　线性空间上的线性算子

定义 2-1　设M与M'为两个集合，对应于每个$\vec{x}\in M$，如果根据某种法则T，在M'中有确定的$\vec{x'}$与之对应，那么称T为由M到M'的一个**映射**，或称**算子**. 记为$T:M\to M'$，或$T(\vec{x})=\vec{x'}$. 此时，$\vec{x'}$叫做$\vec{x}$在T下的像，$\vec{x}$叫做$\vec{x'}$的像源，M是T的定义域，$\vec{x'}$的全体构成T的值域，记为$T(M)$. 由定义可知$T(M)\subseteq M'$.

定义 2-2　设V与V'为数域P上的两个线性空间，T是由V到V'的一个算子，且对于V的任何两个向量$\vec{x_1},\vec{x_2}\in V$和任何数$\lambda\in P$，有$T(\vec{x_1}+\vec{x_2})=T(\vec{x_1})+T(\vec{x_2})$，则称$T$是由$V$到$V'$的**线性算子**（或**线性映射**）.

★ T是线性算子的充要条件是：对任何$\vec{x_1},\vec{x_2}\in V$及$\lambda_1,\lambda_2\in P$，有

$$\underline{T(\lambda_1\vec{x_1}+\lambda_2\vec{x_2})=\lambda_1 T(\vec{x_1})+\lambda_2 T(\vec{x_2})}$$

将线性空间V中每一向量映射成线性空间V'中的零向量的算子θ叫做**零算子**，它是一个线性算子.

在线性空间$P[x]_n$中，求微分是一个线性算子，这个例子通常用D表示，即$D(f(x))=f'(x)(\forall f(x)\in P[x]_n,x\in R)$.

在$[a,b]$上一切实连续函数的线性空间$C[a,b]$中，定义积分算子T，即$T(f(t))=\int_a^x f(t)dt$，$\forall f(t)\in C[a,b]$，则T是线性空间$C[a,b]$上的线性算子.

线性算子的性质：

1° 线性算子T把V中的零向量变为V'中的零向量；把向量$\vec{x}$的负向量$-\vec{x}$变为$\vec{x}$的像$T(\vec{x})$的负向量$-T(\vec{x})$.

2° 线性算子T把线性相关的向量组仍变为线性相关的向量组，即若$\vec{x_1},\vec{x_2},\cdots,\vec{x_r}$线性相关，

则它们的像 $T(\overrightarrow{x_1}),T(\overrightarrow{x_2}),\cdots,T(\overrightarrow{x_r})$ 也线性相关.

★ 线性算子可能把线性无关的向量组变为线性相关的向量组，如零算子 θ .

2.2 同构算子与线性空间同构

定义 2-3 设 T 是由 V 到 $V^{'}$ 的线性算子，且是“一对一”的：① $T(V)=V^{'}$（全映射）② 若 $\overrightarrow{x_1},\overrightarrow{x_2}\in V$，当 $\overrightarrow{x_1}\neq\overrightarrow{x_2}$ 时，$T(\overrightarrow{x_1})\neq T(\overrightarrow{x_2})$；换言之，由 $T(\overrightarrow{x_1})=T(\overrightarrow{x_2})$，即有（可逆映射），那么称 T 为 V 与 $V^{'}$ 间的一个**同构算子**，称 V 与 $V^{'}$ 是同构的线性空间.

★ 由不高于 3 次的实系数多项式组成的线性空间 $P[x]_3$ 与由实数域上四维向量全体组成的线性空间 R^4 同构.

同构线性空间的基本性质：

1° 传递性：设 V_1,V_2,V_3 是数域 P 上的线性空间，如果 V_1 与 V_2 同构，V_2 与 V_3 同构，则 V_1 与 V_3 也同构.

2° 同构线性空间中的零向量必定是互相对应的.

3° 同构空间中的线性相关向量组对应于线性相关向量组，线性无关向量组对应于线性无关向量组.

定理 2-1 数域 P 上两个有限维线性空间同构的充要条件是两空间的维数相等.

推论 数域 P 上任何 n 维线性空间 V^n 都与特殊的线性空间 $K^n=\{(a_1,a_2,\cdots,a_n)\,|\,a_i\in P\}$ 同构.

2.3 线性算子的矩阵表示

定义 2-4 设 $\overrightarrow{e_1},\overrightarrow{e_2},\cdots,\overrightarrow{e_n}$ 是 n 维线性空间 V^n 的一组基，T 是由 V^n 到 m 维线性空间 V^m 的线性算子，则 $T(\overrightarrow{e_1}),T(\overrightarrow{e_2}),\cdots,T(\overrightarrow{e_n})\in V^m$ 叫做 V^n 在算子 T 下的**基像**.

定义 2-5 设 T 与 B 是由 V^n 到 V^m 的两个线性算子，如果对于任何 $\vec{x}\in V^n$ 恒有 $B(\vec{x})=T(\vec{x})\in V^m$，则说 T 与 B **相等**.

定理 2-2 由 V^n 到 V^m 的线性算子 T 由基像 $T(\vec{e}_1),T(\vec{e}_2),\cdots,T(\vec{e}_n)$ 唯一地确定.

设 T 是由 n 维线性空间 V^n 到 m 维线性空间 V^m 的一个线性算子，取 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$ 作为 V^n 的基，$\vec{e}'_1,\vec{e}'_2,\cdots,\vec{e}'_m$ 作为 V^m 的基（有时称它们为**基偶**）. 由于线性算子 T 由基像 $T(\vec{e}_1),T(\vec{e}_2),\cdots,T(\vec{e}_n)$ 唯一地确定，且它们属于 V^m，故可令

$$\begin{cases} T(\vec{e}_1)=a_{11}\vec{e}'_1+a_{21}\vec{e}'_2+\cdots+a_{m1}\vec{e}'_m \\ T(\vec{e}_2)=a_{12}\vec{e}'_1+a_{22}\vec{e}'_2+\cdots+a_{m2}\vec{e}'_m \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T(\vec{e}_n)=a_{1n}\vec{e}'_1+a_{2n}\vec{e}'_2+\cdots+a_{mn}\vec{e}'_m \end{cases}，\text{即 } T(\vec{e}_i)=\sum_{j=1}^{m}a_{ji}\vec{e}'_j\,(i=1,2,\cdots,n)$$

或写成 $T(\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n)=(T(\vec{e}_1),T(\vec{e}_2),\cdots,T(\vec{e}_n))=(\sum_{j=1}^{m}a_{j1}\vec{e}'_j,\sum_{j=1}^{m}a_{j2}\vec{e}'_j,\cdots,\sum_{j=1}^{m}a_{jn}\vec{e}'_j)$

$$=(\vec{e}'_1,\vec{e}'_2,\cdots,\vec{e}'_m)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.\ \text{令 } A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

★ 基像 $T(\vec{e}_i)$ 的坐标恰是矩阵 A 的第 i 列 $(i=1,2,\cdots,n)$，因而 A 的行数等于 $\dim(V^m)=m$，而 A 的列数等于 $\dim(V^n)=n$，简记为 $A_{m\times n}$，$A_{m\times n}$ 也唯一确定.

定义 2-6 上式中矩阵 A 称为线性算子 T 在基偶 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$ 与 $\vec{e}'_1,\vec{e}'_2,\cdots,\vec{e}'_m$ 下的矩阵表示.

定理 2-3 若 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$ 是 n 维线性空间 V^n 的一组基，而 $\vec{y}_1,\vec{y}_2,\cdots,\vec{y}_n$ 是 m 维线性空间 V^m 中任意 n 个向量，则存在一个且只有一个线性算子 T，它把 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$ 分别映射为 $\vec{y}_1,\vec{y}_2,\cdots,\vec{y}_n$，即 $\vec{y}_i=T(\vec{e}_i)(i=1,2,\cdots,n)$.

定理 2-3 设 V^n 到 V^m 的线性算子 T 在基偶 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$ 与 $\vec{e}'_1,\vec{e}'_2,\cdots,\vec{e}'_m$ 下的矩阵为

$A=(a_{ij})_{m\times n}$，向量 $\vec{x}\in V^n$ 在基 $\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$ 下的坐标为 $(x_1,x_2,\cdots,x_n)$，则 $T(\vec{x})$ 在基 $\vec{e_1'},\vec{e_2'},\cdots,\vec{e_m'}$ 下的坐标 $(y_1,y_2,\cdots,y_n)$ 可按公式 $\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}=A\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$ * 来计算.

例： 给定 $R^{2\times 2}$ 的基 $\overrightarrow{X_1}=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\overrightarrow{X_2}=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},\overrightarrow{X_3}=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\overrightarrow{X_4}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$，及线性变换 $T(\overrightarrow{X})=\begin{pmatrix} a & 2b \\ 2c & a \end{pmatrix}\overrightarrow{X}(\overrightarrow{X}\in R^{2\times 2})$，其中 $a,b,c\in R$，求 T 在给定基下的矩阵 A .

解：取 $R^{2\times 2}$ 的简单基 $E_{11},E_{12},E_{21},E_{22}$

由简单基到给定基的过渡矩阵为 $C=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$. 计算

$$T(E_{11})=\begin{pmatrix} a & 2b \\ 2c & a \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} a & 0 \\ 2c & 0 \end{pmatrix},\quad T(E_{12})=\begin{pmatrix} a & 2b \\ 2c & a \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} 0 & a \\ 0 & 2c \end{pmatrix},$$

$$T(E_{21})=\begin{pmatrix} a & 2b \\ 2c & a \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}=\begin{pmatrix} 2b & 0 \\ a & 0 \end{pmatrix},\quad T(E_{22})=\begin{pmatrix} a & 2b \\ 2c & a \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 0 & 2b \\ 0 & a \end{pmatrix}.$$

直接写出 T 在简单基下的矩阵为 $A_0=\begin{pmatrix} a & 0 & 2b & 0 \\ 0 & a & 0 & 2b \\ 2c & 0 & a & 0 \\ 0 & 2c & 0 & a \end{pmatrix}$，于是 T 在给定基下的矩阵为

$A=C^{-1}A_0C$.

而 $$(C|E)=\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}\sim\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix},$$

* 注意矩阵 A 的位置.

$$\therefore C^{-1}=\begin{pmatrix}1&0&0&0\\0&0&0&1\\0&\frac{1}{2}&\frac{1}{2}&0\\0&\frac{1}{2}&-\frac{1}{2}&0\end{pmatrix},$$

于是 $A=\begin{pmatrix}1&0&0&0\\0&0&0&1\\0&\frac{1}{2}&\frac{1}{2}&0\\0&\frac{1}{2}&-\frac{1}{2}&0\end{pmatrix}\begin{pmatrix}a&0&2b&0\\0&a&0&2b\\2c&0&a&0\\0&2c&0&a\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&0&1&1\\0&0&1&-1\\0&1&0&0\end{pmatrix}$

$$=\begin{pmatrix}a&0&2b&0\\0&2c&0&a\\c&\frac{a}{2}&\frac{a}{2}&b\\-c&\frac{a}{2}&-\frac{a}{2}&b\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&0&1&1\\0&0&1&-1\\0&1&0&0\end{pmatrix}=\begin{pmatrix}a&0&2b&-2b\\0&a&2c&2c\\c&b&a&0\\-c&b&0&a\end{pmatrix}.$$

例： 设 $P[t]_2$ 的两个基为：

（Ⅰ） $f_1(t)=1+2t^2,f_2(t)=t+2t^2,f_3(t)=1+2t+5t^2$；

（Ⅱ） $g_1(t)=1-t,g_2(t)=1+t^2,g_3(t)=t+2t^2$.

线性变换 T 满足 $T(f_1(t))=2+t^2,T(f_2(t))=t,T(f_3(t))=1+t+t^2$

（1）求 T 在基（Ⅱ）下的矩阵；（2）设 $f(t)=1+2t+3t^2$，求 $T(f(t))$.

解：（1）取 $P[t]_2$ 的简单基 $1,t,t^2$，设 $T(f_1,f_2,f_3)=(1,t,t^2)A_0$，$(f_1,f_2,f_3)=(1,t,t^2)C_1$，$(g_1,g_2,g_3)=(1,t,t^2)C_2$，那么 $(g_1,g_2,g_3)=(f_1,f_2,f_3)C_1^{-1}C_2$，于是 $T(g_1,g_2,g_3)=T(f_1,f_2,f_3)C_1^{-1}C_2=(1,t,t^2)A_0C_1^{-1}C_2=(g_1,g_2,g_3)C_2^{-1}A_0C_1^{-1}C_2$，$T$ 在基（Ⅱ）下的矩阵即为 $A=C_2^{-1}A_0C_1^{-1}C_2$.

$\because A_0=\begin{pmatrix}2&0&1\\0&1&1\\1&0&1\end{pmatrix},C_1=\begin{pmatrix}1&0&1\\0&1&2\\2&2&5\end{pmatrix},C_2=\begin{pmatrix}1&1&0\\-1&0&1\\0&1&2\end{pmatrix}$，计算得 $C_1^{-1}=\begin{pmatrix}-1&-2&1\\-4&-3&2\\2&2&-1\end{pmatrix}$，

$$C_2^{-1}=\begin{pmatrix}-1&-2&1\\2&2&-1\\-1&-1&1\end{pmatrix},\ \therefore A=C_2^{-1}A_0C_1^{-1}C_2=\begin{pmatrix}1&2&-2\\1&-1&2\\0&1&-1\end{pmatrix}.$$

（2）$\because f(t)=(1,t,t^2)\begin{pmatrix}1\\2\\3\end{pmatrix}=(g_1,g_2,g_3)C_2^{-1}\begin{pmatrix}1\\2\\3\end{pmatrix}=(g_1,g_2,g_3)\begin{pmatrix}-2\\3\\0\end{pmatrix}$

$$\therefore T(f(t))=T(g_1,g_2,g_3)\begin{pmatrix}-2\\3\\0\end{pmatrix}=(g_1,g_2,g_3)A\begin{pmatrix}-2\\3\\0\end{pmatrix}=(g_1,g_2,g_3)\begin{pmatrix}4\\-5\\3\end{pmatrix}=-1-t+t^2.$$

例： 设线性空间V^3的两个基为（Ⅰ）：$\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3}$；（Ⅱ）：$\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3}$，由基（Ⅰ）改变为基（Ⅱ）的过渡矩阵为$C=\begin{pmatrix}1&0&1\\0&-1&0\\-1&0&1\end{pmatrix}$，线性变换$T$满足

$$\left.\begin{aligned}T(\overrightarrow{x_1}+2\overrightarrow{x_2}+3\overrightarrow{x_3})&=\overrightarrow{y_1}+\overrightarrow{y_2}\\T(2\overrightarrow{x_1}+\overrightarrow{x_2}+2\overrightarrow{x_3})&=\overrightarrow{y_2}+\overrightarrow{y_3}\\T(\overrightarrow{x_1}+3\overrightarrow{x_2}+4\overrightarrow{x_3})&=\overrightarrow{y_1}+\overrightarrow{y_3}\end{aligned}\right\}$$

（1）求T在基（Ⅱ）下的矩阵；（2）求$T(\overrightarrow{y_1})$在基（Ⅰ）下的坐标.

解：（1）根据已知条件可得$(T(\overrightarrow{x_1}),T(\overrightarrow{x_2}),T(\overrightarrow{x_3}))\begin{pmatrix}1&2&1\\2&1&3\\3&2&4\end{pmatrix}=(\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3})\begin{pmatrix}1&0&1\\1&1&0\\0&1&1\end{pmatrix}$，即$T(\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3})=(\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3})B$，其中$B=\begin{pmatrix}1&0&1\\1&1&0\\0&1&1\end{pmatrix}\begin{pmatrix}1&2&1\\2&1&3\\3&2&4\end{pmatrix}^{-1}=\begin{pmatrix}-1&-2&2\\-1&-5&4\\2&5&-4\end{pmatrix}$. 由于$(\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3})=(\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3})C$，$\therefore T(\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3})=T(\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3})C=(\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3})BC$，那么$T$在基（Ⅱ）下的矩阵为$A=BC=\begin{pmatrix}-1&-2&2\\-1&-5&4\\2&5&-4\end{pmatrix}\begin{pmatrix}1&0&1\\0&-1&0\\-1&0&1\end{pmatrix}=\begin{pmatrix}-3&2&1\\-5&5&3\\6&-5&-2\end{pmatrix}$.

（2）$\because T(\overrightarrow{y_1})=T(\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3})\begin{pmatrix}1\\0\\0\end{pmatrix}=(\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3})A\begin{pmatrix}1\\0\\0\end{pmatrix}=(\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3})CA\begin{pmatrix}1\\0\\0\end{pmatrix}=(\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3})\begin{pmatrix}3\\5\\9\end{pmatrix}$，

$\therefore T(\overrightarrow{y_1})$ 在基（Ⅰ）下的坐标为 $(3,5,9)^T$.

2.4 线性变换

一、线性变换的定义

定义 2-7 由 V 到 V 的线性算子 T 叫做 V 上的**线性变换**.
换言之，线性变换是线性空间 V 到自身的线性算子，与之对应的矩阵是方阵.

定义 2-8 如果对于任何 $\vec{x}\in V$ 恒有 $T(\vec{x})=\vec{x}$，则说 T 是**恒等变换**或称**单位变换**，记为 F . 与之对应的矩阵是单位矩阵 I 或 E . $TF=FT=T$（**单位变换的性质**）

定义线性变换的

加　法：$(T+B)\vec{\xi}=T(\vec{\xi})+B(\vec{\xi}),\vec{\xi}\in V$

乘　法：$(TB)(\vec{\alpha})=T\left(B\left(\vec{\alpha}\right)\right),\vec{\alpha}\in V$，$(TB)C=T(BC)$，$T(B+C)=TB+TC$

零变换：$\theta(\vec{\alpha})=\vec{0},\forall\vec{\alpha}\in V$，$T+\theta=T$

数　乘：$kT\equiv KT,K(\vec{\alpha})=k\vec{\alpha},(\forall\vec{\alpha}\in V)$

逆变换：$TB=BT=F$，则称 B 为 T 的逆变换，记为 T^{-1} . 如果线性变换 T 是可逆的，那么它的逆变换 T^{-1} 也是线性变换.

幂：$T^0\equiv F,T^m\equiv\overbrace{TT\cdots T}^{m}$，$T^{-n}\equiv(T^{-1})^n(n\in Z^+)$，一般 $(TB)^n\neq T^nB^n$

定理 2-5 设 $\overrightarrow{\xi_1},\overrightarrow{\xi_2},\cdots,\overrightarrow{\xi_n}$ 是数域 P 上 n 维线性空间 V 的一组基，在这组基下，每个线性变换对应一个 n 阶矩阵. 这个对应具有以下性质：
(1) 线性变换的和对应矩阵的和；(2) 线性变换的乘积对应矩阵的乘积；
(3) 线性变换与数的乘积对应矩阵与数的乘积；
(4) 可逆的线性变换与可逆矩阵对应，求逆变换对应着逆矩阵.

定义 2-9 设 T 是数域 P 线性空间 V 的线性变换. 如果对于 $\lambda_0\in P$，存在 $\vec{0}\neq\vec{\xi}\in V$ 使得 $T(\vec{\xi})=\lambda_0\vec{\xi}$，那么 λ_0 称为 T 的一个特征值，而 $\vec{\xi}$ 称为 T 的属于特征值 λ_0 的特征向量.

设$\overrightarrow{\xi_1},\overrightarrow{\xi_2},\cdots,\overrightarrow{\xi_n}$是$V$的一组基，线性变换$T$在基$\overrightarrow{\xi_1},\overrightarrow{\xi_2},\cdots,\overrightarrow{\xi_n}$下的矩阵是$A$，若$\lambda_0$是$T$的一个特征值，则它一定是矩阵$A$的一个特征值.

例： 设在R^3中定义线性变换T：它将基$\overrightarrow{e_1}=(-1,0,2),\overrightarrow{e_2}=(0,1,1),\overrightarrow{e_3}=(3,-1,0)$变为$T(\overrightarrow{e_1})=(-5,0,3),T(\overrightarrow{e_2})=(0,-1,6),T(\overrightarrow{e_3})=(-5,-1,9)$，试求：

（1）线性变换T在基$\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}$下的矩阵；（2）线性变换$T$在自然基$\overrightarrow{\varepsilon_1},\overrightarrow{\varepsilon_2},\overrightarrow{\varepsilon_3}$下的矩阵.

解：（1）不妨令$T(\overrightarrow{e_i})=k_{1i}\overrightarrow{e_1}+k_{2i}\overrightarrow{e_2}+k_{3i}\overrightarrow{e_3}(i=1,2,3)$，于是可得：

$\begin{cases}T(\overrightarrow{e_1})=k_{11}\overrightarrow{e_1}+k_{21}\overrightarrow{e_2}+k_{31}\overrightarrow{e_3}\\T(\overrightarrow{e_2})=k_{12}\overrightarrow{e_1}+k_{22}\overrightarrow{e_2}+k_{32}\overrightarrow{e_3}\\T(\overrightarrow{e_3})=k_{13}\overrightarrow{e_1}+k_{23}\overrightarrow{e_2}+k_{33}\overrightarrow{e_3}\end{cases}$，得到三个三阶线性代数方程组$\begin{cases}-k_{11}+3k_{31}=-5\\k_{21}-k_{31}=0\\2k_{11}+k_{21}=3\end{cases}$，

$\begin{cases}-k_{12}+3k_{32}=0\\k_{22}-k_{32}=-1\\2k_{12}+k_{22}=6\end{cases}$，$\begin{cases}-k_{13}+3k_{33}=-5\\k_{23}-k_{33}=-1\\2k_{13}+k_{23}=9\end{cases}$，分别解得$\begin{cases}k_{11}=2\\k_{21}=-1\\k_{31}=-1\end{cases}$，$\begin{cases}k_{12}=3\\k_{22}=0\\k_{32}=1\end{cases}$，$\begin{cases}k_{13}=5\\k_{23}=-1\\k_{33}=0\end{cases}$，

所以$\begin{cases}T(\overrightarrow{e_1})=2\overrightarrow{e_1}-\overrightarrow{e_2}-\overrightarrow{e_3}\\T(\overrightarrow{e_2})=3\overrightarrow{e_1}+0\overrightarrow{e_2}+\overrightarrow{e_3}\\T(\overrightarrow{e_3})=5\overrightarrow{e_1}-\overrightarrow{e_2}+0\overrightarrow{e_3}\end{cases}$，所以线性变换$T$在基$\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}$下的矩阵为$A=\begin{pmatrix}2&3&5\\-1&0&-1\\-1&1&0\end{pmatrix}$

（2）根据已知条件可得线性变换T在自然基$\overrightarrow{\varepsilon_1},\overrightarrow{\varepsilon_2},\overrightarrow{\varepsilon_3}$下的矩阵为$A_0=CAC^{-1}$，其中

$$C=\begin{pmatrix}-1&0&3\\0&1&-1\\2&1&0\end{pmatrix}，那么C^{-1}=\begin{pmatrix}-\frac{1}{7}&-\frac{3}{7}&\frac{3}{7}\\\frac{2}{7}&\frac{6}{7}&\frac{1}{7}\\\frac{2}{7}&-\frac{1}{7}&\frac{1}{7}\end{pmatrix}，\therefore A_0=\begin{pmatrix}-\frac{5}{7}&\frac{20}{7}&-\frac{20}{7}\\-\frac{4}{7}&-\frac{5}{7}&-\frac{2}{7}\\\frac{27}{7}&\frac{18}{7}&\frac{24}{7}\end{pmatrix}.$$

二、相似矩阵的几何解释

定理 2-6 假定线性空间V^n上的线性变换T对于基$\overrightarrow{e_1},\overrightarrow{e_2},\cdots,\overrightarrow{e_n}$下的矩阵为$A$，而对于另一组基$\overrightarrow{e_1'},\overrightarrow{e_2'},\cdots,\overrightarrow{e_m'}$下的矩阵为$B$，且由基$\overrightarrow{e_1},\overrightarrow{e_2},\cdots,\overrightarrow{e_n}$到基$\overrightarrow{e_1'},\overrightarrow{e_2'},\cdots,\overrightarrow{e_m'}$的过渡矩阵为$C$，则有$B=C^{-1}AC$.

定义 2-10 如果 A 与 B 是数域 P 上的两个 n 阶矩阵，且可以找到 P 上的 n 阶非奇异矩阵 C，使得 $B=C^{-1}AC$，则称 A 与 B **相似**．记为 $A\sim B$．

★ 线性变换在不同基下的矩阵是相似的；反之，如果两个矩阵相似，则它们可以看成同一个线性变换在两组不同基下的矩阵．

定义 2-11 设 A 与 B 都是 $m\times n$ 矩阵，如果存在非奇异的 m 阶方阵 D 和 n 阶方阵 C，使 $B=DAC$ 成立，则称 A 与 B 是**相抵**的，记为 $A\simeq B$．

相抵关系在几何上可解释为：在两个不同维的线性空间 V^n 和 V^m 中，同一个线性算子 T 在不同的基偶下所对应的矩阵 A 与 B 之间的关系．

定义 2-12 设 A 与 B 是两个 n 阶方阵，如果存在非奇异的 n 阶方阵 C，使得 $B=C^TAC$，则称 A 与 B 是**相合**（或**合同**）的．

2.5 线性变换的不变子空间

定义 2-13 设 V 是数域 P 上的 n 维线性空间，T 是 V 上的线性变换，V_1 是 V 的子空间，如果 $\forall\vec{x}\in V_1$，恒有 $T(\vec{x})\in V_1$（或 $T(V_1)\subseteq V_1$），则说 V_1 是关于 T 的**不变子空间**．

★ $P[x]_{n-1}$ 是 D 的不变子空间．

设 T 是线性空间 V 上的一个线性变换，V 中所有向量的像构成的集合称为线性变换 T 的**值域**，记为 $R(T)$，即 $R(T)=\{\vec{y}=T(\vec{x})\,|\,\vec{x}\in V\}$，所有被 T 变成零向量的原像构成的集合称为 T 的**核**，记作 $N(T)$，即 $N(T)=\{\vec{x}\in V\,|\,T(\vec{x})=\vec{0}\}$，易知，$T$ 的值域和核都是 T 的不变子空间．

一般称 $R(T)$ 的维数是线性变换 T 的**秩**，称 $N(T)$ 的维数是 T 的**零度**，记作 $null(T)$．所以，有时又称 $R(T)$ 为 T 的秩空间，$N(T)$ 为 T 的核空间．

定义 2-14 以 C^m 表示全体 m 维复向量在复数域 C 上做成的线性空间；A 为 $m\times n$ 复矩阵，其列向量为 $\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\cdots,\overrightarrow{\alpha_n}$，显然 $\overrightarrow{\alpha_i}\in C^m,i=1,2,\cdots,n$．子空间 $Span(\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\cdots,\overrightarrow{\alpha_n})$ 称为矩阵 A 的列空间（值域），记作 $R(A)$，即 $R(A)=Span(\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\cdots,\overrightarrow{\alpha_n})$，记 $A=(\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\cdots,\overrightarrow{\alpha_n})$，

$\vec{y}=(y_1,y_2,\cdots,y_n)^T\in C^n$，则 $R(A)$ 可表示成 $R(A)=\{A\vec{y}\,|\,\vec{y}\in C^n\}$，根据定义，显然有 $rank(A)=\dim(R(A))$.

定义 2-15 设 A 为 $m\times n$ 复矩阵，称线性方程组 $A\vec{x}=\vec{0}$ 在复数域上的解空间为 A 的化零空间（核），记作 $N(A)$，即 $N(A)=\{\vec{x}\,|\,A\vec{x}=\vec{0}\}$，显然 $N(A)$ 是 C^n 的一个字空间．且有 $null(A)=\dim(N(A))$.

★ 矩阵的秩与零度之间存在关系 $rank(A)+null(A)=n$.

设 T 是数域 P 上线性空间 V 上的一个线性变换，$\lambda_0\in P$ 是 T 的一个特征值．那么，全体数域特征值 λ_0 的 T 的特征向量和零向量作成的集合是 V 的一个子空间，称为线性变换 T 的一个（属于 λ_0 的）**特征子空间**，记作 V_{λ_0}． V_{λ_0} 是 T 的不变子空间.

定理 2-7 设 T 是 V^n 的线性变换，且 V^n 可分解为 s 个 T 的不变子空间的直和 $V^n=V_1\oplus V_2\oplus\cdots\oplus V_s$，又在每一个不变子空间 V_i 中取基 $\overrightarrow{e_{i1}},\overrightarrow{e_{i2}},\cdots,\overrightarrow{e_{in_i}},i=1,2,\cdots,s$，其中 $n_1+n_2+\cdots n_s=n$．把这些基集中起来作为 V^n 的基，则在该基下 T 的矩阵是准对角矩阵 $A=\begin{pmatrix}A_1 & & & \\ & A_2 & & \\ & & \cdots & \\ & & & A_s\end{pmatrix}$，其中 $A_i(i=1,2,\cdots,s)$ 就是 T 在 V_i 的基 $\overrightarrow{e_{i1}},\overrightarrow{e_{i2}},\cdots,\overrightarrow{e_{in_i}},i=1,2,\cdots,s$ 下的矩阵.

例： 设 R^3 中的向量为 $\vec{\alpha}=(\xi_1,\xi_2,\xi_3)$，线性变换为 $T(\vec{\alpha})=(-2\xi_2-2\xi_3,-2\xi_1+3\xi_2-\xi_3,-2\xi_1-\xi_2+3\xi_3)$，求 R^3 的一个基，使 T 在该基下的矩阵为对角矩阵.

解：取 R^3 的简单基 $\vec{e_1},\vec{e_2},\vec{e_3}$，根据已知条件可得

$T(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3})=(-2e_2-2e_3,-2e_1+3e_2-e_3,-2e_1-e_2+3e_3)=(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3})\begin{pmatrix}0&-2&-2\\-2&3&-1\\-2&-1&3\end{pmatrix}$

那么，T 在简单基 $\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}$ 下的矩阵为 $A=\begin{pmatrix}0&-2&-2\\-2&3&-1\\-2&-1&3\end{pmatrix}$. 下面求 A 的特征值及特征向量. $|A-\lambda E|=\begin{vmatrix}-\lambda&-2&-2\\-2&3-\lambda&-1\\-2&-1&3-\lambda\end{vmatrix}=-(\lambda-4)^2(\lambda+2)=0$，特征值为 $\lambda_1=\lambda_2=4,$ $\lambda_3=-2$. 当 $\lambda=4$ 时，$A-4E=\begin{pmatrix}-4&-2&-2\\-2&-1&-1\\-2&-1&-1\end{pmatrix}\sim\begin{pmatrix}2&1&1\\0&0&0\\0&0&0\end{pmatrix}$，对应的特征向量为 $(1,-2,0)^T$ 和 $(1,0,-2)^T$，当 $\lambda=-2$ 时，$A+2E=\begin{pmatrix}2&-2&-2\\-2&5&-1\\-2&-1&5\end{pmatrix}\sim\begin{pmatrix}1&-1&-1\\0&1&-1\\0&0&0\end{pmatrix}$，对应的特征向量为 $(2,1,1)^T$，令 $P=\begin{pmatrix}1&1&2\\-2&0&1\\0&-2&1\end{pmatrix},\Lambda=\begin{pmatrix}4&&\\&4&\\&&-2\end{pmatrix}$，则有 $P^{-1}AP=\Lambda$.

由 $(\overrightarrow{\beta_1},\overrightarrow{\beta_2},\overrightarrow{\beta_3})=(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3})P$ 求得 R^3 的另一个基为 $\overrightarrow{\beta_1}=\overrightarrow{e_1}-2\overrightarrow{e_2}=(1,-2,0)^T$

$\overrightarrow{\beta_2}=\overrightarrow{e_1}-2\overrightarrow{e_3}=(1,0,-2)^T,\overrightarrow{\beta_3}=2\overrightarrow{e_1}+\overrightarrow{e_2}+\overrightarrow{e_3}=(2,1,1)^T$，$T$ 在该基下的矩阵为 Λ.

例： 已知 $B=\begin{pmatrix}1&1\\0&1\end{pmatrix}$，线性空间 $V=\{\overrightarrow{X}=(x_{ij})_{2\times2}\,|x_{11}+x_{22}=0,x_{ij}\in R\}$ 的变换为 $T(\overrightarrow{X})=B^T\overrightarrow{X}-\overrightarrow{X}^TB(\overrightarrow{X}\in V)$

(1) 验证 T 是线性变换；(2) 求 V 的一个基，使 T 在该基下的矩阵为对角矩阵.

解：（1）$\because\forall\overrightarrow{X},\overrightarrow{Y}\in V$ 及 $k,l\in R$ 有 $T(k\overrightarrow{X}+l\overrightarrow{Y})=B^T(k\overrightarrow{X}+l\overrightarrow{Y})-(k\overrightarrow{X}+l\overrightarrow{Y})^TB$

$=k(B^T\overrightarrow{X}-\overrightarrow{X}^TB)+l(B^T\overrightarrow{Y}-\overrightarrow{Y}^TB)=kT(\overrightarrow{X})+lT(\overrightarrow{Y})$，$\therefore T$ 是线性变换.

（2）取 V 的简单基 $\overrightarrow{X_1}=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\overrightarrow{X_2}=\begin{pmatrix}0&1\\0&0\end{pmatrix},\overrightarrow{X_3}=\begin{pmatrix}0&0\\1&0\end{pmatrix}$，计算得到

$T(\overrightarrow{X_1})=\begin{pmatrix}0&-1\\1&0\end{pmatrix},T(\overrightarrow{X_2})=\begin{pmatrix}0&1\\-1&0\end{pmatrix},T(\overrightarrow{X_3})=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$. T 在简单基 $\overrightarrow{X_1},\overrightarrow{X_2},\overrightarrow{X_3}$ 下的矩阵为 $A=\begin{pmatrix}0&0&0\\-1&1&-1\\1&-1&1\end{pmatrix}$. $|A-\lambda E|=\begin{vmatrix}-\lambda&0&0\\-1&1-\lambda&-1\\1&-1&1-\lambda\end{vmatrix}=\lambda^2(2-\lambda)=0$，$\therefore A$ 的特征值为 $\lambda_1=\lambda_2=0,\lambda_3=2$，与之对应的特征向量为 $(1,1,0)^T,(0,1,1)^T,(0,1,-1)^T$，令 $P=\begin{pmatrix}1&0&0\\1&1&1\\0&1&-1\end{pmatrix},\Lambda=\begin{pmatrix}0&&\\&0&\\&&2\end{pmatrix}$，则有 $P^{-1}AP=\Lambda$. 由 $(\overrightarrow{Y_1},\overrightarrow{Y_2},\overrightarrow{Y_3})=(\overrightarrow{X_1},\overrightarrow{X_2},\overrightarrow{X_3})P$ 求得 V 的另一个基为 $\overrightarrow{Y_1}=\overrightarrow{X_1}+\overrightarrow{X_2}=\begin{pmatrix}1&1\\0&-1\end{pmatrix},\overrightarrow{Y_2}=\overrightarrow{X_2}+\overrightarrow{X_3}=\begin{pmatrix}0&1\\1&0\end{pmatrix}$,

$\overrightarrow{Y_3}=\overrightarrow{X_2}-\overrightarrow{X_3}=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$，$T$ 在该基下的矩阵为 Λ.

例： 设线性空间 V^n 的线性变换 T_1 与 T_2 满足 $T_1T_2=T_1+T_2$，证明：

(1) 1 不是 T_1 的特征值；(2) 若 T_1 的 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 互不相同，则存在 V^n 的一个基，使 T_1 与 T_2 在该基下的矩阵是对角矩阵.

证：取 V^n 的一个基为 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n}$，设 $\begin{matrix}T_1(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})=(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})A\\T_2(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})=(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})B\end{matrix}$，则有

$(T_1T_2)(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})=(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})(AB),(T_1+T_2)(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})=(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})(A+B)$,

由于 $T_1T_2=T_1+T_2$，$\therefore AB=A+B\Rightarrow B^TA^T=A^T+B^T$.

（1）假设 1 是 T_1 的特征值，那么 1 也是 A 的特征值，从而 1 也是 A^T 的特征值. 设 A^T 的对应于特征值 1 的特征向量为 $\vec{\beta}$，即 $A^T\vec{\beta}=\vec{\beta}(\vec{\beta}\neq\vec{0})$，由 $(B^TA^T)\vec{\beta}=(A^T+B^T)\vec{\beta}$ $\Rightarrow B^T\vec{\beta}=\vec{\beta}+B^T\vec{\beta}\Rightarrow\vec{\beta}=\vec{0}$. 这与 $\vec{\beta}$ 是 A^T 的特征向量矛盾，故 1 不是 T_1 的特征值.

（2）因为 T_1 有 n 个互不相同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$，所以 T_1 有 n 个线性无关的特征向量. 设 T_1 的对应于特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的线性无关的特征向量为 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n}$，即

$T_1\overrightarrow{x_i}=\lambda_i\overrightarrow{x_i}(i=1,2,\cdots,n)$，那么以 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n}$ 作为 V^n 的基时，T_1 的矩阵 $A=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$．由 $AB=A+B$ 及 $\lambda_i\neq 1$ 可得

$B=(A-E)^{-1}A=diag(\frac{\lambda_1}{\lambda_1-1},\frac{\lambda_2}{\lambda_2-1},\cdots,\frac{\lambda_n}{\lambda_n-1})$，即 T_1 与 T_2 在该基下的矩阵是对角矩阵.

本章测试题

一、设$\overrightarrow{\alpha_1}=(1,2,1,0),\overrightarrow{\alpha_2}=(-1,1,1,1),\overrightarrow{\beta_1}=(2,-1,0,1),\overrightarrow{\beta_2}=(1,-1,3,7)$，求

$V_1=Span\{\overrightarrow{\alpha_1},\overrightarrow{\alpha_2}\}$与$V_2=Span\{\overrightarrow{\beta_1},\overrightarrow{\beta_2}\}$的和及交的维数和它们的基.

解：$V_1+V_2=Span\{\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\beta_1},\overrightarrow{\beta_2}\}$，考察$(\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\beta_1},\overrightarrow{\beta_2})=\begin{pmatrix}1&-1&2&1\\2&1&-1&-1\\1&1&0&3\\0&1&1&7\end{pmatrix}$

$\sim\begin{pmatrix}1&0&3&8\\0&1&0&4\\0&0&1&3\\0&0&0&0\end{pmatrix}$，$\therefore rank(\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\beta_1},\overrightarrow{\beta_2})=3$，向量组$\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\beta_1},\overrightarrow{\beta_2}$的一个最大无关组为$\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\beta_1}$，从而可得$V_1+V_2=Span\{\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\beta_1}\},\dim(V_1+V_2)=3$，$\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\beta_1}$即为它的一组基.

设有向量$\vec{\alpha}\in V_1\cap V_2$，则有$k_1,k_2,l_1,l_2\in R$满足$\vec{\alpha}=k_1\overrightarrow{\alpha_1}+k_2\overrightarrow{\alpha_2}=l_1\overrightarrow{\beta_1}+l_2\overrightarrow{\beta_2}$，即$k_1\overrightarrow{\alpha_1}+k_2\overrightarrow{\alpha_2}-l_1\overrightarrow{\beta_1}-l_2\overrightarrow{\beta_2}=\vec{0}$，于是可得$\begin{cases}k_1-k_2-2l_1-l_2=0\\2k_1+k_2+l_1+l_2=0\\k_1+k_2-3l_2=0\\k_2-l_1-7l_2=0\end{cases}$，系数矩阵

$A=\begin{pmatrix}1&-1&-2&-1\\2&1&1&1\\1&1&0&-3\\0&1&-1&-7\end{pmatrix}\sim\begin{pmatrix}1&0&0&1\\0&1&0&-4\\0&0&1&3\\0&0&0&0\end{pmatrix}$，$rankA=3$，基础解系为$(1,-4,3,-1)^T$，

$\therefore\dim(V_1\cap V_2)=n-rankA=4-3=1$，一组基为$\overrightarrow{\alpha_1}-4\overrightarrow{\alpha_2}=(5,-2,-3,-4)$.

二、设线性空间V^n的基为$\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n}$，$T$是$V^n$的线性变换，则

$R(T)=Span(T(\overrightarrow{x_1}),T(\overrightarrow{x_2}),\cdots,T(\overrightarrow{x_n}))$，证明：$\dim R(T)+\dim N(T)=n$.

证明：先证$R(T)\subset Span(T(\overrightarrow{x_1}),T(\overrightarrow{x_2}),\cdots,T(\overrightarrow{x_n}))$.

$\forall \vec{y} \in R(T) \Rightarrow \exists \vec{x} \in V^n, sb.\vec{y} = T(\vec{x})$

$\vec{x} = c_1\vec{x_1} + c_2\vec{x_2} + \cdots + c_n\vec{x_n} \Rightarrow$

$\vec{y} = c_1 T\left(\vec{x_1}\right) + c_2 T\left(\vec{x_2}\right) + \cdots + c_n T\left(\vec{x_n}\right) \in Span(T(\vec{x_1}), T(\vec{x_2}), \cdots, T(\vec{x_n}))$

再证 $R(T) \supset Span(T(\vec{x_1}), T(\vec{x_2}), \cdots, T(\vec{x_n}))$.

$\forall \vec{y} \in Span(T(\vec{x_1}), T(\vec{x_2}), \cdots, T(\vec{x_n})) \Rightarrow$

$\exists c_1, c_2, \cdots, c_n \in R^n, sb.\vec{y} = c_1 T\left(\vec{x_1}\right) + c_2 T\left(\vec{x_2}\right) + \cdots + c_n T\left(\vec{x_n}\right)$

$\vec{x_i} \in V^n \Rightarrow T(\vec{x_i}) \in R(T) \Rightarrow \vec{y} \in R(T)$

$\therefore R(T) = Span(T(\vec{x_1}), T(\vec{x_2}), \cdots, T(\vec{x_n}))$

设 $\dim N(T) = m$，且 $N(T)$ 的一组基为 $\vec{y_1}, \vec{y_2}, \cdots, \vec{y_m}$，扩充为 V^n 的基：

$\vec{y_1}, \vec{y_2}, \cdots, \vec{y_m}, \overrightarrow{y_{m+1}}, \cdots, \vec{y_n}$，则

$R(T) = Span(T\left(\vec{y_1}\right), T\left(\vec{y_2}\right), \cdots, T\left(\vec{y_m}\right), T\left(\overrightarrow{y_{m+1}}\right), \cdots, T\left(\vec{y_n}\right)) = Span(T\left(\overrightarrow{y_{m+1}}\right), \cdots, T\left(\vec{y_n}\right))$

设有数组 $k_{m+1}, \cdots, k_n$ 使得 $k_{m+1}T\left(\overrightarrow{y_{m+1}}\right) + \cdots + k_n T\left(\vec{y_n}\right) = \vec{0}$，

则 $T\left(k_{m+1}\overrightarrow{y_{m+1}} + \cdots + k_n\vec{y_n}\right) = \vec{0}$. $\therefore T$ 为线性变换，$\therefore k_{m+1}\overrightarrow{y_{m+1}} + \cdots + k_n\vec{y_n} \in N(T)$，

$\therefore k_{m+1}\overrightarrow{y_{m+1}} + \cdots + k_n\vec{y_n} = l_1\vec{y_1} + \cdots + l_m\vec{y_m}$，

即 $\therefore (-l_1)\vec{y_1} + \cdots + (-l_m)\vec{y_m} + k_{m+1}\overrightarrow{y_{m+1}} + \cdots + k_n\vec{y_n} = \vec{0}$

$\because \vec{y_1}, \vec{y_2}, \cdots, \vec{y_n}$ 为 V^n 的基，$\therefore k_{m+1} = \cdots = k_n = 0$，$\therefore T\left(\overrightarrow{y_{m+1}}\right), \cdots, T\left(\vec{y_n}\right)$ 线性无关，

从而 $\dim R(T) = n - m$，故 $\dim R(T) + \dim N(T) = n$.

三、向量空间 R^4 中，$\vec{x} = (\xi_1, \xi_2, \xi_3, \xi_4)$ 线性变换 T 为

$T(\vec{x}) = (\xi_1 + \xi_2 - 3\xi_3 - \xi_4, 3\xi_1 - \xi_2 - 3\xi_3 + 4\xi_4, 0, 0)$，求 $R(T)$ 和 $N(T)$ 的基与维数.

解：取 R^4 的简单基 $\vec{e_1}, \vec{e_2}, \vec{e_3}, \vec{e_4}$，则有 $T\left(\vec{e_1}\right) = (1,3,0,0), T\left(\vec{e_2}\right) = (1,-1,0,0)$

$T\left(\vec{e_3}\right) = (-3,-3,0,0), T\left(\vec{e_4}\right) = (-1,4,0,0)$，$T\left(\vec{e_1}, \vec{e_2}, \vec{e_3}, \vec{e_4}\right) = \left(\vec{e_1}, \vec{e_2}, \vec{e_3}, \vec{e_4}\right)A$，考察

$A = \begin{pmatrix} 1 & 1 & -3 & -1 \\ 3 & -1 & -3 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -3 & -1 \\ 0 & -4 & 6 & 7 \end{pmatrix}, \therefore rankA = 2$，$\therefore \dim R(T) = 4 - 2 = 2$，

一组基为$\boldsymbol{T}\left(\overrightarrow{e_1}\right),\boldsymbol{T}\left(\overrightarrow{e_2}\right)$.

而$\boldsymbol{N}(\boldsymbol{T})=\{\vec{x}\,|\,\boldsymbol{T}\left(\vec{x}\right)=\vec{0}\}=\{\vec{x}\,|\,\boldsymbol{A}(\xi_1,\xi_2,\xi_3,\xi_4)^T=\vec{0}\}$，

其基础解系为$\left(3,3,2,0\right)^T,\left(-3,7,0,4\right)^T$，

$\therefore \dim \boldsymbol{N}(\boldsymbol{T})=2$，一组基为$\left(3,3,2,0\right)^T,\left(-3,7,0,4\right)^T$.

四、设三维线性空间V^3上的线性变换T在基$\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}$下的矩阵$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

（1）求T在基$\overrightarrow{e_3},\overrightarrow{e_2},\overrightarrow{e_1}$下的矩阵；（2）求$T$在基$\overrightarrow{e_1},k\overrightarrow{e_2},\overrightarrow{e_3}$下的矩阵，$k\neq 0$；

（3）求T在基$\overrightarrow{e_1}+\overrightarrow{e_2},\overrightarrow{e_2},\overrightarrow{e_3}$下的矩阵.

解：由已知可得$\boldsymbol{T}\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}\right)=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}\right)\boldsymbol{A}$

（1）由于$\left(\overrightarrow{e_3},\overrightarrow{e_2},\overrightarrow{e_1}\right)=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}\right)\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},\boldsymbol{T}\left(\overrightarrow{e_3},\overrightarrow{e_2},\overrightarrow{e_1}\right)=\boldsymbol{T}\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}\right)\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

$=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}\right)\boldsymbol{A}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}=\left(\overrightarrow{e_3},\overrightarrow{e_2},\overrightarrow{e_1}\right)\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1}\boldsymbol{A}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$，所以$\boldsymbol{T}$在基$\overrightarrow{e_3},\overrightarrow{e_2},\overrightarrow{e_1}$

下的矩阵为$\boldsymbol{B}=\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1}\boldsymbol{A}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}=\begin{pmatrix} a_{33} & a_{32} & a_{31} \\ a_{23} & a_{22} & a_{21} \\ a_{13} & a_{12} & a_{11} \end{pmatrix}$.

（2）$\left(\overrightarrow{e_1},k\overrightarrow{e_2},\overrightarrow{e_3}\right)=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3}\right)\begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}$，所以$\boldsymbol{T}$在基$\overrightarrow{e_1},k\overrightarrow{e_2},\overrightarrow{e_3}$下的矩阵为：

$$\boldsymbol{C}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1}\boldsymbol{A}\begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}=\begin{pmatrix} a_{11} & ka_{12} & a_{13} \\ \frac{a_{21}}{k} & a_{22} & \frac{a_{23}}{k} \\ a_{31} & ka_{32} & a_{33} \end{pmatrix}.$$

（3）$\boldsymbol{T}$在基$\overrightarrow{e_1}+\overrightarrow{e_2},\overrightarrow{e_2},\overrightarrow{e_3}$下的矩阵为：

$$\boldsymbol{D}=\begin{pmatrix}1&0&0\\1&1&0\\0&0&1\end{pmatrix}^{-1}\boldsymbol{A}\begin{pmatrix}1&0&0\\1&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}\boldsymbol{a}_{11}+\boldsymbol{a}_{12}&\boldsymbol{a}_{12}&\boldsymbol{a}_{13}\\\boldsymbol{a}_{21}+\boldsymbol{a}_{22}-\boldsymbol{a}_{11}-\boldsymbol{a}_{12}&\boldsymbol{a}_{22}-\boldsymbol{a}_{12}&\boldsymbol{a}_{23}-\boldsymbol{a}_{13}\\\boldsymbol{a}_{31}+\boldsymbol{a}_{32}&\boldsymbol{a}_{32}&\boldsymbol{a}_{33}\end{pmatrix}.$$

五、在 V^3 中，线性变换 T 在基 $\overrightarrow{\eta_1}=(-1,1,1),\overrightarrow{\eta_2}=(1,0,-1),\overrightarrow{\eta_3}=(0,1,1)$ 下的矩阵是 $A=\begin{pmatrix}1&0&1\\1&1&0\\-1&2&0\end{pmatrix}$，求 T 在自然基 $\overrightarrow{\varepsilon_1},\overrightarrow{\varepsilon_2},\overrightarrow{\varepsilon_3}$ 下的矩阵 B.

解：易知 $\left(\overrightarrow{\boldsymbol{\eta}_1},\overrightarrow{\boldsymbol{\eta}_2},\overrightarrow{\boldsymbol{\eta}_3}\right)=\left(\overrightarrow{\boldsymbol{\varepsilon}_1},\overrightarrow{\boldsymbol{\varepsilon}_2},\overrightarrow{\boldsymbol{\varepsilon}_3}\right)\boldsymbol{C}$，其中 $\boldsymbol{C}=\begin{pmatrix}-1&1&0\\1&0&1\\1&-1&1\end{pmatrix}$.

且 $\boldsymbol{T}\left(\overrightarrow{\boldsymbol{\eta}_1},\overrightarrow{\boldsymbol{\eta}_2},\overrightarrow{\boldsymbol{\eta}_3}\right)=\left(\overrightarrow{\boldsymbol{\eta}_1},\overrightarrow{\boldsymbol{\eta}_2},\overrightarrow{\boldsymbol{\eta}_3}\right)\boldsymbol{A}$

$$\left(\overrightarrow{\boldsymbol{\varepsilon}_1},\overrightarrow{\boldsymbol{\varepsilon}_2},\overrightarrow{\boldsymbol{\varepsilon}_3}\right)=\left(\overrightarrow{\boldsymbol{\eta}_1},\overrightarrow{\boldsymbol{\eta}_2},\overrightarrow{\boldsymbol{\eta}_3}\right)\boldsymbol{C}^{-1}\Rightarrow\boldsymbol{T}\left(\overrightarrow{\boldsymbol{\varepsilon}_1},\overrightarrow{\boldsymbol{\varepsilon}_2},\overrightarrow{\boldsymbol{\varepsilon}_3}\right)=\boldsymbol{T}\left(\overrightarrow{\boldsymbol{\eta}_1},\overrightarrow{\boldsymbol{\eta}_2},\overrightarrow{\boldsymbol{\eta}_3}\right)\boldsymbol{C}^{-1}=\left(\overrightarrow{\boldsymbol{\eta}_1},\overrightarrow{\boldsymbol{\eta}_2},\overrightarrow{\boldsymbol{\eta}_3}\right)\boldsymbol{AC}^{-1}$$
$$=\left(\overrightarrow{\boldsymbol{\varepsilon}_1},\overrightarrow{\boldsymbol{\varepsilon}_2},\overrightarrow{\boldsymbol{\varepsilon}_3}\right)\boldsymbol{CAC}^{-1}$$

而 $\boldsymbol{C}^{-1}=\begin{pmatrix}-1&1&-1\\0&1&-1\\1&0&1\end{pmatrix}$

$$\therefore\boldsymbol{B}=\boldsymbol{CAC}^{-1}=\begin{pmatrix}-1&1&-2\\1&2&-1\\2&0&1\end{pmatrix}.$$

六、如果矩阵 A 非奇异，证明矩阵 AB 与 BA 相似.

证明：由于 $\boldsymbol{A}$ 非奇异，则 $\boldsymbol{A}^{-1}$ 存在，那么 $\boldsymbol{A}^{-1}(\boldsymbol{AB})\boldsymbol{A}=\left(\boldsymbol{A}^{-1}\boldsymbol{A}\right)(\boldsymbol{BA})=\boldsymbol{I}(\boldsymbol{BA})=\boldsymbol{BA}$，
$\therefore\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 相似.

七、如果矩阵 A 与 B 相似，C 与 D 相似，证明 $\begin{pmatrix}A&0\\0&C\end{pmatrix}$ 与 $\begin{pmatrix}B&0\\0&D\end{pmatrix}$ 相似.

证明：$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似 $\Rightarrow$ 存在可逆矩阵 $\boldsymbol{P}$，满足 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$；

$\boldsymbol{C}$ 与 $\boldsymbol{D}$ 相似 $\Rightarrow$ 存在可逆矩阵 $\boldsymbol{Q}$，满足 $\boldsymbol{Q}^{-1}\boldsymbol{C}\boldsymbol{Q}=\boldsymbol{D}$. 而 $\begin{pmatrix}\boldsymbol{P} & 0\\ 0 & \boldsymbol{Q}\end{pmatrix}$ 显然非奇异，从而有

$$\begin{pmatrix}\boldsymbol{P} & 0\\ 0 & \boldsymbol{Q}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{A} & 0\\ 0 & \boldsymbol{C}\end{pmatrix}\begin{pmatrix}\boldsymbol{P} & 0\\ 0 & \boldsymbol{Q}\end{pmatrix}=\begin{pmatrix}\boldsymbol{P}^{-1} & 0\\ 0 & \boldsymbol{Q}^{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{AP} & 0\\ 0 & \boldsymbol{CQ}\end{pmatrix}=\begin{pmatrix}\boldsymbol{P}^{-1}\boldsymbol{AP} & 0\\ 0 & \boldsymbol{Q}^{-1}\boldsymbol{CQ}\end{pmatrix}=\begin{pmatrix}\boldsymbol{B} & 0\\ 0 & \boldsymbol{D}\end{pmatrix},$$

$\therefore \begin{pmatrix}\boldsymbol{A} & 0\\ 0 & \boldsymbol{C}\end{pmatrix}$ 与 $\begin{pmatrix}\boldsymbol{B} & 0\\ 0 & \boldsymbol{D}\end{pmatrix}$ 相似.

八、已知 $\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}$ 是四维线性空间 V^4 的一组基，V^4 的线性变换 T 在这组基下的矩阵为

$A=\begin{pmatrix}1 & 0 & 2 & 1\\ -1 & 2 & 1 & 3\\ 1 & 2 & 5 & 5\\ 2 & -2 & 1 & -2\end{pmatrix}$.（1）求 T 在基 $\overrightarrow{e_1'}=\overrightarrow{e_1}-2\overrightarrow{e_2}+\overrightarrow{e_4},\overrightarrow{e_2'}=3\overrightarrow{e_2}-\overrightarrow{e_3}-\overrightarrow{e_4},$

$\overrightarrow{e_3'}=\overrightarrow{e_3}+\overrightarrow{e_4},\overrightarrow{e_4'}=2\overrightarrow{e_4}$ 下的矩阵；（2）求 T 的值域和核；（3）在 T 的核中选一组基，把它扩充成 V 的一组基，并求 T 在这组基下的矩阵；（4）在 T 的值域中选一组基，把它扩充成 V 的一组基，并求 T 在这组基下的矩阵.

解：（1）由已知可得 $\left(\overrightarrow{e_1'},\overrightarrow{e_2'},\overrightarrow{e_3'},\overrightarrow{e_4'}\right)=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}\right)\boldsymbol{C}$，其中 $\boldsymbol{C}=\begin{pmatrix}1 & 0 & 0 & 0\\ -2 & 3 & 0 & 0\\ 0 & -1 & 1 & 0\\ 1 & -1 & 1 & 2\end{pmatrix}$，

则 $\boldsymbol{C}^{-1}=\begin{pmatrix}1 & 0 & 0 & 0\\ \frac{2}{3} & \frac{1}{3} & 0 & 0\\ \frac{2}{3} & \frac{1}{3} & 1 & 0\\ -\frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2}\end{pmatrix}$，且 $\boldsymbol{T}\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}\right)=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}\right)\boldsymbol{A}$.

所以 $\boldsymbol{T}\left(\overrightarrow{e_1'},\overrightarrow{e_2'},\overrightarrow{e_3'},\overrightarrow{e_4'}\right)=\boldsymbol{T}\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}\right)\boldsymbol{C}=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}\right)\boldsymbol{AC}=\left(\overrightarrow{e_1'},\overrightarrow{e_2'},\overrightarrow{e_3'},\overrightarrow{e_4'}\right)\boldsymbol{C}^{-1}\boldsymbol{AC}$，则

$\boldsymbol{T}$ 在基 $\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}$ 下的矩阵为 $\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix} 2 & -3 & 3 & 2 \\ \frac{2}{3} & -\frac{4}{3} & \frac{10}{3} & \frac{10}{3} \\ \frac{10}{3} & -\frac{16}{3} & \frac{40}{3} & \frac{40}{3} \\ 0 & 1 & -7 & -6 \end{pmatrix}$.

（2）先求 $\boldsymbol{N}(\boldsymbol{T})$，设 $\vec{\xi}\in \boldsymbol{N}(\boldsymbol{T})$，它在 $\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}$ 下的坐标为 x_1,x_2,x_3,x_4，由于 $\boldsymbol{T}\left(\vec{\xi}\right)=\vec{0}$，则 $\boldsymbol{T}\left(\vec{\xi}\right)$ 在 $\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}$ 下的坐标为 $0,0,0,0$，$\therefore \boldsymbol{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\boldsymbol{\theta}$，其基础解系为 $\left(-2,-\frac{2}{3},1,0\right)^T,(-1,-2,0,1)^T$，$\therefore \overrightarrow{\xi_1}=-2\overrightarrow{e_1}-\frac{2}{3}\overrightarrow{e_2}+\overrightarrow{e_3},\overrightarrow{\xi_2}=-\overrightarrow{e_1}-2\overrightarrow{e_2}+\overrightarrow{e_4}$ 为 $\boldsymbol{N}(\boldsymbol{T})$ 的一组基，即 $\boldsymbol{N}(\boldsymbol{T})=\boldsymbol{Span}(\overrightarrow{\xi_1},\overrightarrow{\xi_2})$ 且 $\boldsymbol{null}(\boldsymbol{T})=2$，那么 $\dim \boldsymbol{R}(\boldsymbol{T})=4-\boldsymbol{null}(\boldsymbol{T})=2$，又 $\boldsymbol{T}\left(\overrightarrow{e_1}\right)=\overrightarrow{e_1}-\overrightarrow{e_2}+\overrightarrow{e_3}+2\overrightarrow{e_4},\boldsymbol{T}\left(\overrightarrow{e_2}\right)=2\overrightarrow{e_2}+2\overrightarrow{e_3}-2\overrightarrow{e_4}$，$\therefore \boldsymbol{T}\left(\overrightarrow{e_1}\right),\boldsymbol{T}\left(\overrightarrow{e_2}\right)$ 线性无关，从而 $\boldsymbol{R}(\boldsymbol{T})=\boldsymbol{Span}\left(\boldsymbol{T}\left(\overrightarrow{e_1}\right),\boldsymbol{T}\left(\overrightarrow{e_2}\right)\right)$.

（3）$\because \left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{\xi_1},\overrightarrow{\xi_2}\right)=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}\right)\begin{pmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & -2/3 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}\right)\boldsymbol{D}_1$，

而 $\begin{vmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & -2/3 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}=1\neq 0$，$\therefore \boldsymbol{D}_1$ 可逆，从而 $\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{\xi_1},\overrightarrow{\xi_2}$ 线性无关，即为 $\boldsymbol{V}$ 的一组基，$\boldsymbol{T}$ 在 $\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{\xi_1},\overrightarrow{\xi_2}$ 下的矩阵为 $\boldsymbol{D}_1^{-1}\boldsymbol{A}\boldsymbol{D}_1=\begin{pmatrix} 5 & 2 & 0 & 0 \\ 9/2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 2 & -2 & 0 & 0 \end{pmatrix}$.

（4）$\because \left(\boldsymbol{T}\left(\overrightarrow{e_1}\right),\boldsymbol{T}\left(\overrightarrow{e_2}\right),\overrightarrow{e_3},\overrightarrow{e_4}\right)=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}\right)\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 2 & -2 & 0 & 1 \end{pmatrix}=\left(\overrightarrow{e_1},\overrightarrow{e_2},\overrightarrow{e_3},\overrightarrow{e_4}\right)\boldsymbol{D}_2$，而

$\begin{vmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 2 & -2 & 0 & 1 \end{vmatrix} = 2 \neq 0$，$\therefore \boldsymbol{D}_2$ **可逆，从而** $\boldsymbol{T}\left(\overrightarrow{\boldsymbol{e}_1}\right), \boldsymbol{T}\left(\overrightarrow{\boldsymbol{e}_2}\right), \overrightarrow{\boldsymbol{e}_3}, \overrightarrow{\boldsymbol{e}_4}$ **线性无关，即为** $\boldsymbol{V}$ **的一组基，** $\boldsymbol{T}$ **在** $\boldsymbol{T}\left(\overrightarrow{\boldsymbol{e}_1}\right), \boldsymbol{T}\left(\overrightarrow{\boldsymbol{e}_2}\right), \overrightarrow{\boldsymbol{e}_3}, \overrightarrow{\boldsymbol{e}_4}$ **下的矩阵为** $\boldsymbol{D}_2^{-1}\boldsymbol{A}\boldsymbol{D}_2 = \begin{pmatrix} 5 & 2 & 2 & 1 \\ 9/2 & 1 & 3/2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

九、线性空间 V^n 中，设由基（Ⅰ）：$\overrightarrow{x_1}, \overrightarrow{x_2}, \cdots, \overrightarrow{x_n}$ 改变为基（Ⅱ）：$\overrightarrow{y_1}, \overrightarrow{y_2}, \cdots, \overrightarrow{y_n}$ 的过渡矩阵为 C，给定 n 阶矩阵 B，线性变换 T 满足：$(T(\overrightarrow{y_1}), T(\overrightarrow{y_2}), \cdots, T(\overrightarrow{y_n})) = (\overrightarrow{x_1}, \overrightarrow{x_2}, \cdots, \overrightarrow{x_n})B$，则 T 在基（Ⅱ）下的矩阵是______________________________

解： $\left(\overrightarrow{\boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_2}, \cdots, \overrightarrow{\boldsymbol{y}_n}\right) = \left(\overrightarrow{\boldsymbol{x}_1}, \overrightarrow{\boldsymbol{x}_2}, \cdots, \overrightarrow{\boldsymbol{x}_n}\right)\boldsymbol{C} \Rightarrow \boldsymbol{T}\left(\overrightarrow{\boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_2}, \cdots, \overrightarrow{\boldsymbol{y}_n}\right) = \boldsymbol{T}\left(\overrightarrow{\boldsymbol{x}_1}, \overrightarrow{\boldsymbol{x}_2}, \cdots, \overrightarrow{\boldsymbol{x}_n}\right)\boldsymbol{C}$

$$= \left(\overrightarrow{\boldsymbol{x}_1}, \overrightarrow{\boldsymbol{x}_2}, \cdots, \overrightarrow{\boldsymbol{x}_n}\right)\boldsymbol{B} \Rightarrow \begin{cases} \boldsymbol{T}\left(\overrightarrow{\boldsymbol{x}_1}, \overrightarrow{\boldsymbol{x}_2}, \cdots, \overrightarrow{\boldsymbol{x}_n}\right) = \left(\overrightarrow{\boldsymbol{x}_1}, \overrightarrow{\boldsymbol{x}_2}, \cdots, \overrightarrow{\boldsymbol{x}_n}\right)\boldsymbol{B}\boldsymbol{C}^{-1} \\ \boldsymbol{T}\left(\overrightarrow{\boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_2}, \cdots, \overrightarrow{\boldsymbol{y}_n}\right) = \left(\overrightarrow{\boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_2}, \cdots, \overrightarrow{\boldsymbol{y}_n}\right)\boldsymbol{C}^{-1}\boldsymbol{B} \end{cases}.$$

十、给定线性空间 V^6 的基 $\overrightarrow{x_1}, \overrightarrow{x_2}, \cdots, \overrightarrow{x_6}$，线性变换 T 满足：$T(\overrightarrow{x_i}) = \overrightarrow{x_i} + 2\overrightarrow{x_{7-i}} (i = 1, 2, \cdots, 6)$

1．求 T 的全部特征值与特征向量；

2．判断是否存在另一组基，使 T 在该基下的矩阵为对角矩阵？若存在，求之．

解：1．根据已知条件可得， $\boldsymbol{T}$ **在基** $\overrightarrow{\boldsymbol{x}_1}, \overrightarrow{\boldsymbol{x}_2}, \cdots, \overrightarrow{\boldsymbol{x}_6}$ **下的矩阵为**

$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$，**则有** $\left|\boldsymbol{\lambda E} - \boldsymbol{A}\right| = \left|(\lambda+1)^3(\lambda-3)^3\left[(\lambda-1)^2 - 2(\lambda-1) + 4\right]\right|$，

$\therefore \boldsymbol{A}$ **的特征值为** $\lambda_1 = \lambda_2 = \lambda_3 = -1, \lambda_4 = \lambda_5 = \lambda_6 = 3$，

即 $\boldsymbol{T}$ **的特征值为** $\lambda_1 = \lambda_2 = \lambda_3 = -1, \lambda_4 = \lambda_5 = \lambda_6 = 3$.

当 $\lambda = -1$ **时，一组基础解系为** $\overrightarrow{\boldsymbol{x}_1} + \overrightarrow{\boldsymbol{x}_6}, \overrightarrow{\boldsymbol{x}_2} + \overrightarrow{\boldsymbol{x}_5}, \overrightarrow{\boldsymbol{x}_3} + \overrightarrow{\boldsymbol{x}_4}$，**所对应的特征向量为：**

$\vec{y}=k_1\left(\overrightarrow{x_1}+\overrightarrow{x_6}\right)+k_2\left(\overrightarrow{x_2}+\overrightarrow{x_5}\right)+k_3\left(\overrightarrow{x_3}+\overrightarrow{x_4}\right)$（$k_1,k_2,k_3$ 不同时为零）

当 $\lambda=3$ 时，一组基础解系为 $\overrightarrow{x_1}-\overrightarrow{x_6},\overrightarrow{x_2}-\overrightarrow{x_5},\overrightarrow{x_3}-\overrightarrow{x_4}$，所对应的特征向量为：

$\vec{y}=k_4\left(\overrightarrow{x_1}-\overrightarrow{x_6}\right)+k_5\left(\overrightarrow{x_2}-\overrightarrow{x_5}\right)+k_6\left(\overrightarrow{x_3}-\overrightarrow{x_4}\right)$（$k_4,k_5,k_6$ 不同时为零）

2．根据上面的计算结果，T 在 $(\overrightarrow{x_1}+\overrightarrow{x_6},\overrightarrow{x_2}+\overrightarrow{x_5},\overrightarrow{x_3}+\overrightarrow{x_4},\overrightarrow{x_1}-\overrightarrow{x_6},\overrightarrow{x_2}-\overrightarrow{x_5},\overrightarrow{x_3}-\overrightarrow{x_4})$ 下

的矩阵为 $\Lambda=\begin{pmatrix}-1&&&&&\\&-1&&&&\\&&-1&&&\\&&&3&&\\&&&&3&\\&&&&&3\end{pmatrix}$.

十一、设线性空间 $V=\{f(t)\big|f(t)=k_1+k_2t+k_3t^2,k_i\in R\}$ 的线性变换 T 为 $T(f(t))=(k_2+k_3)+(k_1+k_3)t+(k_1+k_2)t^2$，求 V 的一个基，使 T 在该基下的矩阵为对角矩阵.

解：取简单基 $(1,t,t^2),\Rightarrow f(t)=(1,t,t^2)\begin{pmatrix}k_1\\k_2\\k_3\end{pmatrix}$,

且 $T\left(1,t,t^2\right)=\left(1,t,t^2\right)A,A=\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}$,

那么 $\left|\lambda E-A\right|=\left|(\lambda-2)(\lambda+1)^2\right|,\therefore\lambda_1=2,\lambda_2=\lambda_3=-1$，对应 $\lambda_1=2$ 的特征向量为

$(1,1,1)^T$，对应 $\lambda_2=\lambda_3=-1$ 的特征向量为 $(1,-1,0)^T,(1,0,-1)^T,\therefore P=\begin{pmatrix}1&1&1\\1&-1&0\\1&0&-1\end{pmatrix}$,

$\Lambda=\begin{pmatrix}2&&\\&-1&\\&&-1\end{pmatrix}$ 满足 $P^{-1}AP=\Lambda$，所求的基为 $\left(f_1,f_2,f_3\right)=\left(1,t,t^2\right)P$，即有：

$f_1(t)=1+t+t^2,f_2(t)=1-t,f_3(t)=1-t^2$.

十二、设 A 为 $m\times n$ 矩阵，B 为 $n\times k$ 矩阵，$R(A)$ 和 $R(AB)$ 分别表示 A 和 AB 的值域，证

明：$R(AB)=R(A)$ 的充要条件是存在 $k\times n$ 矩阵 C，使得 $ABC=A$.

证明：必要性.

划分 $\boldsymbol{A}=\left(\overrightarrow{\boldsymbol{\alpha}_1},\overrightarrow{\boldsymbol{\alpha}_2},\cdots,\overrightarrow{\boldsymbol{\alpha}_n}\right)$，由 $\overrightarrow{\boldsymbol{\alpha}_i}\in \boldsymbol{R}(\boldsymbol{A})=\boldsymbol{R}(\boldsymbol{AB})$ 可得

$\exists\overrightarrow{\boldsymbol{x}_i}\in \boldsymbol{C}^k, sb.\boldsymbol{\alpha}_i=\boldsymbol{AB}\overrightarrow{\boldsymbol{x}_i}(\boldsymbol{i}=1,2,\cdots,\boldsymbol{n})$，于是有

$\boldsymbol{A}=\left(\overrightarrow{\boldsymbol{\alpha}_1},\overrightarrow{\boldsymbol{\alpha}_2},\cdots,\overrightarrow{\boldsymbol{\alpha}_n}\right)=\boldsymbol{AB}\left(\overrightarrow{\boldsymbol{x}_1},\overrightarrow{\boldsymbol{x}_2},\cdots,\overrightarrow{\boldsymbol{x}_n}\right)=\boldsymbol{ABC}$，其中 $\boldsymbol{C}=\left(\overrightarrow{\boldsymbol{x}_1},\overrightarrow{\boldsymbol{x}_2},\cdots,\overrightarrow{\boldsymbol{x}_n}\right)$.

充分性．由 $\boldsymbol{ABC}=\boldsymbol{A}$ 得 $\boldsymbol{R}(\boldsymbol{A})=\boldsymbol{R}(\boldsymbol{ABC})\subset\boldsymbol{R}(\boldsymbol{AB})\subset\boldsymbol{R}(\boldsymbol{A})$，

$\boldsymbol{R}(\boldsymbol{AB})=\boldsymbol{R}(\boldsymbol{A})$.

十三、设 $C^{n\times n}$ 是 $n\times n$ 复矩阵全体在通常的运算下所构成的复数域 C 上的线性空间，

$F=\begin{pmatrix}0&0&\cdots&0&-a_n\\1&0&\cdots&0&-a_{n-1}\\0&1&\cdots&0&-a_{n-2}\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&1&-a_1\end{pmatrix}$.（1）假设 $A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_2&\cdots&a_{nn}\end{pmatrix}$，若 $AF=FA$，证明：

$A=a_{n1}F^{n-1}+a_{n-1,1}F^{n-2}+\cdots+a_{21}F+a_{11}E$；（2）求 $C^{n\times n}$ 的子空间

$C(F)=\{X\in C^{n\times n}\,|FX=XF\}$ 的维数.

证明：记

$A=(\alpha_1,\alpha_2,\cdots,\alpha_n),M=a_{n1}F^{n-1}+a_{n-1,1}F^{n-2}+\cdots+a_{21}F+a_{11}E,\vec{\beta}=(-a_n,-a_{n-1},\cdots,-a_1)^T$

则 $F=(\overrightarrow{e_2},\overrightarrow{e_3},\cdots,\overrightarrow{e_n},\vec{\beta})$，

而 $F\overrightarrow{e_1}=\overrightarrow{e_2},F^2\overrightarrow{e_1}=F\overrightarrow{e_2}=\overrightarrow{e_3},\cdots,F^{n-1}\overrightarrow{e_1}=F(F^{n-2}\overrightarrow{e_1})=F\overrightarrow{e_{n-1}}=\overrightarrow{e_n}$.

$$\begin{aligned}\therefore M\overrightarrow{e_1}&=(a_{n1}F^{n-1}+a_{n-1,1}F^{n-2}+\cdots+a_{21}F+a_{11}E)\overrightarrow{e_1}\\&=a_{n1}F^{n-1}\overrightarrow{e_1}+a_{n-1,1}F^{n-2}\overrightarrow{e_1}+\cdots+a_{21}F\overrightarrow{e_1}+a_{11}E\overrightarrow{e_1}\\&=a_{n1}\overrightarrow{e_n}+a_{n-1,1}\overrightarrow{e_{n-1}}+\cdots+a_{21}\overrightarrow{e_2}+a_{11}\overrightarrow{e_1}=\overrightarrow{\alpha_1}=A\overrightarrow{e_1}\end{aligned}$$

由此类推， $M\overrightarrow{e_2}=MF\overrightarrow{e_1}=FM\overrightarrow{e_1}=FA\overrightarrow{e_1}=AF\overrightarrow{e_1}=A\overrightarrow{e_2}$

$$M\overrightarrow{e_3}=MF^2\overrightarrow{e_1}=F^2M\overrightarrow{e_1}=F^2A\overrightarrow{e_1}=AF^2\overrightarrow{e_1}=A\overrightarrow{e_3}$$

……………………………

$$M\overrightarrow{e_n}=MF^{n-1}\overrightarrow{e_1}=F^{n-1}M\overrightarrow{e_1}=F^{n-1}A\overrightarrow{e_1}=AF^{n-1}\overrightarrow{e_1}=A\overrightarrow{e_n}$$

所以 $M=A$，也即 $A=a_{n1}F^{n-1}+a_{n-1,1}F^{n-2}+\cdots+a_{21}F+a_{11}E$

$C(F)=Span\{E,F,F^2,\cdots,F^{n-1}\}$，设 $x_0E+x_1F+x_2F^2+\cdots+x_{n-1}F^{n-1}=O$，等式两边同右乘 $\overrightarrow{e_1}$，利用 $F\overrightarrow{e_1}=\overrightarrow{e_2},F^2\overrightarrow{e_1}=F\overrightarrow{e_2}=\overrightarrow{e_3},\cdots,F^{n-1}\overrightarrow{e_1}=F(F^{n-2}\overrightarrow{e_1})=F\overrightarrow{e_{n-1}}=\overrightarrow{e_n}$ 可得

$$\theta=O\overrightarrow{e_1}=(x_0E+x_1F+x_2F^2+\cdots+x_{n-1}F^{n-1})\overrightarrow{e_1}=x_0E\overrightarrow{e_1}+x_1F\overrightarrow{e_1}+x_2F^2\overrightarrow{e_1}+\cdots+x_{n-1}F^{n-1}\overrightarrow{e_1}$$

$$=x_0\overrightarrow{e_1}+x_1\overrightarrow{e_2}+\cdots+x_{n-1}\overrightarrow{e_n}$$

由于 $\overrightarrow{e_1},\overrightarrow{e_2},\cdots,\overrightarrow{e_n}$ 线性无关，故 $x_0=x_1=x_2=\cdots=x_{n-1}=0$，所以 $E,F,F^2,\cdots,F^{n-1}$ 线性无关，因此 $E,F,F^2,\cdots,F^{n-1}$ 是 $C(F)$ 的基，$\dim C(F)=n$.

十四、A 半正定，求证：$|A+I|\geq 1$ 且等号成立的充要条件为 $A=O$.

证：存在正交矩阵 P，使 $P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n\end{pmatrix}$，$\lambda_i$ 为 A 的特征值，因 A 半正定，$\therefore \lambda_i\geq 0$，$\therefore |A+I|=\det\begin{pmatrix}\lambda_1+1 & & & \\ & \lambda_2+1 & & \\ & & \cdots & \\ & & & \lambda_n+1\end{pmatrix}\geq 1$，显然只有当 $\lambda_i=0$ 时，等号成立.

第二章　内积空间上的等积变换

§1　内积空间

1.1　内积与欧氏空间

定义 1-1　设 V 是实数域 R 上的线性空间，对于 $\forall \vec{x},\vec{y}\in V$（$\vec{y}$ 可以等于 $\vec{x}$）如能给定某种规则，使 $\vec{x}$ 与 $\vec{y}$ 对应着一个实数，记为 $(\vec{x},\vec{y})$，它能满足以下条件：

(1) $(\vec{x},\vec{y})=(\vec{y},\vec{x})$；(2) $(\vec{x}+\vec{y},\vec{z})=(\vec{x},\vec{z})+(\vec{y},\vec{z})$；(3) $(k\vec{x},\vec{y})=k(\vec{x},\vec{y}),\forall k\in R$；

(4) $(\vec{x},\vec{x})\geq 0$，当且仅当 $\vec{x}=\vec{0}$ 时，$(\vec{x},\vec{x})=0$.

则称该实数 $(\vec{x},\vec{y})$ 是向量 $\vec{x}$ 与 $\vec{y}$ 的**内积**．此时的实线性空间 V 叫做 *Euclid* **空间**，简称**欧氏空间**（或**实内积空间**）.

$\forall \vec{x},\vec{y}\in V,\vec{x}=(a_1,a_2,\cdots,a_n),\vec{y}=(b_1,b_2,\cdots,b_n)$，规定

$$(\vec{x},\vec{y})=a_1b_1+a_2b_2+\cdots+a_nb_n=\sum_{k=1}^{n}a_kb_k .$$

对于函数 $f(x),g(x)\in C[a,b](b>a)$ 定义内积为：$(f,g)=\int_a^b f(x)g(x)dx$.

内积的基本性质：

1° $(\vec{x},k\vec{y})=k(\vec{x},\vec{y})$；$2^\circ$ $(\vec{x},\vec{y}+\vec{z})=(\vec{x},\vec{y})+(\vec{x},\vec{z})$；$3^\circ$ $(\vec{x},\vec{0})=(\vec{0},\vec{x})=0$；

4° $(\sum_{i=1}^{n}\lambda_i\vec{x_i},\sum_{j=1}^{n}\mu_j\vec{y_j})=\sum_{i,j=1}^{n}\lambda_i\mu_j(\vec{x_i},\vec{y_j})$.

一、向量的长度与夹角

定义 1-2　非负实数 $\sqrt{(\vec{x},\vec{x})}$ 叫做向量 $\vec{x}$ 的**长度**或**模**，记为 $|\vec{x}|$．长度等于 1 的向量叫做单位向量．零向量的长度为 0.

$\vec{x}$ **单位化**或**规范化**，$\dfrac{\vec{x}}{|\vec{x}|}=\dfrac{\vec{x}}{\sqrt{(\vec{x},\vec{x})}}$.

定义 1-3 非零向量 $\vec{x}$ 与 $\vec{y}$ 的**夹角** $<\vec{x},\vec{y}>$ 规定为

$$<\vec{x},\vec{y}>=\arccos\frac{(\vec{x},\vec{y})}{|\vec{x}||\vec{y}|},0\le<\vec{x},\vec{y}>\le\pi .$$

柯西-许瓦兹 (Cauchy-Schwarz) **不等式**：$|(\vec{x},\vec{y})|\le|\vec{x}||\vec{y}|$ 对任意两个向量 $\vec{x}$ 与 $\vec{y}$ 均成立，

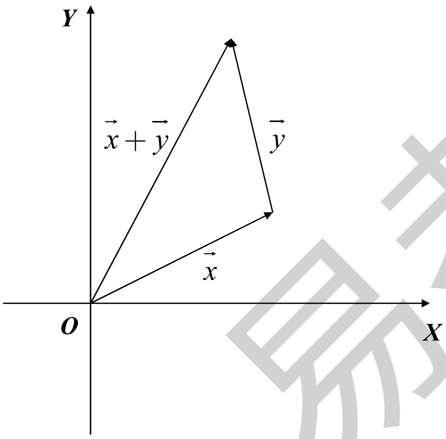
其中当且仅当 $\vec{x}$，$\vec{y}$ 线性相关时，等号才成立.

其几何解释是：$<\vec{x},\vec{y}>$ 满足 $|\cos<\vec{x},\vec{y}>|\le 1$.

Schwarz **不等式**：$\left|\sum_{i=1}^{n}x_iy_i\right|\le\sqrt{\sum_{i=1}^{n}x_i^2}\bullet\sqrt{\sum_{i=1}^{n}y_i^2}$，

$$\left|\int_a^b f(x)g(x)dx\right|\le\sqrt{\int_a^b f^2(x)dx}\bullet\sqrt{\int_a^b g^2(x)dx}$$

三角不等式及推广：$|\vec{x}+\vec{y}|\le|\vec{x}|+|\vec{y}|$



$$|\vec{x}+\vec{y}+\cdots+\vec{z}|\le|\vec{x}|+|\vec{y}|+\cdots+|\vec{z}|$$

$|\vec{x}-\vec{y}|\ge|\vec{x}|-|\vec{y}|$，$|\vec{x}-\vec{z}|\le|\vec{x}-\vec{y}|+|\vec{y}-\vec{z}|$，$|\vec{x}-\vec{y}|$ 称为向量 $\vec{x}$ 与 $\vec{y}$ 之间的距离.

二、度量矩阵

假定 $\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$ 是 n 维欧氏空间 V^n 的一组基．对于任意两个向量 $\vec{x},\vec{y}\in V^n$，有

$\vec{x}=x_1\vec{e_1}+x_2\vec{e_2}+\cdots+x_n\vec{e_n},\vec{y}=y_1\vec{e_1}+y_2\vec{e_2}+\cdots+y_n\vec{e_n}$，则根据内积的性质可得：

$(\vec{x},\vec{y})=(\sum_{i=1}^{n}x_i\vec{e_i},\sum_{j=1}^{n}y_j\vec{e_j})=\sum_{i,j=1}^{n}x_iy_j(\vec{e_i},\vec{e_j})$，令 $a_{ij}=(\vec{e_i},\vec{e_j})(i,j=1,2,\cdots,n)$，并构造矩

阵和列向量 $A=(a_{ij})=\begin{pmatrix}(\vec{e_1},\vec{e_1}) & (\vec{e_1},\vec{e_2}) & \cdots & (\vec{e_1},\vec{e_n})\\ (\vec{e_2},\vec{e_1}) & (\vec{e_2},\vec{e_2}) & \cdots & (\vec{e_2},\vec{e_n})\\ \cdots & \cdots & \cdots & \cdots\\ (\vec{e_n},\vec{e_1}) & (\vec{e_n},\vec{e_2}) & \cdots & (\vec{e_n},\vec{e_n})\end{pmatrix}$，$\vec{X}=(x_1,x_2,\cdots,x_n)^T$ 和

$\vec{Y}=(y_1,y_2,\cdots,y_n)^T$，于是内积 $(\vec{x},\vec{y})$ 可表示成 $(\vec{x},\vec{y})=\vec{X}^T A\vec{Y}$，$A\in R^{n\times n}$ 叫做基 $\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$ 的**度量矩阵**，又叫作 *Gram* **矩阵**.

度量矩阵的性质：

1° 度量矩阵是对称正定矩阵.

2° 两组不同基的度量矩阵是不同的，但它们之间呈相合关系.

三、正交性

（1）正交的概念

定义 1-4 设 $\vec{x}$，$\vec{y}$ 为欧氏空间的两个向量，如果 $(\vec{x},\vec{y})=0$，则说 $\vec{x}$ 与 $\vec{y}$ **正交**，记为 $\vec{x}\perp\vec{y}$.

正交与否与欧式空间的内积如何定义有关.

定理 1-1 如果向量 $\vec{x}$ 与 $\vec{y}$ 正交，则有 $\left|\vec{x}+\vec{y}\right|^2=\left|\vec{x}\right|^2+\left|\vec{y}\right|^2$.

定义 1-5 欧式空间中一组非零向量，如果它们两两正交，则称其为一个**正交向量组**.

若 $\vec{x_1},\vec{x_2},\cdots,\vec{x_m}$ 是正交向量组，则有 $\left|\vec{x_1}+\vec{x_2}+\cdots+\vec{x_m}\right|^2=\left|\vec{x_1}\right|^2+\left|\vec{x_2}\right|^2+\cdots+\left|\vec{x_m}\right|^2$.

定理 1-2 如果 $\vec{x_1},\vec{x_2},\cdots,\vec{x_m}$ 是一组两两正交的非零向量，它们必是线性无关的.

（2）*Schmidt* 正交化方法

设 $\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$ 是一组线性无关的向量. *Schmidt* 正交规范化步骤：

首先，把它们正交化；

第一步，取 $\vec{y_1'}=\vec{x_1}$；

第二步，令 $\vec{y_2'}=\vec{x_2}+k_{21}\vec{y_1'}$，由 $(\vec{y_2'},\vec{y_1'})=0$ 得 $(\vec{x_2}+k_{21}\vec{y_1'},\vec{y_1'})=(\vec{x_2},\vec{y_1'})$ $+k_{21}(\vec{y_1'},\vec{y_1'})=0$，则 $k_{21}=-\dfrac{(\vec{x_2},\vec{y_1'})}{(\vec{y_1'},\vec{y_1'})}$，于是 $\vec{y_2'}=\vec{x_2}-\dfrac{(\vec{x_2},\vec{y_1'})}{(\vec{y_1'},\vec{y_1'})}\vec{y_1'}$.

第三部，令 $\vec{y_3'}=\vec{x_3}+k_{31}\vec{y_2'}+k_{32}\vec{y_1'}$，采用第二步中的方法同样可以得到

$\overrightarrow{y_3}=\overrightarrow{x_3}-\frac{(\overrightarrow{x_3},\overrightarrow{y_2})}{(\overrightarrow{y_2},\overrightarrow{y_2})}\overrightarrow{y_2}-\frac{(\overrightarrow{x_3},\overrightarrow{y_1})}{(\overrightarrow{y_1},\overrightarrow{y_1})}\overrightarrow{y_1}$.

第四步，如此继续进行下去，得到一般式是：$\begin{cases}\overrightarrow{y_m}=\overrightarrow{x_m}+k_{m1}\overrightarrow{y_{m-1}}+k_{m2}\overrightarrow{y_{m-2}}+\cdots+k_{m,m-1}\overrightarrow{y_1}\\ k_{mi}=-\frac{(\overrightarrow{x_m},\overrightarrow{y_{m-i}})}{(\overrightarrow{y_{m-i}},\overrightarrow{y_{m-i}})},i=1,2,\cdots,m-1\end{cases}$ 直到 $m=n$，这样得到一组向量 $\overrightarrow{y_1},\overrightarrow{y_2},\cdots,\overrightarrow{y_n}$ 显然是两两正交的.

再单位化，$\overrightarrow{y_i}=\frac{\overrightarrow{y_i}}{|\overrightarrow{y_i}|},i=1,2,\cdots,n$.

（3）标准正交基

定义 1-6 在欧氏空间 V^n 中，由 n 个向量构成的正交向量组称为 V^n 的**正交基**；由单位向量构成的正交基叫做**标准正交基**.

定理 1-3 任何 n 维欧氏空间都有正交基和标准正交基.

定理 1-4 一组基为标准正交基的充要条件是，它的度量矩阵为单位矩阵.

例： 在实数域 R 上的多项式空间 $P[t]_n(n\geq 1)$ 中，对于多项式 $f(t)$ 与 $g(t)$，定义实数 $(f,g)=\int_0^1 f(t)g(t)dt$.

(1) 验证 (f,g) 是 $P[t]_n$ 中 $f(t)$ 与 $g(t)$ 的内积；

(2) 当 $n=1$ 时，取 $f(t)=t,g(t)=t+a$，问 a 为何值时，$f(t)$ 与 $g(t)$ 正交？

解：(1) 设 $h(t)\in P[t]_n,k\in R$，则有 $(f,g)=\int_0^1 f(t)g(t)dt=\int_0^1 g(t)f(t)dt=(g,f)$

$(f,g+h)=\int_0^1 f(t)[g(t)+h(t)]dt=\int_0^1 f(t)g(t)dt+\int_0^1 f(t)h(t)dt=(f,g)+(f,h)$

$(kf,g)=\int_0^1 kf(t)g(t)dt=k\int_0^1 f(t)g(t)dt=k(f,g)$

当 $f(t)=0$ 时，$(f,f)=0$；当 $f(t)\neq 0$ 时，$f(t)$ 在区间 $(0,1)$ 上不恒为零，从而有 $(f,f)=\int_0^1 f^2(t)dt>0$，因此，$(f,g)$ 是 $P[t]_n$ 中 $f(t)$ 与 $g(t)$ 的内积.

(2) 由 $(f,g)=0$ 可得 $\int_0^1 t(t+a)dt=(\frac{t^3}{3}+\frac{a}{2}t^2)|_0^1=\frac{1}{3}+\frac{a}{2}=0$，所以 $a=-\frac{2}{3}$.

例： 证明：在欧氏空间V^n中，采用矩阵乘法形式计算两个元素的内积时，其值与选取的基无关.

证明：设V^n的两个基为(Ⅰ)：$\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n}$ ；(Ⅱ) $\overrightarrow{y_1},\overrightarrow{y_2},\cdots,\overrightarrow{y_n}$，它们的度量矩阵分别为$A$和$B$，由基(Ⅰ)改变为基(Ⅱ)的过渡矩阵为$C$，$\vec{x}\in V^n$在基(Ⅰ)和基(Ⅱ)下的坐标（列向量）分别为$\overrightarrow{\alpha_1}$和$\overrightarrow{\alpha_2}$，$\vec{y}\in V^n$在基(Ⅰ)和基(Ⅱ)下的坐标（列向量）分别为$\overrightarrow{\beta_1}$和$\overrightarrow{\beta_2}$. 则需证明$\overrightarrow{\alpha_1}^T A\overrightarrow{\beta_1}=\overrightarrow{\alpha_2}^T B\overrightarrow{\beta_2}$.

由坐标变换公式得$\overrightarrow{\alpha_1}=C\overrightarrow{\alpha_2},\overrightarrow{\beta_1}=C\overrightarrow{\beta_2}$，而$B=C^TAC$，于是有

$$(\vec{x},\vec{y})=\overrightarrow{\alpha_1}^T A\overrightarrow{\beta_1}=(C\overrightarrow{\alpha_2})^T A(C\overrightarrow{\beta_2})=\overrightarrow{\alpha_2}^T(C^TAC)\overrightarrow{\beta_2}=\overrightarrow{\alpha_2}^T B\overrightarrow{\beta_2}$$

例： 设欧氏空间$P[t]_2$中的内积为$(f,g)=\int_{-1}^{1}f(t)g(t)dt$

(1) 求基$1,t,t^2$的度量矩阵；(2) 计算$f(t)=1-t+t^2$与$g(t)=1-4t-5t^2$的内积.

解：(1) 设基$1,t,t^2$的度量矩阵为$A=(a_{ij})$，则有$a_{11}=(1,1)=\int_{-1}^{1}dt=2$，

$a_{12}=(1,t)=\int_{-1}^{1}tdt=0$ ，$a_{13}=(1,t^2)=\int_{-1}^{1}t^2dt=\frac{2}{3}$ ，$a_{22}=(t,t)=\int_{-1}^{1}t^2dt=\frac{2}{3}$ ，

$a_{23}=(t,t^2)=\int_{-1}^{1}t^3dt=0$，$a_{33}=(t^2,t^2)=\int_{-1}^{1}t^4dt=\frac{2}{5}$. 根据度量矩阵的对称性可得：

$$A=\begin{pmatrix}2 & 0 & \frac{2}{3}\\ 0 & \frac{2}{3} & 0\\ \frac{2}{3} & 0 & \frac{2}{5}\end{pmatrix}.$$

(2) 易知$f(t),g(t)$在基$1,t,t^2$下的坐标分别为$\vec{\alpha}=(1,-1,1)^T,\vec{\beta}=(1,-4,-5)^T$，于是有

$$(f,g)=\vec{\alpha}^T A\vec{\beta}=(1,-1,1)\begin{pmatrix}2 & 0 & \frac{2}{3}\\ 0 & \frac{2}{3} & 0\\ \frac{2}{3} & 0 & \frac{2}{5}\end{pmatrix}\begin{pmatrix}1\\ -4\\ -5\end{pmatrix}=0.$$

例： 设向量空间 R^2 按照某种内积方式构成欧氏空间，记作 V^2．已知 V^2 的两个基为

（Ⅰ）$\vec{\alpha_1}=(1,1),\vec{\alpha_2}=(1,-1)$ （Ⅱ）$\vec{\beta_1}=(0,2),\vec{\beta_2}=(6,12)$

且 $\vec{\alpha_i}$ 与 $\vec{\beta_i}$ 的内积为 $(\vec{\alpha_1},\vec{\beta_1})=1,(\vec{\alpha_1},\vec{\beta_2})=15,(\vec{\alpha_2},\vec{\beta_1})=-1,(\vec{\alpha_2},\vec{\beta_2})=3$

(1) 求基(Ⅰ)的度量矩阵 A；(2) 求基(Ⅱ)的度量矩阵 B；(3) 求欧氏空间 V^2 的一个标准正交基．

解：（1）根据已知条件，需先将 $\vec{\alpha_i}$ 用 $\vec{\beta_i}$ 线性表示．不妨设 $\vec{\alpha_1}=k_1\vec{\beta_1}+k_2\vec{\beta_2}$，则有 $(1,1)=k_1(0,2)+k_2(6,12)$，解得 $k_1=-\frac{1}{2},k_2=\frac{1}{6}$，所以 $\vec{\alpha_1}=-\frac{1}{2}\vec{\beta_1}+\frac{1}{6}\vec{\beta_2}$，同理可有 $\vec{\alpha_2}=-\frac{3}{2}\vec{\beta_1}+\frac{1}{6}\vec{\beta_2}$.

$$\therefore(\vec{\alpha_1},\vec{\alpha_1})=(\vec{\alpha_1},-\frac{1}{2}\vec{\beta_1}+\frac{1}{6}\vec{\beta_2})=-\frac{1}{2}(\vec{\alpha_1},\vec{\beta_1})+\frac{1}{6}(\vec{\alpha_1},\vec{\beta_2})=-\frac{1}{2}\times1+\frac{1}{6}\times15=2$$

$$(\vec{\alpha_1},\vec{\alpha_2})=(\vec{\alpha_1},-\frac{3}{2}\vec{\beta_1}+\frac{1}{6}\vec{\beta_2})=-\frac{3}{2}(\vec{\alpha_1},\vec{\beta_1})+\frac{1}{6}(\vec{\alpha_1},\vec{\beta_2})=-\frac{3}{2}\times1+\frac{1}{6}\times15=1$$

$$(\vec{\alpha_2},\vec{\alpha_1})=(\vec{\alpha_1},\vec{\alpha_2})=1$$

$$(\vec{\alpha_2},\vec{\alpha_2})=(\vec{\alpha_2},-\frac{3}{2}\vec{\beta_1}+\frac{1}{6}\vec{\beta_2})=-\frac{3}{2}(\vec{\alpha_2},\vec{\beta_1})+\frac{1}{6}(\vec{\alpha_2},\vec{\beta_2})=-\frac{3}{2}\times(-1)+\frac{1}{6}\times3=2$$

$$\therefore A=\begin{pmatrix}2&1\\1&2\end{pmatrix}.$$

（2）由 $\vec{\alpha_1}=-\frac{1}{2}\vec{\beta_1}+\frac{1}{6}\vec{\beta_2}$，$\vec{\alpha_2}=-\frac{3}{2}\vec{\beta_1}+\frac{1}{6}\vec{\beta_2}$ 可得 $\vec{\beta_1}=\vec{\alpha_1}-\vec{\alpha_2},\vec{\beta_2}=9\vec{\alpha_1}-3\vec{\alpha_2}$，所以由基(Ⅰ)到基(Ⅱ)的过渡矩阵为 $C=\begin{pmatrix}1&9\\-1&-3\end{pmatrix}$，于是 $B=C^TAC=\begin{pmatrix}2&12\\12&126\end{pmatrix}$.

（3）对基(Ⅰ)进行 *Schmidt* 正交化可得

$$\vec{x_1}=\vec{\alpha_1}=(1,1)$$

$$\vec{x_2}=\vec{\alpha_2}-\frac{(\vec{\alpha_2},\vec{x_1})}{(\vec{x_1},\vec{x_1})}\vec{x_1}=\vec{\alpha_2}-\frac{(\vec{\alpha_2},\vec{\alpha_1})}{(\vec{\alpha_1},\vec{\alpha_1})}\vec{\alpha_1}=\vec{\alpha_2}-\frac{1}{2}\vec{\alpha_1}=(\frac{1}{2},-\frac{3}{2})$$

对 $\vec{x_1},\vec{x_2}$ 进行单位化可得

$$\vec{y_1}=\frac{\vec{x_1}}{|\vec{x_1}|}=\frac{\vec{x_1}}{\sqrt{(\vec{x_1},\vec{x_1})}}=\frac{\vec{x_1}}{\sqrt{(\vec{\alpha_1},\vec{\alpha_1})}}=\frac{\vec{x_1}}{\sqrt{2}}=(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})$$

$$\vec{y_2}=\frac{\vec{x_2}}{|\vec{x_2}|}=\frac{\vec{x_2}}{\sqrt{(\vec{x_2},\vec{x_2})}}=\frac{\vec{x_2}}{\sqrt{(\vec{\alpha_2}-\frac{1}{2}\vec{\alpha_1},\vec{\alpha_2}-\frac{1}{2}\vec{\alpha_1})}}=\frac{\vec{x_2}}{\sqrt{(\vec{\alpha_2},\vec{\alpha_2})-(\vec{\alpha_1},\vec{\alpha_2})+\frac{1}{4}(\vec{\alpha_1},\vec{\alpha_1})}}$$

$$=\frac{\vec{x_2}}{\sqrt{2-1+\frac{1}{4}\times 2}}=\sqrt{\frac{2}{3}}(\frac{1}{2},-\frac{3}{2})=(\frac{1}{\sqrt{6}},-\frac{3}{\sqrt{6}})$$

于是$\vec{y_1},\vec{y_2}$为欧氏空间V^2的一个标准正交基.

§2　等积变换及其矩阵

2.1　正交变换与正交矩阵

一、正交变换的定义及其性质

定义 2-1　设V是一个欧式空间，T是V上的线性变换，如果对于$\forall\vec{x},\vec{y}\in V$，变换$T$恒能使下式成立：$(T(\vec{x}),T(\vec{y}))=(\vec{x},\vec{y})$，则说$T$是$V$上的**正交变换**.

★ 恒等变换即是一个正交变换；坐标平面上的旋转变换也是一个正交变换.

定理 2-1　设T是欧式空间V上的线性变换，下面的任意一条都是使T成为正交变换的充要条件：1°　T使向量长度保持不变，即：对于$\forall\vec{x}\in V$，有$(T(\vec{x}),T(\vec{x}))=(\vec{x},\vec{x})$；

2°　任一组标准正交基经T变换后的基像仍是一组标准正交基；

3°　T在任一组标准正交基下的矩阵A满足$A^TA=AA^T=I$或$A^{-1}=A^T$.

★ 一组基为标准正交基的充要条件是它的度量矩阵为单位矩阵.

★ 两组标准正交基的过渡矩阵C为正交矩阵.（P87 例 2-7）

二、正交矩阵

定义 2-2　满足$A^TA=AA^T=I$的任何方阵A叫做**正交矩阵**.

★ 常用$A^TA=AA^T=I$来判断A是不是正交矩阵.

定理 2-2　在欧氏空间中，正交变换在标准正交基下的矩阵是正交矩阵；反过来，如果线

性变换 T 在标准正交基下的矩阵是正交矩阵，则 T 是正交变换.

★ 正交矩阵的一些常用性质：

1° 正交矩阵是非奇异的，且 $\det A=1$（或 -1）；

2° 正交矩阵的逆矩阵仍是正交矩阵；

3° 两个正交矩阵的乘积仍为正交矩阵；

4° 实数域上方阵 A 是正交矩阵的充要条件是： A 的行（或列）向量为标准正交向量.

2.2 两类常用的正交变换及其矩阵

一、初等旋转变换

一般地，在 n 维 $Euclid$ 空间 V 中取一组标准正交基 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$，沿平面 $[\vec{e}_i,\vec{e}_j]$ 旋转，它的矩阵表示是

$$R_{ij}=\left(\begin{array}{cccccccccccc}
1 & & & & & & & & & & & \\
 & \cdots & & & & & & & & & & \\
 & & \cdots & & & & & & & & & \\
 & & & 1 & & & & & & & & \\
 & & & & \cos\theta & \cdots & \cdots & \cdots & \sin\theta & & & \\
 & & & & \cdots & 1 & & & \cdots & & & \\
 & & & & \cdots & & \cdots & & \cdots & & & \\
 & & & & \cdots & & & 1 & \cdots & & & \\
 & & & & -\sin\theta & \cdots & \cdots & \cdots & \cos\theta & & & \\
 & & & & & & & & & 1 & & \\
 & & & & & & & & & & \cdots & \\
 & & & & & & & & & & & 1
\end{array}\right)\begin{array}{l}
\\ \\ \\ \\ \leftarrow R_i \\ \\ i<j \\ \\ \leftarrow R_j \\ \\ \\ \\
\end{array}$$

其元素满足 $r_{ii}=\cos\theta,r_{ij}=\sin\theta,r_{ji}=-\sin\theta,r_{jj}=\cos\theta,r_{pp}=1(p\neq i,j),r_{pq}=0$， $\theta-$ 旋转角，记 $\cos\theta=C,\sin\theta=S$，从而有

$$R_{ij}=\begin{pmatrix}1 & & & & & & & & & & \\ & \cdots & & & & & & & & & \\ & & \cdots & & & & & & & & \\ & & & 1 & & & & & & & \\ & & & & C & \cdots & \cdots & \cdots & S & & \\ & & & & \cdots & 1 & & & \cdots & & \\ & & & & \cdots & & \cdots & & \cdots & & \\ & & & & \cdots & & & 1 & \cdots & & \\ & & & & -S & \cdots & \cdots & \cdots & C & & \\ & & & & & & & & & 1 & \\ & & & & & & & & & & \cdots \\ & & & & & & & & & & & 1\end{pmatrix}$$

定义 2-4 R_{ij} 叫做初等旋转矩阵；它所确定的线性变换叫做**初等旋转变换**（*Givens* **变换**）.

定理 2-3 设 R_{ij} 是初等旋转矩阵，则有 $\det R_{ij}=1$；R_{ij} 对应的初等旋转变换是正交变换，R_{ij} 是正交矩阵.

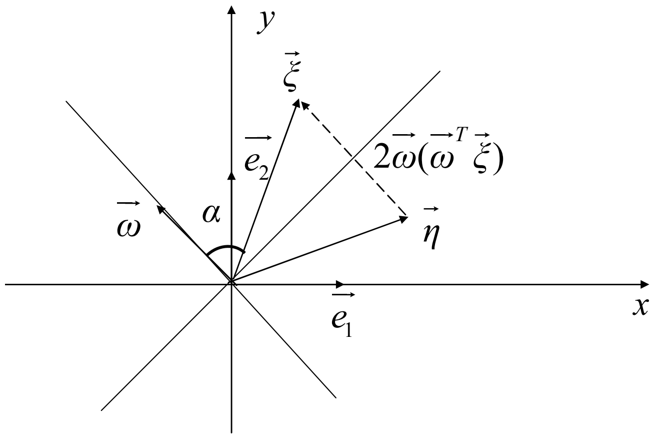
非零 n 维列（行）向量 $\vec{x}$，连续左（右）乘 $R_{r,r+1},R_{r,r+2},\cdots,R_{r,n}$，可使 $\vec{x}$ 的第 r 个分量变为一新的正数，而第 $r+1$ 个到第 n 个分量全部消为 0，即分别有

$$\vec{y}=R_{r,n}\cdots R_{r,r+2}R_{r,r+1}\vec{x}=\begin{pmatrix}x_1\\ \cdots\\ x_{r-1}\\ y_r\\ 0\\ \cdots\\ 0\end{pmatrix}$$（此时的 $C=\dfrac{x_i}{\sqrt{x_i^2+x_j^2}},S=\dfrac{x_j}{\sqrt{x_i^2+x_j^2}}$）

或 $\vec{y}=\vec{x}R_{r,r+1}R_{r,r+2}\cdots R_{r,n}=(x_1,\cdots,x_{r-1},y_r,0,\cdots,0)$

（此时的 $C=\dfrac{x_i}{\sqrt{x_i^2+x_j^2}},S=-\dfrac{x_j}{\sqrt{x_i^2+x_j^2}}$） $y_r=\sqrt{x_r^2+x_{r+1}^2+\cdots+x_n^2}$

二、镜像变换



根据几何关系可有：$\vec{\xi}-\vec{\eta}=2\vec{\omega}\cdot|\vec{\xi}|\cdot\cos\alpha=2\vec{\omega}\cdot|\vec{\xi}|\cdot\dfrac{(\vec{\omega},\vec{\xi})}{|\vec{\omega}|\cdot|\vec{\xi}|}=2\vec{\omega}(\vec{\omega},\vec{\xi})=2\vec{\omega}(\vec{\omega}^T\vec{\xi})$

（$\vec{\omega}$ 为单位坐标向量，$|\vec{\omega}|=1$）. $\vec{\omega}^T\vec{\xi}$ — 向量 $\vec{\xi}$ 在 $\vec{\omega}$ 方向上的投影长度（即内积）.

$\vec{\eta}=\vec{\xi}-2\vec{\omega}(\vec{\omega}^T\vec{\xi})=(I-2\vec{\omega}\vec{\omega}^T)\vec{\xi}$

令 $H=I-2\vec{\omega}\vec{\omega}^T$，则 $\vec{\eta}=H\vec{\xi}$，顺时针方向旋转向量 $\vec{\omega}$（角度 θ），使 $\vec{\omega}$ 与 $\vec{e}_2$ 重合，则使 $\vec{\omega}=\begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix}\to\begin{pmatrix}0\\1\end{pmatrix}=\vec{e}_2$ 的初等旋转矩阵为 $Q=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$，且 $Q^{-1}=Q^T$，于是有

$Q\vec{\eta}=QHI\vec{\xi}=QHQ^{-1}Q\vec{\xi}$，$QHQ^{-1}=Q(I-2\vec{\omega}\vec{\omega}^T)Q^T=I-2Q\vec{\omega}\vec{\omega}^TQ^T$

$=I-2(Q\vec{\omega})(Q\vec{\omega})^T=I-2\begin{pmatrix}0\\1\end{pmatrix}(0,1)=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$.

定义 2-5 在欧氏空间 R^n 中，设有变换将向量 $\vec{\xi}$ 映射成与单位向量 $\vec{\omega}$ 正交的 $n-1$ 维子空间对称的像 $\vec{\eta}$，且有 $\vec{\eta}=(I-2\vec{\omega}\vec{\omega}^T)\vec{\xi}=H\vec{\xi}$，则称这种变换为**镜像变换**，或**豪斯荷德**（$Householder$）**变换**，其中的矩阵 $H=I-2\vec{\omega}\vec{\omega}^T$ 称为**初等反射矩阵**，或 $Householder$ **矩阵**.

定理 2-4 设 $H=I-2\vec{\omega}\vec{\omega}^T$ 是 R^n 中的初等反射矩阵，则 H 是对称的正交矩阵；$\det H=-1$.

定理 2-5 镜像变换可以使任何非零向量 $\vec{\xi}$ 变到与给定单位向量 $\vec{\zeta}$ 同方向的向量 $\vec{\eta}$，且有

$\vec{\eta}=H\vec{\xi}=(I-2\vec{\omega}\vec{\omega}^T)\vec{\xi}$ （$\vec{\omega}=\dfrac{\vec{\xi}-\vec{\eta}}{2(\vec{\omega}^T\vec{\xi})}=\dfrac{\vec{\xi}-|\vec{\xi}|\vec{\zeta}}{|\vec{\xi}-|\vec{\xi}|\vec{\zeta}|}$）.

★ 任何实的非奇异矩阵都可以用镜像变换化为上三角矩阵.

定理 2-6 初等旋转矩阵（变换）是两个初等反射矩阵（变换）的乘积.

§3 其他几种线性变换及其矩阵

3.1 对称变换与厄尔密特变换

一、对称变换与对称矩阵

定义 3-1 设 T 是欧氏空间 V 中的一个线性变换，且对 $\forall\vec{x},\vec{y}\in V$ 都有 $(T(\vec{x}),\vec{y})=(\vec{x},T(\vec{y}))$ 成立，则称 T 是欧氏空间 V 中的**对称变换**.

定理 3-1 n 维欧氏空间 V 中的线性变换 T 是对称变换 $\Leftrightarrow$ T 在标准正交基下的矩阵 A 是对称矩阵，即有 $A^T=A$.

<u>矩阵在内积运算中的转移规则：</u>

1° 若 A 是对称矩阵，且 $A\in R^{n\times n},\vec{x}=X\in R^n,\vec{y}=Y\in R^n$，则 A 在内积中的转移规则为：

$(AX,Y)=(X,AY)$.

2° 若 A 不是对称矩阵，$A\in R^{m\times n},\vec{x}=X\in R^n,\vec{y}=Y\in R^m$，则有 $(AX,Y)=(X,A^TY)$.

二、厄尔密特变换与厄尔密特矩阵

定义 3-2 如果酉空间 V 上的线性变换 T 对 $\forall\vec{x},\vec{y}\in V$ 都有 $(T(\vec{x}),\vec{y})=(\vec{x},T(\vec{y}))$ 成立，则称 T 是**厄尔密特** $(Hermite)$ **变换**.

若厄尔密特变换 T 关于标准正交基的矩阵为 $A=(a_{ij})_{n\times n}$，则对于 $\forall i,j$ 恒有 $a_{ij}=\overline{a_{ji}}$ 或 $A^H=A$. A 称为**厄尔密特** $(Hermite)$ **矩阵**；又若满足 $A^H=-A$，则称 A 为**反厄尔密特** $(Hermite)$ **矩阵**.

引理 设$\overrightarrow{\xi_1}$为n维单位列向量，则存在n维列向量$\overrightarrow{\xi_1},\cdots,\overrightarrow{\xi_n}$使得矩阵$P=(\overrightarrow{\xi_1},\overrightarrow{\xi_2},\cdots,\overrightarrow{\xi_n})$为酉矩阵.

定理 3-2 (Schur 定理) 任何n阶矩阵都酉相似于一个上三角阵，即存在一个n阶酉矩阵U和一个上三角阵R，使得$U^H AU=R$(或$A=URU^H$). 式中R的主对角元是A的特征值，它们可以按所要求的次序排列.

推论 若A为$Hermite$矩阵，则A必酉相似于对角阵，其对角元(A的特征值)均为实数.

例： 设$\overrightarrow{x_0}$是欧氏空间V中的单位元素，定义变换$T(\vec{x})=\vec{x}-2(\vec{x},\overrightarrow{x_0})\overrightarrow{x_0}(\vec{x}\in V)$.

(1) 验证T是线性变换；(2) 验证T既是正交变换，又是对称变换；

(3) 验证$\overrightarrow{x_0}$是T的一个特征向量，并求其对应的特征值.

解：(1) 设$\vec{x},\vec{y}\in V,k,l\in R$，则有$T(k\vec{x}+l\vec{y})=(k\vec{x}+l\vec{y})-2(k\vec{x}+l\vec{y},\overrightarrow{x_0})\overrightarrow{x_0}$

$=k[\vec{x}-2(\vec{x},\overrightarrow{x_0})\overrightarrow{x_0}]+l[\vec{y}-2(\vec{y},\overrightarrow{x_0})\overrightarrow{x_0}]=kT(\vec{x})+lT(\vec{y})$，故$T$是线性变换.

(2) $\because (T(\vec{x}),T(\vec{x}))=(\vec{x},\vec{x})-4(\vec{x},\overrightarrow{x_0})(\vec{x},\overrightarrow{x_0})+4(\vec{x},\overrightarrow{x_0})^2(\overrightarrow{x_0},\overrightarrow{x_0})$且$(\overrightarrow{x_0},\overrightarrow{x_0})=1$

$\therefore (T(\vec{x}),T(\vec{x}))=(\vec{x},\vec{x})$，所以$T$是正交变换.

设$\vec{y}\in V$，则$T(\vec{y})=\vec{y}-2(\vec{y},\overrightarrow{x_0})\overrightarrow{x_0}$，于是有$(T(\vec{x}),\vec{y})=(\vec{x},\vec{y})-2(\vec{x},\overrightarrow{x_0})(\overrightarrow{x_0},\vec{y})$

$(\vec{x},T(\vec{y}))=(\vec{x},\vec{y})-2(\vec{y},\overrightarrow{x_0})(\vec{x},\overrightarrow{x_0})=(T(\vec{x}),\vec{y})$，故$T$又是对称变换.

(3) 根据已知条件可得$T(\overrightarrow{x_0})=\overrightarrow{x_0}-2(\overrightarrow{x_0},\overrightarrow{x_0})\overrightarrow{x_0}=(-1)\overrightarrow{x_0}$，故$\overrightarrow{x_0}$是$T$的对应于特征值$-1$的特征向量.

例： 证明：欧氏空间V^n的线性变换T为反对称变换，即

$(T(\vec{x}),\vec{y})=-(\vec{x},T(\vec{y}))(\vec{x},\vec{y}\in V^n)$的充要条件是$T$在$V^n$的标准正交基下的矩阵为反对称矩阵.

证明：设V^n的一个标准正交基为$\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n}$，线性变换$T$在该基下的矩阵为$A=(a_{ij})_{n\times n}$，

即$T(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})=(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_n})A$，于是有：

$$T(\vec{x_i})=a_{1i}\vec{x_1}+a_{2i}\vec{x_2}+\cdots+a_{ni}\vec{x_n},(T(\vec{x_i}),\vec{x_j})=a_{ji}$$
$$T(\vec{x_j})=a_{1j}\vec{x_1}+a_{2j}\vec{x_2}+\cdots+a_{nj}\vec{x_n},(\vec{x_i},T(\vec{x_j}))=a_{ij}$$

必要性$\Rightarrow$设T为反对称变换，则有$(T(\vec{x_i}),\vec{x_j})=-(\vec{x_i},T(\vec{x_j}))$，即$a_{ji}=-a_{ij}$

$(i,j=1,2,\cdots,n)$，故$A^T=-A$.

充分性$\Rightarrow$设T在V^n的标准正交基下的矩阵为反对称矩阵，即$A^T=-A$，则$\forall\vec{x},\vec{y}\in V^n$有：

$$\vec{x}=(\vec{x_1},\vec{x_2},\cdots,\vec{x_n})\begin{pmatrix}\xi_1\\\xi_2\\\cdots\\\xi_n\end{pmatrix},T(\vec{x})=(\vec{x_1},\vec{x_2},\cdots,\vec{x_n})A\begin{pmatrix}\xi_1\\\xi_2\\\cdots\\\xi_n\end{pmatrix}$$

$$\vec{y}=(\vec{x_1},\vec{x_2},\cdots,\vec{x_n})\begin{pmatrix}\eta_1\\\eta_2\\\cdots\\\eta_n\end{pmatrix},T(\vec{y})=(\vec{x_1},\vec{x_2},\cdots,\vec{x_n})A\begin{pmatrix}\eta_1\\\eta_2\\\cdots\\\eta_n\end{pmatrix}$$

$\because\vec{x_1},\vec{x_2},\cdots,\vec{x_n}$是标准正交基

$$\therefore(T(\vec{x}),\vec{y})=(\xi_1,\xi_2,\cdots,\xi_n)A^T\cdot\begin{pmatrix}\eta_1\\\eta_2\\\cdots\\\eta_n\end{pmatrix}=-(\xi_1,\xi_2,\cdots,\xi_n)A\cdot\begin{pmatrix}\eta_1\\\eta_2\\\cdots\\\eta_n\end{pmatrix}=-(\vec{x},T(\vec{y}))$$

$\therefore T$是反对称变换.

3.2 Hermite 正定、半正定矩阵

定义 2-3 设A为n阶 Hermite 矩阵，如果$\forall\vec{x}\in C^n$有$\vec{x}^H A\vec{x}\geq 0$，则称A为非负定（半正定）矩阵，记作$A\geq 0$. 如$\vec{x}^H A\vec{x}>0$，则称A为正定矩阵，记作$A>0$.

定理 3-3 矩阵A为正定（非负定）矩阵$\Leftrightarrow$ A所有特征值为正（非负）数.

推论 若$A>0$，则$t_r(A)>\lambda_i(i=1,2,\cdots,n)$. 若$A\geq 0$，则$t_r(A)>\lambda_i(i=1,2,\cdots,n)$.

λ_i – A的特征值；$t_r(A)=\lambda_1+\lambda_2+\cdots+\lambda_n$称为$A$的迹.

定理 3-4 矩阵A为正定矩阵$\Leftrightarrow$存在n阶矩阵P，使得$A=P^H P$.

推论 1 若 $A>0$，则 A 可逆且 $A^{-1}>0$.

推论 2 若 $A>0$，C 是任一 n 阶非奇异矩阵，则 $C^H AC>0$.

推论 3 若 $A\geq 0$，C 是任一 $n\times m$ 阶矩阵，则 $C^H AC\geq 0$.

定理 3-5 正定矩阵 A 的各阶顺序主子矩阵都是正定矩阵.

定理 3-6 矩阵 A 为正定矩阵 $\Leftrightarrow$ A 的各阶顺序主子式都是正数，即 $|A_k|>0, k=1,2,\cdots,n$.

定理 3-7 设 A,B 都是 n 阶 Hermite 矩阵，且 $B>0$，则存在非奇异矩阵 Q，使得 $Q^T BQ=I$，$Q^H AQ=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$.

3.3 矩阵不等式

定义 3-4 设 A,B 都是 n 阶 Hermite 矩阵，且 $A-B\geq 0$，则称 A 大于或等于 B，记作 $A\geq B$.

★ 矩阵不等式的性质：

（1）$A\geq B, B\geq C\Rightarrow A\geq C$；$A\geq B, k\in R^+ \Rightarrow kA\geq kB$；

$A_1\geq B_1, A_2\geq B_2\Rightarrow A_1+A_2\geq B_1+B_2$；$A\geq B, P$ 为 $n\times m$ 矩阵 $\Rightarrow P^H AP\geq P^H BP$.

（2）若 $A\geq 0$，则 $A\leq t_r(A)I$.

（3）设 A,B 为同阶的正定矩阵 $A\geq B$，则 $B^{-1}\geq A^{-1}$.

（4）若正定矩阵 A,B 可交换：$AB=BA$，且 $A\geq B$，则 $A^2\geq B^2$.

（5）矩阵型的Schwartz 不等式

设 A,B 分别为 $n\times m$ 和 $m\times l$ 阶矩阵，且 AA^H 非奇异 $(n\leq m, rank(A)=n)$，则

$$B^H B\geq (AB)^H(AA^H)^{-1}(AB).$$

3.4 正规矩阵

定义 3-7 设 A 是复数域上的方阵，如果有 $AA^H = A^H A$，则称 A 为**正规矩阵**.

定理 3-10 设 $A \in C^{n\times n}$，则 A 酉相似于对角矩阵 $\Leftrightarrow$ A 为正规矩阵.

本章测试题

一、设 $A=(a_{ij})$ 是一个 n 阶实对称正定矩阵，而 $\vec{\alpha}=(x_1,x_2,\cdots,x_n),\vec{\beta}=(y_1,y_2,\cdots,y_n)$ 在 R^n 中定义实函数为 $(\vec{\alpha},\vec{\beta})=\vec{\alpha}A\vec{\beta}^T$

(1) 证明在这个定义下，R^n 构成一欧氏空间；

(2) 求向量 $\overrightarrow{\varepsilon_1}=(1,0,\cdots,0),\overrightarrow{\varepsilon_2}=(0,1,\cdots,0),\cdots,\overrightarrow{\varepsilon_n}=(0,0,\cdots,1)$ 的度量矩阵；

(3) 具体写出这个空间中的柯西-布涅夫斯基不等式.

解：(1) 设有 $\vec{\gamma}\in R^n,k\in R$，则有

$$(\vec{\alpha},\vec{\beta})=\vec{\alpha}A\vec{\beta}^T=(\vec{\alpha}A\vec{\beta}^T)^T=\vec{\beta}A^T\vec{\alpha}^T=\vec{\beta}A\vec{\alpha}^T=(\vec{\beta},\vec{\alpha})$$

$$(\vec{\alpha},\vec{\beta}+\vec{\gamma})=\vec{\alpha}A(\vec{\beta}+\vec{\gamma})^T=\vec{\alpha}A(\vec{\beta}^T+\vec{\gamma}^T)=\vec{\alpha}A\vec{\beta}^T+\vec{\alpha}A\vec{\gamma}^T=(\vec{\alpha},\vec{\beta})+(\vec{\alpha},\vec{\gamma})$$

$$(k\vec{\alpha},\vec{\beta})=k\vec{\alpha}A\vec{\beta}^T=k(\vec{\alpha}A\vec{\beta}^T)=k(\vec{\alpha},\vec{\beta})$$

当 $\vec{\alpha}=\vec{0}$ 时，$(\vec{\alpha},\vec{\alpha})=0$；当 $\vec{\alpha}\neq\vec{0}$ 时，由 A 对称正定可知 $(\vec{\alpha},\vec{\alpha})=\vec{\alpha}A\vec{\alpha}^T>0$，因此，在这个定义下，$R^n$ 构成一欧氏空间.

(2) 根据已知可知 $(\overrightarrow{\varepsilon_i},\overrightarrow{\varepsilon_j})=\overrightarrow{\varepsilon_i}A\overrightarrow{\varepsilon_j}=a_{ij}(i,j=1,2,\cdots,n)$，所以此向量的度量矩阵即为 $A=(a_{ij})$.

(3) 根据柯西-布涅夫斯基不等式的定义有 $\left|(\vec{\alpha},\vec{\beta})\right|\leq\left|\vec{\alpha}\right|\left|\vec{\beta}\right|$，当且仅当 $\vec{\alpha},\vec{\beta}$ 线性相关时，等号成立．而 $(\vec{\alpha},\vec{\beta})=\vec{\alpha}A\vec{\beta}^T=\sum_{i,j=1}^{n}a_{ij}x_iy_j,\left|\vec{\alpha}\right|=\sqrt{\vec{\alpha}A\vec{\alpha}^T}=\sqrt{\sum_{i,j=1}^{n}a_{ij}x_ix_j}$，

$$\left|\vec{\beta}\right|=\sqrt{\vec{\beta}A\vec{\beta}^T}=\sqrt{\sum_{i,j=1}^{n}a_{ij}y_iy_j}\Rightarrow\left|\sum_{i,j=1}^{n}a_{ij}x_iy_j\right|\leq\sqrt{\sum_{i,j=1}^{n}a_{ij}x_ix_j}\cdot\sqrt{\sum_{i,j=1}^{n}a_{ij}y_iy_j}.$$

二、在 R^4 中，求一单位向量与 $(1,1,-1,1),(1,-1,-1,1)$ 及 $(2,1,1,3)$ 均正交.

解：设所求向量为$\vec{\alpha}=\left(x_1,x_2,x_3,x_4\right)$，则有$\begin{cases}x_1+x_2+x_3+x_4=0\\x_1-x_2-x_3+x_4=0\\2x_1+x_2+x_3+3x_4=0\end{cases}$，求得此方程组的一组解为 $x_1=-4,x_2=0,x_3=-1,x_4=3$，又 $\vec{\alpha}$ 为单位向量，$\therefore \vec{\alpha}=\pm\frac{1}{\sqrt{26}}\left(-4,0,-1,3\right)$.

三、在$P[x]_3$中定义内积为$(f,g)=\int_{-1}^{1}f(x)g(x)dx$，求$P[x]_3$的一个标准正交基.

解：取$P[x]_3$的自然基$1,t,t^2,t^3$，先正交化

$$\overrightarrow{x_1}=1$$

$$\overrightarrow{x_2}=t-\frac{(t,\overrightarrow{x_1})}{(\overrightarrow{x_1},\overrightarrow{x_1})}\overrightarrow{x_1}=t$$

$$\overrightarrow{x_3}=t^2-\frac{(t^2,\overrightarrow{x_1})}{(\overrightarrow{x_1},\overrightarrow{x_1})}\overrightarrow{x_1}-\frac{(t^2,\overrightarrow{x_2})}{(\overrightarrow{x_2},\overrightarrow{x_2})}\overrightarrow{x_2}=t^2-\frac{1}{3}$$

$$\overrightarrow{x_4}=t^3-\frac{(t^3,\overrightarrow{x_1})}{(\overrightarrow{x_1},\overrightarrow{x_1})}\overrightarrow{x_1}-\frac{(t^3,\overrightarrow{x_2})}{(\overrightarrow{x_2},\overrightarrow{x_2})}\overrightarrow{x_2}-\frac{(t^3,\overrightarrow{x_3})}{(\overrightarrow{x_3},\overrightarrow{x_3})}\overrightarrow{x_3}=t^3-\frac{3}{5}t$$

再单位化可得：

$$\overrightarrow{y_1}=\frac{\overrightarrow{x_1}}{\left|\overrightarrow{x_1}\right|}=\frac{1}{\sqrt{2}}$$

$$\overrightarrow{y_2}=\frac{\overrightarrow{x_2}}{\left|\overrightarrow{x_2}\right|}=\frac{\sqrt{6}}{2}t$$

$$\overrightarrow{y_3}=\frac{\overrightarrow{x_3}}{\left|\overrightarrow{x_3}\right|}=\frac{\sqrt{10}}{2}(t^2-\frac{1}{3})$$

$$\overrightarrow{y_4}=\frac{\overrightarrow{x_4}}{\left|\overrightarrow{x_4}\right|}=\frac{5\sqrt{14}}{4}(t^3-\frac{3}{5}t)$$

$\therefore \overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3},\overrightarrow{y_4}$ 即为 $P[x]_3$ 的一个标准正交基.

四、设$\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_m}$是欧式空间$V^m$中的一组向量，而

$$\Delta=\begin{pmatrix}(\overrightarrow{x_1},\overrightarrow{x_1}) & (\overrightarrow{x_1},\overrightarrow{x_2}) & \cdots & (\overrightarrow{x_1},\overrightarrow{x_m})\\(\overrightarrow{x_2},\overrightarrow{x_1}) & (\overrightarrow{x_2},\overrightarrow{x_2}) & \cdots & (\overrightarrow{x_2},\overrightarrow{x_m})\\ \cdots & \cdots & \cdots & \cdots\\(\overrightarrow{x_m},\overrightarrow{x_1}) & (\overrightarrow{x_m},\overrightarrow{x_2}) & \cdots & (\overrightarrow{x_m},\overrightarrow{x_m})\end{pmatrix}$$，试证明 $\det\Delta\neq0$ 的充要条件是 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_m}$ 线性无关.

证明：必要性.

设 $k_1\overrightarrow{x_1}+k_2\overrightarrow{x_2}+\cdots+k_m\overrightarrow{x_m}=\vec{0}$，等式两端与 $\overrightarrow{x_i}(i=1,2,\cdots,m)$ 作内积得，

$k_1\left(\overrightarrow{x_i},\overrightarrow{x_1}\right)+k_2\left(\overrightarrow{x_i},\overrightarrow{x_2}\right)+\cdots+k_m\left(\overrightarrow{x_i},\overrightarrow{x_m}\right)=\vec{0}$，此方程组的系数矩阵即为 Δ，当 $\det\Delta\neq0$ 时，方程组只有零解，故 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_m}$ 线性无关.

充分性.

若 $\overrightarrow{k_1},\overrightarrow{k_2},\cdots,\overrightarrow{k_m}$ 是方程组 $\begin{cases}k_1(\overrightarrow{x_1},\overrightarrow{x_1}) & +k_2(\overrightarrow{x_1},\overrightarrow{x_2}) & +\cdots & +k_m(\overrightarrow{x_1},\overrightarrow{x_m})=0\\k_1(\overrightarrow{x_2},\overrightarrow{x_1}) & +k_2(\overrightarrow{x_2},\overrightarrow{x_2}) & +\cdots & +k_m(\overrightarrow{x_2},\overrightarrow{x_m})=0\\ \cdots & \cdots & \cdots & \cdots\\k_1(\overrightarrow{x_m},\overrightarrow{x_1}) & +k_2(\overrightarrow{x_m},\overrightarrow{x_2}) & +\cdots & +k_m(\overrightarrow{x_m},\overrightarrow{x_m})=0\end{cases}$ 的解，则

$\left(\overrightarrow{x_i},k_1\overrightarrow{x_1}+k_2\overrightarrow{x_2}+\cdots+k_m\overrightarrow{x_m}\right)=k_1(\overrightarrow{x_i},\overrightarrow{x_1})+k_2(\overrightarrow{x_i},\overrightarrow{x_2})+\cdots+k_m(\overrightarrow{x_1},\overrightarrow{x_m})=0$，从而得

$$\begin{aligned}&\left(k_1\overrightarrow{x_1}+k_2\overrightarrow{x_2}+\cdots+k_m\overrightarrow{x_m},k_1\overrightarrow{x_1}+k_2\overrightarrow{x_2}+\cdots+k_m\overrightarrow{x_m}\right)\\&=k_1\left(\overrightarrow{x_1},k_1\overrightarrow{x_1}+k_2\overrightarrow{x_2}+\cdots+k_m\overrightarrow{x_m}\right)+\cdots+k_m\left(\overrightarrow{x_m},k_1\overrightarrow{x_1}+k_2\overrightarrow{x_2}+\cdots+k_m\overrightarrow{x_m}\right)=0\end{aligned}$$

所以 $k_1\overrightarrow{x_1}+k_2\overrightarrow{x_2}+\cdots+k_m\overrightarrow{x_m}=\vec{0}$，由于 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_m}$ 线性无关，则 $k_1=k_2=\cdots=k_m=0$，方程组只有零解，所以 $\det\Delta\neq0$.

五、求齐次线性方程组 $\begin{cases}2x_1+x_2-x_3+x_4-3x_5=0\\x_1+x_2-x_3+x_5=0\end{cases}$ 的解空间的一组标准正交基.

六、设$\overrightarrow{e_1},\overrightarrow{e_2},\cdots,\overrightarrow{e_n}$为$n$维欧式空间$V$的一组基底，证明任二向量$\overrightarrow{\mu}=(\overrightarrow{e_1},\overrightarrow{e_2},\cdots,\overrightarrow{e_n})\begin{pmatrix}x_1\\ \cdots\\ x_n\end{pmatrix}$

$\vec{v}=(\overrightarrow{e_1},\overrightarrow{e_2},\cdots,\overrightarrow{e_n})\begin{pmatrix}y_1\\ \cdots\\ y_n\end{pmatrix}$的内积由等式$(\overrightarrow{\mu},\vec{v})=x_1y_1+x_2y_2+\cdots+x_ny_n$表达的充要条件是$\overrightarrow{e_1},\overrightarrow{e_2},\cdots,\overrightarrow{e_n}$为标准正交基.

七、设$\overrightarrow{\varepsilon_1},\overrightarrow{\varepsilon_2},\overrightarrow{\varepsilon_3}$是三维欧式空间中的一组标准正交基，证明：
$$\begin{aligned}\overrightarrow{\alpha_1}&=\frac{1}{3}(2\overrightarrow{\varepsilon_1}+2\overrightarrow{\varepsilon_2}-\overrightarrow{\varepsilon_3})\\ \overrightarrow{\alpha_2}&=\frac{1}{3}(2\overrightarrow{\varepsilon_1}-\overrightarrow{\varepsilon_2}+2\overrightarrow{\varepsilon_3})\\ \overrightarrow{\alpha_3}&=\frac{1}{3}(\overrightarrow{\varepsilon_1}-2\overrightarrow{\varepsilon_2}-2\overrightarrow{\varepsilon_3})\end{aligned}$$
也是一组标准正交基.

解：已知可得$\left(\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\alpha_3}\right)=\left(\overrightarrow{\varepsilon_1},\overrightarrow{\varepsilon_2},\overrightarrow{\varepsilon_3}\right)\begin{pmatrix}\frac{2}{3}&\frac{2}{3}&\frac{1}{3}\\ \frac{2}{3}&-\frac{1}{3}&-\frac{2}{3}\\ -\frac{1}{3}&\frac{2}{3}&-\frac{2}{3}\end{pmatrix}$，$A=\begin{pmatrix}\frac{2}{3}&\frac{2}{3}&\frac{1}{3}\\ \frac{2}{3}&-\frac{1}{3}&-\frac{2}{3}\\ -\frac{1}{3}&\frac{2}{3}&-\frac{2}{3}\end{pmatrix}$ **而**

$$A^TA=\begin{pmatrix}\frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3}\\ \frac{1}{3} & -\frac{2}{3} & -\frac{2}{3}\end{pmatrix}\begin{pmatrix}\frac{2}{3} & \frac{2}{3} & \frac{1}{3}\\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3}\\ -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3}\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}=I_3$$，即 A 为正交矩阵，而

$\overrightarrow{\varepsilon_1},\overrightarrow{\varepsilon_2},\overrightarrow{\varepsilon_3}$ 是三维欧式空间中的一组标准正交基，所以 $\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\alpha_3}$ 也是一组标准正交基.

八、在 R^n 中，$d\left(\vec{\alpha},\vec{\beta}\right)=\left|\vec{\alpha}-\vec{\beta}\right|$ 通常称为 $\vec{\alpha}$ 与 $\vec{\beta}$ 的距离，证明：$d\left(\vec{\alpha},\vec{\gamma}\right)\le d\left(\vec{\alpha},\vec{\beta}\right)+d\left(\vec{\beta},\vec{\gamma}\right)$.

九、设 $\vec{y}$ 是欧氏空间 V 中的单位向量，$\vec{x}\in V$，定义线性变换 $T:T(\vec{x})=\vec{x}-2(\vec{y},\vec{x})\vec{y}$，证明 T 是正交变换.

十、求两个正交矩阵，以 $(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})$ 及 $(\frac{1}{6},\frac{1}{6},\frac{1}{2},-\frac{5}{6})$ 为前两行.

十一、证明上三角的正交矩阵必为对角矩阵，且对角线上的元素为$+1$或-1.

十二、证明：欧氏空间中不同基的度量矩阵是合同的；任一欧氏空间都存在标准正交基.

十三、设A和B都是正交矩阵，$|A|=-|B|$，求证$|A+B|=0$.

十四、若 $A=(a_{ij})$ 为正交矩阵，则方程组 $\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases}$ 的解为：

$$\begin{cases} x_1=a_{11}b_1+a_{21}b_2+\cdots+a_{n1}b_n \\ x_2=a_{12}b_1+a_{22}b_2+\cdots+a_{n2}b_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n=a_{1n}b_1+a_{2n}b_2+\cdots+a_{nn}b_n \end{cases}.$$

十五、设 $A=(a_{ij})$ 是行列式为 1 的三阶正交方正，证明：$\begin{cases} a_{11}=a_{22}a_{33}-a_{23}a_{32} \\ a_{21}=a_{13}a_{32}-a_{12}a_{33} \\ a_{31}=a_{12}a_{23}-a_{13}a_{22} \end{cases}$.

十六、求将 $(1,0,0),(0,1,0)$ 变为 $(\frac{2}{3},\frac{2}{3},-\frac{1}{3}),(\frac{2}{3},-\frac{1}{3},\frac{2}{3})$ 之正交变换.

十七、已知单位向量$\vec{\mu}=(\frac{1}{\sqrt{3}},-\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}})^T$，做一个初等反射矩阵$H$，并验证有$H^{-1}=H$.

十八、试用初等旋转变换化向量$\vec{x}=(2,3,0,5)^T$与$\vec{e_1}=(1,0,0,0)^T$同方向.

十九、设$\vec{x}=\frac{1}{\sqrt{2}}(1,1,0)^T$，求一正交矩阵$A$，使$A\vec{x}=\vec{e_1}(\vec{e_1}=(1,0,0)^T)$.

第三章　矩阵的 Jordan 标准型

§1　约当（Jordan）标准型

1.1　$\lambda-$矩阵的概念

定义 1-1　设 $a_{ij}(\lambda)(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 为数域 P 上的多项式，以 $a_{ij}(\lambda)$ 为元素的

$m\times n$ 矩阵 $A(\lambda)=\begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{12}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \cdots & a_{2n}(\lambda) \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \cdots & a_{mn}(\lambda) \end{pmatrix}$ 称为 $\lambda-$**矩阵**或**多项式矩阵**.

◆ 显然，数字矩阵和特征矩阵 $\lambda I-A$ 均为 $\lambda-$矩阵的特例.

定义 1-2　$\lambda-$矩阵 $A(\lambda)$ 不恒等于零的子式的最大阶数叫做 $A(\lambda)$ 的**秩**，即若 $A(\lambda)$ 的秩为 r，则至少有一个 r 阶子式不恒等于零，而 $r+1$ 阶及以上各阶的子式均恒等于零.

定义 1-3　下列各种类型的变换，叫做 $\lambda-$矩阵的**初等变换**：

1° 对调两行（列）；2° 用 $(k\neq 0)$ 乘任一行（列）；3° 第 i 行（列）的 $\varphi(\lambda)$ 倍加到第 j 行（列）上，其中 $\varphi(\lambda)$ 是 λ 的多项式.

★ 对单位矩阵 I 施行上述三种变换便得到相应的三种 $\lambda-$矩阵的**初等矩阵**.

$$I(i,j)=\begin{pmatrix} 1 &&&&&&&& \\ & \cdots &&&&&&& \\ && 0 && \cdots && 1 && \\ &&& 1 &&&&& \\ && \cdots && \cdots && \cdots && \\ &&&&& 1 &&& \\ && 1 && \cdots && 0 && \\ &&&&&&& \cdots & \\ &&&&&&&& 1 \end{pmatrix}\begin{matrix} \\ \\ \leftarrow R_i \\ \\ \\ \\ \leftarrow R_j \\ \\ \\ \end{matrix},\quad I(i(k))=\begin{pmatrix} 1 &&&& \\ & \cdots &&& \\ && k && \\ &&& \cdots & \\ &&&& 1 \end{pmatrix}\begin{matrix} \\ \\ \leftarrow R_i \\ \\ \\ \end{matrix},$$

$$I(i.j(\varphi))=\begin{pmatrix}1 & & & & & & \\ & \cdots & & & & & \\ & & 1 & \cdots & \varphi(\lambda) & & \\ & & & \cdots & \cdots & & \\ & & & & 1 & & \\ & & & & & \cdots & \\ & & & & & & 1\end{pmatrix}\begin{array}{l} \\ \\ \leftarrow R_i \\ \\ \leftarrow R_j \\ \\ \\ \end{array}.$$

◆ 初等矩阵均可逆，且 $I(i,j)^{-1}=I(i,j),I(i(k))^{-1}=I(i(\frac{1}{k})),I(i,j(\varphi))^{-1}=I(i,j(-\varphi))$.

◆ 对一个 m 阶 $\lambda-$矩阵 $A(\lambda)$ 的行（列）做初等变换，相当于用相应的 m 阶初等矩阵左（右）乘 $A(\lambda)$.

定义 1-4 如果 $A(\lambda)$ 经过有限次初等变换后变成 $B(\lambda)$，则称 $A(\lambda)$ 与 $B(\lambda)$ **等价**，记之为 $A(\lambda)\simeq B(\lambda)$. $\lambda-$矩阵的等价关系满足 1° 自反性：每一 $\lambda-$矩阵与自己等价；2° 对称性：若 $A(\lambda)\simeq B(\lambda)$，则 $B(\lambda)\simeq A(\lambda)$；$3^\circ$ 传递性：若 $A(\lambda)\simeq B(\lambda)$，$B(\lambda)\simeq C(\lambda)$，则 $A(\lambda)\simeq C(\lambda)$.

定理 1-1 $A(\lambda)\simeq B(\lambda)\Leftrightarrow$
$\exists P(\lambda),Q(\lambda)(|P(\lambda)|\neq 0,|Q(\lambda)|\neq 0),sub.B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$

1.2 $\lambda-$矩阵的标准形

★ 任何一个 $\lambda-$矩阵经过有限次初等变换，总可以化为对角形，即标准形.

引理 1-1 设 $\lambda-$矩阵的左上角元素 $a_{11}(\lambda)\neq 0$，$A(\lambda)$ 中至少有一个元素不能被它整除，那么一定可以找到与 $A(\lambda)$ 等级的矩阵 $B(\lambda)$，它的左上角元素也不为零，但次数比 $a_{11}(\lambda)$ 低.

定理 1-2 n 阶 $\lambda-$矩阵 $A(\lambda)$ 经过有限次初等变换后，总可以化为如下对角形，称为 $A(\lambda)$ 的**标准形**.

$$A(\lambda)\simeq\begin{pmatrix} d_1(\lambda) & & & & & & \\ & d_2(\lambda) & & & & & \\ & & \cdots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \cdots & \\ & & & & & & 0 \end{pmatrix}$$

，其中 $r\le n, d_i(\lambda)$ 是首项系数为 1 的多项式，且 $d_{i-1}(\lambda)$ 能整除 $d_i(\lambda)$（前面的 n 个 $d_i(\lambda)$ 可能是 1）.

1.3 不变因子和初等因子

定义 1-5 $\lambda-$矩阵 $A(\lambda)$ 最后化成的标准形

$$A(\lambda)\simeq\begin{pmatrix} d_1(\lambda) & & & & & & \\ & d_2(\lambda) & & & & & \\ & & \cdots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \cdots & \\ & & & & & & 0 \end{pmatrix}$$

，其对角线的元素 $d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$ 称为 $A(\lambda)$ 的**不变因子**.

定义 1-6 设 $A(\lambda)$ 的秩为 r，对于 $k\in Z^+(1\le k\le r)$，$A(\lambda)$ 中必有非零的 k 阶子式，且不是唯一的，我们把 $A(\lambda)$ 中全部 k 阶子式的最高公因式称为 $A(\lambda)$ 的 k **阶行列式因子**，记为 $D_k(\lambda)$，且有 $D_k(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda)$，不变因子和行列式因子有如下关系：

$d_1(\lambda)=D_1(\lambda),d_2(\lambda)=\dfrac{D_2(\lambda)}{D_1(\lambda)},\cdots,d_r(\lambda)=\dfrac{D_r(\lambda)}{D_{r-1}(\lambda)}$，也就是说 $A(\lambda)$ 的不变因子由行列式因子唯一确定.

定理 1-3 $A(\lambda)\simeq B(\lambda)\Leftrightarrow A(\lambda)$ 与 $B(\lambda)$ 有相同的不变因子 $d_i(\lambda)(i=1,2,\cdots,r)$.

由于每个不变因子$d_i(\lambda)$是首项系数为 1 的多项式，$\therefore$当$A(\lambda)$的秩为r时，r个不恒等于零的$d_i(\lambda)$在复数域内总可分解为一次因式的乘积，即：

$$\begin{cases} d_1(\lambda)=(\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdots(\lambda-\lambda_S)^{e_{1s}} \\ d_2(\lambda)=(\lambda-\lambda_1)^{e_{21}}(\lambda-\lambda_2)^{e_{22}}\cdots(\lambda-\lambda_S)^{e_{2s}} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ d_r(\lambda)=(\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}}\cdots(\lambda-\lambda_S)^{e_{rs}} \end{cases}$$

其中$\lambda_1,\lambda_2,\cdots,\lambda_S$是$d_r(\lambda)$中一切相异的根，且有些可能为复数，另外，由于$d_i(\lambda)$可以被$d_{i-1}(\lambda)$整除，所以有关系式$0\le e_{1j}\le e_{2j}\le\cdots\le e_{rj}(j=1,2,\cdots,s)$，也就是说，$e_{ij}\ (1\le i<r,1\le j\le s)$可能为零，当$e_{r1},e_{r2},\cdots,e_{rs}$无一为零.

定义 1-7　当$e_{ij}\ne 0$，因子$(\lambda-\lambda_j)^{e_{ij}}(i=1,2,\cdots,r;j=1,2,\cdots,s)$的全体叫做$A(\lambda)$的**初等因子**.

★　n阶约当块$J_i=\begin{pmatrix}\lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \cdots & \cdots & \\ & & & \cdots & 1 \\ & & & & \lambda_i\end{pmatrix}_{n\times n}$的特征矩阵$\lambda I-J_i$的不变因子为$d_1(\lambda)=d_2(\lambda)=\cdots=d_{n-1}(\lambda)=1,d_n(\lambda)=(\lambda-\lambda_i)^n$，初等因子只有一个$(\lambda-\lambda_i)^n$.

定理 1-4　设$\lambda-$矩阵为准对角形$A(\lambda)=\begin{pmatrix}A_1(\lambda) & & & \\ & A_2(\lambda) & & \\ & & \cdots & \\ & & & A_m(\lambda)\end{pmatrix}$，则$A_i(\lambda)$ $(i=1,2,\cdots,m)$的所有初等因子的集合是$A(\lambda)$的初等因子.

定理 1-5　$A(\lambda)\simeq B(\lambda)\Leftrightarrow A(\lambda)$与$B(\lambda)$的秩相等，且初等因子相同.

1.4 利用初等因子化为约当标准形

定理 1-6 矩阵 $A\sim B \Leftrightarrow$ 它们相应的特征矩阵 $\lambda I-A\simeq \lambda I-B$.

定理 1-7 设 A 的特征矩阵 $\lambda I-A$ 的初等因子为 $(\lambda-\lambda_1)^{n_1},(\lambda-\lambda_2)^{n_2},\cdots,(\lambda-\lambda_m)^{n_m}$ 且

$\sum\limits_{i=1}^{m} n_i=n$，则 $A\sim J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \cdots & \\ & & & J_m \end{pmatrix}$，其中 $J_i=\begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \cdots & \cdots & \\ & & & \cdots & 1 \\ & & & & \lambda_i \end{pmatrix}_{n_i\times n_i}$，式中

$i=1,2,\cdots,m$， J 就是 A 的约当标准形.

★ 利用初等因子将 A 化为约当标准形的步骤：

(1) 求 n 阶方阵 $A=(a_{ij})$ 的特征矩阵 $\lambda I-A$ 的初等因子(行列式因子 $D_i(\lambda)\Rightarrow$ 不变因子 $d_i(\lambda)=\dfrac{D_i(\lambda)}{D_{i-1}(\lambda)}\Rightarrow$ 初等因子)，设为 $(\lambda-\lambda_1)^{n_1},(\lambda-\lambda_2)^{n_2},\cdots,(\lambda-\lambda_m)^{n_m}$，其中 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 可能相同，指数 $n_1,n_2,\cdots,n_m$ 也可能相同，且 $\sum\limits_{i=1}^{m} n_i=n$；

(2) 写出每个初等因子 $(\lambda-\lambda_i)^{n_i}(i=1,2,\cdots,m)$ 对应的 *Jordan* 块

$$J_i=\begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \cdots & \cdots & \\ & & & \cdots & 1 \\ & & & & \lambda_i \end{pmatrix}_{n_i\times n_i},(i=1,2,\cdots,m);$$

(3) 写出这些 *Jordan* 块构成的 *Jordan* 标准形 $J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \cdots & \\ & & & J_m \end{pmatrix}$.

例： 设线性空间 $V^3\in C$ 的一个基为 $\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3}$，线性变换 T 在该基下的矩阵为

$A=\begin{pmatrix}-1 & 1 & 1\\ -5 & 21 & 17\\ 6 & -26 & -21\end{pmatrix}$，求$V^3$的另一组基$\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3}$（用$\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3}$表示），使$T$在该基下的矩阵为$Jordan$标准形.

解：先求A的$Jordan$标准形及相似变换矩阵P.

A的特征矩阵为$\lambda I-A=\begin{pmatrix}\lambda+1 & -1 & -1\\ 5 & \lambda-21 & -17\\ -6 & 26 & \lambda+21\end{pmatrix}$，易知$D_1(\lambda)=1$，而$\lambda I-A$的其中 2 个二阶子式$\begin{vmatrix}-1 & -1\\ \lambda-21 & -17\end{vmatrix}=\lambda-4$，$\begin{vmatrix}5 & \lambda-21\\ -6 & 26\end{vmatrix}=6\lambda+4$互质，故而$D_2(\lambda)=1$.

$$\lambda I-A=\begin{pmatrix}\lambda+1 & -1 & -1\\ 5 & \lambda-21 & -17\\ -6 & 26 & \lambda+21\end{pmatrix}\simeq\begin{pmatrix}-1 & -1 & \lambda+1\\ -17 & \lambda-21 & 5\\ \lambda+21 & 26 & -6\end{pmatrix}\simeq\begin{pmatrix}1 & 1 & -\lambda-1\\ 0 & \lambda-4 & -17\lambda-12\\ 0 & -\lambda+5 & \lambda^2+22\lambda+15\end{pmatrix}$$

$$\simeq\begin{pmatrix}1 & 1 & -\lambda-1\\ 0 & \lambda-4 & -17\lambda-12\\ 0 & 1 & \lambda^2+5\lambda+3\end{pmatrix}\simeq\begin{pmatrix}1 & 1 & -\lambda-1\\ 0 & 1 & \lambda^2+5\lambda+3\\ 0 & \lambda-4 & -17\lambda-12\end{pmatrix}\simeq\begin{pmatrix}1 & 1 & -\lambda-1\\ 0 & 1 & \lambda^2+5\lambda+3\\ 0 & 0 & \lambda^2(\lambda+1)\end{pmatrix}$$

$\therefore D_3(\lambda)=\lambda^2(\lambda+1)$，于是，不变因子为$d_1(\lambda)=\dfrac{D_1(\lambda)}{1}=1, d_2(\lambda)=\dfrac{D_2(\lambda)}{D_1(\lambda)}=1$，

$d_3(\lambda)=\dfrac{D_3(\lambda)}{D_2(\lambda)}=\lambda^2(\lambda+1)$

$\therefore A$的初等因子为$\lambda^2,\lambda+1$，从而A的$Jordan$标准形为$J=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{pmatrix}$.

$(0\cdot I-A)\vec{x}=\vec{0}$的一个非零解为$\overrightarrow{x_1^{(1)}}=(1,-3,4)^T$；

$(0\cdot I-A)\vec{x}=-\overrightarrow{x_1^{(1)}}$的一个非零解为$\overrightarrow{x_2^{(1)}}=(-1,-2,2)^T$；

$(-1\cdot I-A)\vec{x}=\vec{0}$的一个非零解为$\overrightarrow{x_1^{(2)}}=(1,1,-1)^T$.

于是可得$P=(\overrightarrow{x_1^{(1)}},\overrightarrow{x_2^{(1)}},\overrightarrow{x_1^{(2)}})=\begin{pmatrix}1 & -1 & 1\\ -3 & -2 & 1\\ 4 & 2 & -1\end{pmatrix}$，且有$P^{-1}AP=J$.

由$(\overrightarrow{y_1},\overrightarrow{y_2},\overrightarrow{y_3})=(\overrightarrow{x_1},\overrightarrow{x_2},\overrightarrow{x_3})P$求得$V^3$的另一组基为$\begin{matrix}\overrightarrow{y_1}=\overrightarrow{x_1}-3\overrightarrow{x_2}+4\overrightarrow{x_3}\\ \overrightarrow{y_2}=-\overrightarrow{x_1}-2\overrightarrow{x_2}+2\overrightarrow{x_3}\\ \overrightarrow{y_3}=\overrightarrow{x_1}+\overrightarrow{x_2}-\overrightarrow{x_3}\end{matrix}$，$T$在该基下的矩阵为$J$.

§2　约当标准形的应用

例： 解线性微分方程组$\begin{cases}\dfrac{dy_1}{dx}=-y_1+y_2\\ \dfrac{dy_2}{dx}=-4y_1+3y_2\\ \dfrac{dy_3}{dx}=y_1+2y_3\end{cases}$.

解：方程组右边的系数方阵为$A=\begin{pmatrix}-1&1&0\\-4&3&0\\1&0&2\end{pmatrix}$，令$\vec{Y}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$，则原方程组改写成$\dfrac{d\vec{Y}}{dx}=A\vec{Y}$. 下面把矩阵$A$化为$Jordan$标准形.

$$\lambda I-A=\begin{pmatrix}\lambda+1&-1&0\\4&\lambda-3&0\\-1&0&\lambda-2\end{pmatrix}\simeq\begin{pmatrix}1&0&0\\\lambda-3&(\lambda-1)^2&0\\0&-1&\lambda-2\end{pmatrix}\simeq\begin{pmatrix}1&0&0\\0&(\lambda-1)^2&0\\0&-1&\lambda-2\end{pmatrix}$$

$$\simeq\begin{pmatrix}1&0&0\\0&0&(\lambda-2)(\lambda-1)^2\\0&-1&\lambda-2\end{pmatrix}\simeq\begin{pmatrix}1&0&0\\0&1&2-\lambda\\0&0&(\lambda-2)(\lambda-1)^2\end{pmatrix}$$

$\therefore d_1(\lambda)=d_2(\lambda)=1,d_3(\lambda)=(\lambda-2)(\lambda-1)^2$，初等因子为$\lambda-2,(\lambda-1)^2$，从而$A$的$Jordan$标准形为$J=\begin{pmatrix}2&0&0\\0&1&1\\0&0&1\end{pmatrix}$.

$(2\cdot I-A)\overrightarrow{X_1}=\vec{0}$的一个非零解为$\overrightarrow{X_1}=(0,0,1)^T$；

$(I-A)\overrightarrow{X_2}=\vec{0}$的一个非零解为$\overrightarrow{X_2}=(1,2,-1)^T$；

$(I-A)\overrightarrow{X_3}=-\overrightarrow{X_2}$的一个非零解为$\overrightarrow{X_3}=(0,1,-1)^T$.

$\therefore C=\begin{pmatrix}0 & 1 & 0\\ 0 & 2 & 1\\ 1 & -1 & -1\end{pmatrix}$，使得 $C^{-1}AC=J$.

令 $\vec{Y}=C\vec{Z},\vec{Z}=\begin{pmatrix}z_1\\ z_2\\ z_3\end{pmatrix}$，则由 $\frac{d\vec{Y}}{dx}=A\vec{Y}$ 得

$\frac{dC\vec{Z}}{dx}=AC\vec{Z}\Rightarrow C^{-1}\frac{dC\vec{Z}}{dx}=C^{-1}AC\vec{Z}\Rightarrow\frac{d\vec{Z}}{dx}=J\vec{Z}$，即 $\frac{d\vec{Z}}{dx}=\begin{pmatrix}2 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}\vec{Z}$，也即

$\begin{cases}\frac{dz_1}{dx}=2z_1\\ \frac{dz_2}{dx}=z_2+z_3\\ \frac{dz_3}{dx}=z_3\end{cases}$，显然有 $\begin{cases}z_1=k_1e^{2x}\\ z_3=k_3e^{x}\end{cases}$ 再解第二个方程 $z_2=\int k_3e^xdx+k_2e^x=e^x(k_3x+k_2)$

再由 $\vec{Y}=C\vec{Z}$ 得 $\begin{cases}y_1=z_2=e^x(k_3x+k_2)\\ y_2=2z_2+z_3=2e^x(k_3x+k_2)+k_3e^x\\ y_3=z_1-z_2-z_3=k_1e^{2x}-e^x(k_3x+k_2+k_3)\end{cases}$，$k_i(i=1,2,3)$ 为任意常数.

本章测试题

一、已知 $A=\begin{pmatrix}1 & & & \\ 1 & 1 & & \\ 2 & 3 & 2 & \\ 2 & 3 & 1 & 2\end{pmatrix}$，求可逆矩阵 P，使得 $P^{-1}AP=J$.

二、设 n 阶方阵 A 有特征值 $2,4,6,\cdots,2n$. (1) $A-I$ 可否对角化？(2) 求 $|A-I|$.

解：(1) 设 λ 是方阵 A 的特征值，则 $\lambda-1$ 是 $A-I$ 的特征值，由此可知 $A-I$ 的特征值有 $1,3,5,\cdots,2n-1$，所以 $A-I$ 可对角化.

(2) 由(1) 可得 $|A-I|=1\cdot3\cdot5\cdots(2n-1)$.

三、设 A 是 n 阶是对称矩阵，且 $A^2=A$，证明：存在 n 阶正交矩阵 Q，使得

$$Q^{-1}AQ=\begin{pmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0\end{pmatrix}_{n\times n}.$$

证明：设 λ 是方阵 A 的特征值，对应的特征向量为 $\overrightarrow{X}$，则有 $A\overrightarrow{X}=\lambda\overrightarrow{X}$. 由此可得

$A(\lambda\overrightarrow{X})=A(A\overrightarrow{X})=A^2\overrightarrow{X}$，$\lambda^2\overrightarrow{X}=A^2\overrightarrow{X}=A\overrightarrow{X}=\lambda\overrightarrow{X}\Rightarrow(\lambda^2-\lambda)\overrightarrow{X}=\vec{0}$，$\because\overrightarrow{X}\neq\vec{0}$，

$\therefore\lambda^2-\lambda=0\Rightarrow\lambda=0$ 或 $\lambda=1$.

设 $\eta_1,\cdots,\eta_s$ 是 A 对应特征值 0 的极大标准正交特征向量组，$\xi_1,\cdots,\xi_t$ 是 A 对应特征值 1 的极大标准正交特征向量组.

因 A 是实对称矩阵，可对角化，所以 A 应有 n 个标准正交的特征向量。于是 $s+t=n$.

令 $Q=[\eta_1,\cdots,\eta_s,\xi_1,\cdots,\xi_t]$，则 Q 是正交矩阵且 $Q^{-1}AQ=\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix}_{n\times n}$.

四、求解线性微分方程组 $\begin{cases}\xi_1'(t)=-\xi_1+\xi_2+\xi_3\\\xi_2'(t)=-5\xi_1+21\xi_2+17\xi_3\\\xi_3'(t)=6\xi_1-26\xi_2-21\xi_3\end{cases}$.

第四章　矩阵分解

§1　矩阵的三角分解

1.1　消元过程的矩阵描述

设 n 元线性方程组为 $\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases}$ 或 $A\vec{x}=\vec{b}$，其中 $A=(a_{ij})_{n\times n}$，$\vec{x}=(x_1,x_2,\cdots,x_n)^T,\vec{b}=(b_1,b_2,\cdots,b_n)^T$．

Gauss 消去法的基本思想是：利用矩阵的初等行变换化系数矩阵 A 为上三角矩阵.

记 n 阶矩阵 A 的前 $n-1$ 个顺序主子式为

$$\Delta_1=a_{11},\Delta_2=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\cdots,\Delta_{n-1}=\begin{vmatrix} a_{11} & \cdots & a_{1,n-1} \\ \cdots & & \cdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{vmatrix}$$

定理 1-1　当 $\Delta_1,\Delta_2,\cdots,\Delta_{n-1}$ 都不为零时，则 $a_{kk}^{(k)}\neq 0(k=1,2,\cdots,n-1)$．

定理 1-2　Gauss 消元过程能够进行到底 $\Leftrightarrow$ A 的前 $n-1$ 个顺序主子式都不为零，即 $\Delta_k\neq 0(k=1,2,\cdots,n-1)$．

1.2　矩阵的三角分解

$$A=A^{(1)}=\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix} \xrightarrow{a_{11}^{(1)}\neq 0,r_i-r_1\times\frac{a_{i1}^{(1)}}{a_{11}^{(1)}}(i=2,3,\cdots,n)}$$

记

$$A^{(2)}=\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix}(l_{i1}=\frac{a_{i1}^{(1)}}{a_{11}^{(1)}},a_{ij}^{(2)}=a_{ij}^{(1)}-l_{i1}a_{1j}^{(1)})$$

相当于 $A^{(2)}=I(1,n(-l_{n1}))\cdots\cdots I(1,2,(-l_{21}))A^{(1)}\xrightarrow{L^{(1)}=I(1,n(-l_{n1}))\cdots\cdots I(1,2,(-l_{21}))}$
$A^{(2)}=L^{(1)}A^{(1)}$

…………如此下去，可得 $L^{(n-1)}\cdots L^{(2)}L^{(1)}A^{(1)}=A^{(n)}$

令 $U=A^{(n)}$，则有 $L^{(n-1)}\cdots L^{(2)}L^{(1)}A=U$，$\because \det L^{(k)}=1\neq 0(k=1,2,\cdots,n-1)$，

$\therefore A=(L^{(1)})^{-1}(L^{(2)})^{-1}\cdots(L^{(n-1)})^{-1}U$，令

$$L=(L^{(1)})^{-1}(L^{(2)})^{-1}\cdots(L^{(n-1)})^{-1}=\begin{pmatrix}1 & & & & & \\ l_{21} & 1 & & & & \\ l_{31} & l_{32} & 1 & & & \\ \cdots & \cdots & \cdots & & & \\ \cdots & \cdots & \cdots & \cdots & 1 & \\ l_{n1} & l_{n2} & \cdots & \cdots & l_{n,n-1} & 1\end{pmatrix}$$，则 $A=LU$.

定理 1-3 （*LDU* **基本定理**）设 A 为 n 阶方阵，则 A 可以唯一地分解为 $A=LDU$ 的充要条件是 A 的前 $n-1$ 个顺序主子式 $\Delta_k\neq 0(k=1,2,\cdots,n-1)$. 其中 L,U 分别是单位下、上三角矩阵，D 是对角矩阵 $D=diag(d_1,d_2,\cdots,d_n),d_k=\dfrac{\Delta_k}{\Delta_{k-1}},k=1,2,\cdots,n,\Delta_0=1$.

推论 1 设 A 为 n 阶方阵，则 A 可以唯一地进行 Doolittle 分解 $\Leftrightarrow$ A 的前 $n-1$ 个顺序主子式 $\Delta_k=\begin{vmatrix}a_{11} & \cdots & a_{ik}\\ \cdots & \cdots & \cdots\\ a_{k1} & \cdots & a_{kk}\end{vmatrix}\neq 0(k=1,2,\cdots,n-1)$，其中 L 为单位下三角矩阵，$\widetilde{U}$ 是上三角矩阵，即有

$$A=\begin{pmatrix}1 & & & & \\ l_{21} & 1 & & & \\ \cdots & \cdots & \cdots & & \\ \cdots & \cdots & \cdots & 1 & \\ l_{n1} & \cdots & \cdots & l_{n,n-1} & 1\end{pmatrix}\begin{pmatrix}u_{11} & u_{12} & \cdots & u_{1n}\\ & u_{22} & \cdots & u_{2n}\\ & & \cdots & \cdots\\ & & & u_{nn}\end{pmatrix}$$. 并且若 A 为奇异阵，则 $u_{nn}=0$；

若 A 为非奇异阵，则充要条件可换为 $\Delta_k\neq 0(k=1,2,\cdots,n)$.

推论 2 n 阶方阵 A 可唯一地进行 Crout 分解

$$A=\tilde{L}U=\begin{pmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & & \\ \cdots & \cdots & \cdots & & \\ \cdots & \cdots & \cdots & l_{n-1,n-1} & \\ l_{n1} & l_{n2} & \cdots & l_{n,n-1} & l_{nn} \end{pmatrix}\begin{pmatrix} 1 & u_{12} & \cdots & u_{1n} \\ & 1 & \cdots & u_{2n} \\ & & \cdots & \cdots \\ & & & 1 \end{pmatrix} \Leftrightarrow$$ A 的前 $n-1$ 个顺序主子式

$\Delta_k=\begin{vmatrix} a_{11} & \cdots & a_{ik} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}\neq 0(k=1,2,\cdots,n-1)$．并且若 A 为奇异阵，则 $l_{nn}=0$；若 A 为非奇异阵，则充要条件可换为 $\Delta_k\neq 0(k=1,2,\cdots,n)$．

1.3　常用的三角分解公式

一、Crout 分解

Crout 分解公式：$\begin{matrix} u_{ii}=1(i=1,2,\cdots,n) \\ l_{i1}=a_{i1}(i=1,2,\cdots,n) \\ u_{1j}=\dfrac{a_{1j}}{l_{11}}(j=2,3,\cdots,n) \end{matrix}$，对于 $k=2,3,\cdots,n$ 依次计算

$$\begin{cases} l_{ik}=a_{ik}-\sum\limits_{s=1}^{k-1}l_{is}u_{sk}(i=k,k+1,\cdots,n) \\ u_{kj}=\dfrac{1}{l_{kk}}(a_{kj}-\sum\limits_{t=1}^{k-1}l_{kt}u_{tj})(j=k,k+1,\cdots,n) \end{cases}$$

矩阵 A 的 LU 分解计算顺序：

第 1 步：$l_{11},u_{12},u_{13},\cdots,u_{1n}$

第 2 步：$l_{21},l_{22},u_{23},\cdots,u_{2n}$

………………………………

第 n 步：$l_{n1},l_{n2},l_{n3},\cdots,u_{nn}$

例：　已知 $A=\begin{pmatrix} 2 & 1 & -5 & 1 \\ 1 & -3 & 0 & -6 \\ 0 & 2 & -1 & 2 \\ 1 & 4 & -7 & 6 \end{pmatrix},\vec{b}=\begin{pmatrix} 8 \\ 9 \\ 5 \\ 0 \end{pmatrix}$，求矩阵 A 的 Crout 分解，并求解线性方程组 $A\vec{x}=\vec{b}$．

解：根据 Crout 分解的计算公式可有：

$l_{11}=a_{11}=2,u_{11}=1,u_{12}=\frac{a_{12}}{l_{11}}=\frac{1}{2},u_{13}=\frac{a_{13}}{l_{11}}=-\frac{5}{2},u_{14}==\frac{a_{14}}{l_{11}}=\frac{1}{2}$

$l_{21}=a_{12}=1,l_{22}=a_{22}-l_{21}u_{12}=-3-1\times\frac{1}{2}=-\frac{7}{2},u_{22}=1,u_{23}=\frac{a_{23}-l_{21}u_{13}}{l_{22}}=-\frac{5}{7},u_{24}=\frac{a_{24}-l_{21}u_{14}}{l_{22}}=\frac{13}{7}$

$l_{31}=a_{31}=0,l_{32}=a_{32}-l_{31}u_{12}=2,l_{33}=a_{33}-l_{31}u_{13}-l_{32}u_{23}=\frac{3}{7},u_{34}=\frac{a_{34}-l_{31}u_{14}-l_{32}u_{24}}{l_{33}}=-4$

$l_{41}=a_{41}=1,l_{42}=a_{42}-l_{41}u_{12}=\frac{7}{2},l_{43}=a_{43}-l_{41}u_{13}-l_{42}u_{23}=-2,l_{44}=a_{44}-l_{41}u_{14}-l_{42}u_{24}-l_{43}u_{34}=-9$

故有 $A=\begin{pmatrix}2&0&0&0\\1&-\frac{7}{2}&0&0\\0&2&\frac{3}{7}&0\\1&\frac{7}{2}&-2&-9\end{pmatrix}\begin{pmatrix}1&\frac{1}{2}&-\frac{5}{2}&\frac{1}{2}\\0&1&-\frac{5}{7}&\frac{13}{7}\\0&0&1&-4\\0&0&0&1\end{pmatrix}$，将 $A=LU$ 代入 $A\vec{x}=\vec{b}$ 得到

$LU\vec{x}=\vec{b}$，令 $U\vec{x}=\vec{y}$，则有 $L\vec{y}=\vec{b}\Rightarrow\begin{pmatrix}2&0&0&0\\1&-\frac{7}{2}&0&0\\0&2&\frac{3}{7}&0\\1&\frac{7}{2}&-2&-9\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}=\begin{pmatrix}8\\9\\5\\0\end{pmatrix}$，解得 $\vec{y}=\begin{pmatrix}4\\-\frac{10}{7}\\-5\\1\end{pmatrix}$，

代入 $U\vec{x}=\vec{y}\Rightarrow\begin{pmatrix}1&\frac{1}{2}&-\frac{5}{2}&\frac{1}{2}\\0&1&-\frac{5}{7}&\frac{13}{7}\\0&0&1&-4\\0&0&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}4\\-\frac{10}{7}\\-5\\1\end{pmatrix}$ 解得 $\vec{x}=\begin{pmatrix}3\\-4\\-1\\1\end{pmatrix}$.

类似地可以得到 n 阶方阵 A 的 Doolittle 分解的计算公式为：

$$\begin{cases}u_{ij}=a_{ij}-\sum_{k=1}^{i-1}l_{ik}u_{kj}(i=1,2,\cdots,n;j=i,\cdots,n)\\ l_{ij}=\dfrac{a_{ij}-\sum_{k=1}^{j-1}l_{ik}u_{kj}}{u_{jj}}(i=2,\cdots,n;j=1,\cdots,i-1)\end{cases}$$

二、乔累斯基 (Cholesky) 分解

定理 1-4 设 A 为 n 阶对称正定矩阵，则存在一个实的非奇异三角矩阵 L，使 $A=LL^T$，如果限定 L 的对角元素为正时，这种分解是唯一的． $A=LL^T$ 称为对称正定矩阵 A 的

Cholesky 分解，也称平方根分解.

$L=(l_{ij})_{n\times n}$ 是下三角矩阵，则 $l_{11}=\sqrt{a_{11}}, l_{i1}=\dfrac{a_{i1}}{l_{11}}(i=2,3,\cdots,n)$

对于 $k=2,3,\cdots,n$ 依次计算 $\begin{cases} l_{kk}=\sqrt{a_{kk}-\sum\limits_{s=1}^{k-1}l_{ks}^2} \\ l_{ik}=\dfrac{a_{ik}-\sum\limits_{t=1}^{k-1}l_{it}l_{kt}}{l_{kk}}(i=k+1,\cdots,n) \end{cases}$

对于分块矩阵 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$，利用分块初等矩阵可以得到它的分块三角分解：

1° 当 A 可逆时，有 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} I & O \\ CA^{-1} & I \end{pmatrix}\begin{pmatrix} A & O \\ O & D-CA^{-1}B \end{pmatrix}\begin{pmatrix} I & A^{-1}B \\ O & I \end{pmatrix}$；

2° 当 D 可逆时，有 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} I & BD^{-1} \\ O & I \end{pmatrix}\begin{pmatrix} A-BD^{-1}C & O \\ O & D \end{pmatrix}\begin{pmatrix} I & O \\ D^{-1}C & I \end{pmatrix}$.

例： 求对称正定矩阵 $A=\begin{pmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix}$ 的 Cholesky 分解.

解：有 Cholesky 分解的公式可得：

$l_{11}=\sqrt{a_{11}}=\sqrt{5}$

$l_{21}=\dfrac{a_{21}}{l_{11}}=\dfrac{2}{\sqrt{5}}, l_{22}=\sqrt{a_{22}-l_{21}^2}=\dfrac{1}{\sqrt{5}}$

$l_{31}=\dfrac{a_{31}}{l_{11}}=-\dfrac{4}{\sqrt{5}}, l_{32}=\dfrac{a_{32}-l_{31}l_{21}}{l_{22}}=-\dfrac{2}{\sqrt{5}}, l_{33}=\sqrt{a_{33}-l_{31}^2-l_{32}^2}=1$

$$\therefore A=\begin{pmatrix} \sqrt{5} & 0 & 0 \\ \dfrac{2}{\sqrt{5}} & \dfrac{1}{\sqrt{5}} & 0 \\ -\dfrac{4}{\sqrt{5}} & -\dfrac{2}{\sqrt{5}} & 1 \end{pmatrix}\begin{pmatrix} \sqrt{5} & \dfrac{2}{\sqrt{5}} & -\dfrac{4}{\sqrt{5}} \\ 0 & \dfrac{1}{\sqrt{5}} & -\dfrac{2}{\sqrt{5}} \\ 0 & 0 & 1 \end{pmatrix}$$

定理 1-5 设 A 为 n 阶对称正定矩阵，则 A 可唯一地分解为 $A=LDL^T$，其中 L 为下三角

矩阵；D 为对角阵，且对角线元素是 L 对角线元素的倒数，即

$$L=\begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \cdots & & \cdots & \\ l_{n1} & \cdots & \cdots & l_{nn} \end{pmatrix}, D=\begin{pmatrix} \frac{1}{l_{11}} & & & \\ & \frac{1}{l_{22}} & & \\ & & \cdots & \\ & & & \frac{1}{l_{nn}} \end{pmatrix}.\quad 而\ l_{ij}=a_{ij}-\sum_{k=1}^{j-1}\frac{l_{ik}l_{jk}}{l_{kk}},\begin{pmatrix} i=1,\cdots,n \\ j=1,\cdots,i \end{pmatrix},$$

$A=LDL^T$ 称为对称正定矩阵 A 的不带平方根 Cholesky 分解.

例：　求对称正定矩阵 $A=\begin{pmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix}$ 的不带平方根的 Cholesky 分解.

解：根据公式 $l_{ij}=a_{ij}-\sum_{k=1}^{j-1}\frac{l_{ik}l_{jk}}{l_{kk}},\begin{pmatrix} i=1,\cdots,n \\ j=1,\cdots,i \end{pmatrix}$ 可得：

$$l_{11}=a_{11}-0=5$$

$$l_{21}=a_{21}-0=2, l_{22}=a_{22}-\frac{l_{21}l_{21}}{l_{11}}=\frac{1}{5}$$

$$l_{31}=a_{31}-0=-4, l_{32}=a_{32}-\frac{l_{31}l_{21}}{l_{11}}=-\frac{2}{5}, l_{33}=a_{33}-\frac{l_{31}l_{31}}{l_{11}}-\frac{l_{32}l_{32}}{l_{22}}=1$$

$$\therefore L=\begin{pmatrix} 5 & & \\ 2 & \frac{1}{5} & \\ -4 & -\frac{2}{5} & 1 \end{pmatrix}, D=\begin{pmatrix} \frac{1}{5} & & \\ & 5 & \\ & & 1 \end{pmatrix},$$

$$\therefore A=LDL^T=\begin{pmatrix} 5 & & \\ 2 & \frac{1}{5} & \\ -4 & -\frac{2}{5} & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{5} & & \\ & 5 & \\ & & 1 \end{pmatrix}\begin{pmatrix} 5 & 2 & -4 \\ & \frac{1}{5} & -\frac{2}{5} \\ & & 1 \end{pmatrix}.$$

§2　矩阵的 QR（正交三角）分解

2.1　QR 分解的概念

定义 2-1　如果实（复）非奇异矩阵 A 能化成正交（酉）矩阵 Q 与实（复）非奇异上三角矩阵 R 的乘积，即 $A=QR$，则称其为 A 的 **QR 分解**.

定理 2-1　任何实的非奇异 n 阶矩阵 A 可以分解成正交矩阵 Q 和上三角矩阵 R 的乘积，且除去相差一个对角线元素之绝对值全等于 1 的对角矩阵因子 D 外，分解式 $A=QR$ 唯一.

例：　试求矩阵 $A=\begin{pmatrix}0&3&1\\0&4&-2\\2&1&2\end{pmatrix}$ 的 QR 分解.

解：易知 $|A|=-20\neq 0$，令 $\overrightarrow{\alpha_1}=(0,0,2)^T,\overrightarrow{\alpha_2}=(3,4,1)^T,\overrightarrow{\alpha_3}=(1,-2,2)^T$，由 $Schmidt$ 方法可得：$\overrightarrow{\beta_1'}=\overrightarrow{\alpha_1}=(0,0,2)^T$

$$\overrightarrow{\beta_2'}=\overrightarrow{\alpha_2}-\frac{(\overrightarrow{\beta_1'},\overrightarrow{\alpha_2})}{(\overrightarrow{\beta_1'},\overrightarrow{\beta_1'})}\overrightarrow{\beta_1'}=\overrightarrow{\alpha_2}-\frac{1}{2}\overrightarrow{\beta_1'}=(3,4,0)^T$$

$$\overrightarrow{\beta_3'}=\overrightarrow{\alpha_3}-\frac{(\overrightarrow{\beta_1'},\overrightarrow{\alpha_3})}{(\overrightarrow{\beta_1'},\overrightarrow{\beta_1'})}\overrightarrow{\beta_1'}-\frac{(\overrightarrow{\beta_2'},\overrightarrow{\alpha_3})}{(\overrightarrow{\beta_2'},\overrightarrow{\beta_2'})}\overrightarrow{\beta_2'}=\overrightarrow{\alpha_3}-\overrightarrow{\beta_1'}+\frac{1}{5}\overrightarrow{\beta_2'}=(\frac{8}{5},-\frac{6}{5},0)^T$$

再单位化可得 $\overrightarrow{\beta_1}=\frac{1}{2}\overrightarrow{\beta_1'}=(0,0,1)^T$，$\overrightarrow{\beta_2}=\frac{1}{5}\overrightarrow{\beta_2'}=(\frac{3}{5},\frac{4}{5},0)^T$，$\overrightarrow{\beta_3}=\frac{1}{2}\overrightarrow{\beta_3'}=(\frac{4}{5},-\frac{3}{5},0)^T$

于是可得 $\overrightarrow{\alpha_1}=2\overrightarrow{\beta_1},\overrightarrow{\alpha_2}=\overrightarrow{\beta_1}+5\overrightarrow{\beta_2},\overrightarrow{\alpha_3}=2\overrightarrow{\beta_1}-\overrightarrow{\beta_2}+2\overrightarrow{\beta_3}$

$$\therefore A=(\overrightarrow{\alpha_1},\overrightarrow{\alpha_2},\overrightarrow{\alpha_3})=(\overrightarrow{\beta_1},\overrightarrow{\beta_2},\overrightarrow{\beta_3})\begin{pmatrix}2&1&2\\0&5&-1\\0&0&2\end{pmatrix}=\begin{pmatrix}0&\frac{3}{5}&\frac{4}{5}\\0&\frac{4}{5}&-\frac{3}{5}\\1&0&0\end{pmatrix}\begin{pmatrix}2&1&2\\0&5&-1\\0&0&2\end{pmatrix}=QR$$

2.2　QR 分解的实际求法

一、吉温斯（Givens）方法

定理 2-3 任何实非奇异矩阵可通过左连乘初等旋转阵化为上三角.

Givens 方法需要做 $\frac{n(n-1)}{2}$ 个初等旋转矩阵的连乘积.

例： 用 *Givens* 方法求矩阵 $A=\begin{pmatrix}0&3&1\\0&4&-2\\2&1&2\end{pmatrix}$ 的 QR 分解.

解：分析矩阵，以 R_{13} 左乘 A 消去第 3 行，第 1 列处的元素.

$$a_{21}^{'}=\sqrt{0^2+2^2}=2, c=0, s=1, R_{13}=\begin{pmatrix}0&0&1\\0&1&0\\-1&0&0\end{pmatrix},$$

$$A^{(1)}=R_{13}A=\begin{pmatrix}0&0&1\\0&1&0\\-1&0&0\end{pmatrix}\begin{pmatrix}0&3&1\\0&4&-2\\2&1&2\end{pmatrix}=\begin{pmatrix}2&1&2\\0&4&-2\\0&-3&-1\end{pmatrix}$$

再根据 $A^{(1)}$ 中的元素计算得到

$$a_{22}^{'}=\sqrt{4^2+(-3)^2}=5, c=\frac{4}{5}, s=-\frac{3}{5}, R_{23}=\begin{pmatrix}1&0&0\\0&\frac{4}{5}&-\frac{3}{5}\\0&\frac{3}{5}&\frac{4}{5}\end{pmatrix},$$

$$A^{(2)}=R_{23}A^{(1)}=\begin{pmatrix}1&0&0\\0&\frac{4}{5}&-\frac{3}{5}\\0&\frac{3}{5}&\frac{4}{5}\end{pmatrix}\begin{pmatrix}2&1&2\\0&4&-2\\0&-3&-1\end{pmatrix}=\begin{pmatrix}2&1&2\\0&5&-1\\0&0&-2\end{pmatrix}，\text{即有}$$

$$R_{23}R_{13}\begin{pmatrix}0&3&1\\0&4&-2\\2&1&2\end{pmatrix}=\begin{pmatrix}2&1&2\\0&5&-1\\0&0&-2\end{pmatrix}=R_{23}R_{13}A,$$

$$Q=(R_{23}R_{13})^{-1}=\left(\begin{pmatrix}1&0&0\\0&\frac{4}{5}&-\frac{3}{5}\\0&\frac{3}{5}&\frac{4}{5}\end{pmatrix}\begin{pmatrix}0&0&1\\0&1&0\\-1&0&0\end{pmatrix}\right)^{-1}=\begin{pmatrix}0&0&1\\\frac{3}{5}&\frac{4}{5}&0\\-\frac{4}{5}&\frac{3}{5}&0\end{pmatrix}^{-1}=\begin{pmatrix}0&\frac{3}{5}&-\frac{4}{5}\\0&\frac{4}{5}&\frac{3}{5}\\1&0&0\end{pmatrix}$$

$$\therefore A=\begin{pmatrix}0 & \frac{3}{5} & -\frac{4}{5}\\ 0 & \frac{4}{5} & \frac{3}{5}\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}2 & 1 & 2\\ 0 & 5 & -1\\ 0 & 0 & -2\end{pmatrix}=QR$$

二、豪斯蒙德（Householder）方法

定理 2-4　任何实的 n 阶矩阵 A 可用初等反射矩阵 $H=I-2\vec{\omega}\vec{\omega}^T$ 化为上三角矩阵.

例：　用 Householder 方法求矩阵 $A=\begin{pmatrix}0 & 3 & 1\\ 0 & 4 & -2\\ 2 & 1 & 2\end{pmatrix}$ 的 QR 分解.

解：因为 $\overrightarrow{\alpha^{(1)}}=(0,0,2)^T\neq\vec{0}$，所以作 $\vec{\omega}^{(1)}=\dfrac{\overrightarrow{\alpha^{(1)}}-\left|\overrightarrow{\alpha^{(1)}}\right|\vec{e}}{\left|\overrightarrow{\alpha^{(1)}}-\left|\overrightarrow{\alpha^{(1)}}\right|\vec{e}\right|}=(-\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}})^T$，

$$\therefore H^{(1)}=I_3-2\vec{\omega}^{(1)}(\vec{\omega}^{(1)})^T=\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}\Rightarrow H^{(1)}A=\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}0 & 3 & 1\\ 0 & 4 & -2\\ 2 & 1 & 2\end{pmatrix}=\begin{pmatrix}2 & 1 & 2\\ 0 & 4 & -2\\ 0 & 3 & 1\end{pmatrix}$$

再令 $\overrightarrow{\alpha^{(2)}}=(4,3)^T$，再作 $\hat{\omega}^{(2)}=\dfrac{(4,3)^T-\left|(4,3)^T\right|\vec{e}}{\left|(4,3)^T-\left|(4,3)^T\right|\vec{e}\right|}=(-\frac{1}{\sqrt{10}},\frac{3}{\sqrt{10}})^T$

$$\therefore \widehat{H^{(2)}}=I_2-2\hat{\omega}^{(2)}(\hat{\omega}^{(2)})^T=\frac{1}{5}\begin{pmatrix}4 & 3\\ 3 & -4\end{pmatrix}\Rightarrow H^{(2)}=\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{4}{5} & \frac{3}{5}\\ 0 & \frac{3}{5} & -\frac{4}{5}\end{pmatrix}$$

$$\Rightarrow H^{(2)}H^{(1)}A=\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{4}{5} & \frac{3}{5}\\ 0 & \frac{3}{5} & -\frac{4}{5}\end{pmatrix}\begin{pmatrix}2 & 1 & 2\\ 0 & 4 & -2\\ 0 & 3 & 1\end{pmatrix}=\begin{pmatrix}2 & 1 & 2\\ 0 & 5 & -1\\ 0 & 0 & -2\end{pmatrix}=R$$

$$Q=(H^{(2)}H^{(1)})^{-1}=\left(\begin{pmatrix}1&0&0\\0&\frac{4}{5}&\frac{3}{5}\\0&\frac{3}{5}&-\frac{4}{5}\end{pmatrix}\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\right)^{-1}=\begin{pmatrix}0&0&1\\\frac{3}{5}&\frac{4}{5}&0\\-\frac{4}{5}&\frac{3}{5}&0\end{pmatrix}^{-1}=\begin{pmatrix}0&\frac{3}{5}&-\frac{4}{5}\\0&\frac{4}{5}&\frac{3}{5}\\1&0&0\end{pmatrix}$$

$$\therefore A=\begin{pmatrix}0&\frac{3}{5}&-\frac{4}{5}\\0&\frac{4}{5}&\frac{3}{5}\\1&0&0\end{pmatrix}\begin{pmatrix}2&1&2\\0&5&-1\\0&0&-2\end{pmatrix}=QR$$

§3 矩阵的最大秩分解

定义 3-1 设 $m\times n$ 矩阵 $A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}$. 如果当 $m\le n$ 时，存在有 $rankA=n$；或者当 $m\ge n$ 时，存在有 $rankA=m$，则称这两种长方阵为**最大秩长方阵**（**满秩长方阵**），前者又称**行最大秩矩阵**（**行满秩矩阵**），后者又称**列最大秩矩阵**（**列满秩矩阵**或**高矩阵**）.

显然，最大秩长方阵具有如下性质：

$$rank(AA^T)=m,(A=(a_{ij})_{m\times n},m\le n)\text{ 或 }rank(A^TA)=n,(A=(a_{ij})_{m\times n},m\ge n)$$

定义 3-2 设 A 为 $m\times n$ 且秩为 $r>0$ 的复矩阵，记为 $A\in C_r^{m\times n}$，如果 $\exists B\in C_r^{m\times r},C\in C_r^{r\times n}$，$sb.A=BC$，则称此式为矩阵 A 的**最大秩分解**（**满秩分解**）.

定理 3-1 设 $A\in C_r^{m\times n}$，则一定存在 $B\in C_r^{m\times r},C\in C_r^{r\times n}$，使得 $A=BC$.

例： 求矩阵 $A=\begin{pmatrix}0&1&-1&-1&1\\0&-2&2&-2&6\\0&1&-1&-2&3\end{pmatrix}$ 的满秩分解.

解：将 A 只进行行初等变换，化为行标准形

$$A=\begin{pmatrix}0&1&-1&-1&1\\0&-2&2&-2&6\\0&1&-1&-2&3\end{pmatrix}\simeq\begin{pmatrix}0&1&-1&-1&1\\0&0&0&-4&8\\0&0&0&-1&2\end{pmatrix}\simeq\begin{pmatrix}0&1&-1&0&-1\\0&0&0&1&-2\\0&0&0&0&0\end{pmatrix}$$

于是取 A 的第 2、4 列组成矩阵 $B=\begin{pmatrix}1&-1\\-2&-2\\1&-2\end{pmatrix}$，再取 A 的行标准形的前两个非零行组成矩阵 $C=\begin{pmatrix}0&1&-1&0&-1\\0&0&0&1&-2\end{pmatrix}$，$\therefore A=\begin{pmatrix}1&-1\\-2&-2\\1&-2\end{pmatrix}\begin{pmatrix}0&1&-1&0&-1\\0&0&0&1&-2\end{pmatrix}$.

A 的最大秩分解不是唯一的，但最大秩分解之间有如下关系：

定理 3-2 设 $A\in C_r^{m\times n}$，且 $A=BC=\widetilde{B}\widetilde{C}$ 均为 A 的最大秩分解，则有：

1° $\exists Q\in C_r^{r\times r},sb.B=\widetilde{B}Q,C=Q^{-1}\widetilde{C}$；

2° $C^H(CC^H)^{-1}(B^HB)^{-1}B^H=\widetilde{C}^H(\widetilde{C}\widetilde{C}^H)^{-1}(\widetilde{B}^H\widetilde{B})^{-1}\widetilde{B}^H$.

例： 求矩阵 $A=\begin{pmatrix}2&1&-2&3\\0&4&1&1\\2&5&-2&4\end{pmatrix}$ 的最大秩分解 $A=BC$，并计算 $C^T(CC^T)^{-1}(B^TB)^{-1}B^T$.

解：将 A 只进行行初等变换，化为行标准形

$$A=\begin{pmatrix}2&1&-2&3\\0&4&1&1\\2&5&-2&4\end{pmatrix}\simeq\begin{pmatrix}2&1&-2&3\\0&4&1&1\\0&4&0&1\end{pmatrix}\simeq\begin{pmatrix}2&1&-2&3\\0&0&1&0\\0&4&0&1\end{pmatrix}\simeq\begin{pmatrix}1&0&0&\frac{11}{8}\\0&1&0&\frac{1}{4}\\0&0&1&0\end{pmatrix}$$，显 然

$rankA=r=3$， A 为最大秩矩阵，$\therefore A=I_{3\times3}A$，此处 $B=I_{3\times3}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},C=A$，

$$\therefore C^T(CC^T)^{-1}(B^TB)^{-1}B^T=A^T(AA^T)^{-1}(I^TI)^{-1}I^T=A^T(AA^T)^{-1}=A^T(A^T)^{-1}A^{-1}$$

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

§4 矩阵的奇异值分解

命题 4-1 设 $A\in C^{m\times n}$，则有1° A^HA 与 AA^H 的特征值均为非负实数；2° A^HA 与 AA^H 的非零特征值相同.

定义 4-1 设 $A\in C_r^{m\times n}$，A^HA 的特征值为 $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_r>\lambda_{r+1}=\lambda_{r+2}=\cdots=\lambda_n=0$，则称 $\sigma_i=\sqrt{\lambda_i}\,(i=1,2,\cdots,r)$ 为矩阵 A 的**正奇异值**，简称**奇异值**.

定义 4-2 设 $A,B\in C^{m\times n}$，如果存在 m 阶酉矩阵 U 和 n 阶酉矩阵 V，使得 $B=UAV$，则称 A 与 B **酉等价**或**酉相抵**.

定理 4-1 若 A 与 B 酉等价，则 A 与 B 有相同的奇异值.

定理 4-2 设 $A\in C_r^{m\times n}$，则存在 m 阶酉矩阵 U 和 n 阶酉矩阵 V，使得 $U^HAV=\begin{pmatrix}\Delta & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}$或 $A=U\begin{pmatrix}\Delta & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}V^H$，其中 $\Delta=diag(\delta_1,\delta_2,\cdots,\delta_r),\delta_i$ 为复数，且它的模 $|\delta_i|=\sigma_i(i=1,2,\cdots,r)$，而 σ_i 是 A 的全部奇异值.

例： 求矩阵 $A=\begin{pmatrix}1&0&0&-1\\0&1&0&1\\0&0&0&0\end{pmatrix}$ 的奇异值分解.

解：$AA^H=\begin{pmatrix}1&0&0&-1\\0&1&0&1\\0&0&0&0\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\\-1&1&0\end{pmatrix}=\begin{pmatrix}2&-1&0\\-1&2&0\\0&0&0\end{pmatrix}$

$\lambda I-AA^H=\begin{pmatrix}\lambda-2&1&0\\1&\lambda-2&0\\0&0&\lambda\end{pmatrix}\simeq\begin{pmatrix}1&\lambda-2&0\\0&1-(\lambda-2)^2&0\\0&0&\lambda\end{pmatrix}$，由

$|\lambda I-AA^H|=0$ 得 AA^H 的特征值为 $\lambda_1=1,\lambda_2=3,\lambda_3=0$，对应的特征向量依次为：

$p_1=\begin{pmatrix}1\\1\\0\end{pmatrix},p_2=\begin{pmatrix}-1\\1\\0\end{pmatrix},p_3=\begin{pmatrix}0\\0\\1\end{pmatrix}$，$A$的奇异值为$\sigma_1=1,\sigma_2=\sqrt{3}$，则$\Delta=\begin{pmatrix}1&\\&\sqrt{3}\end{pmatrix}$，则

正交单位化可得酉矩阵$U=\begin{pmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\0&0&1\end{pmatrix}$

$V_1=A^HU_1\Delta^{-H}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\\-1&1&0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\0&0\end{pmatrix}\begin{pmatrix}1&\\&\frac{1}{\sqrt{3}}\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}\\0&0\\0&\frac{2}{\sqrt{6}}\end{pmatrix}$，设有单位向量

$q=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}$ 与向量 $\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\\0\end{pmatrix}$ 和 $\begin{pmatrix}-\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\0\\\frac{2}{\sqrt{6}}\end{pmatrix}$ 正交，则可得 $V_2=\begin{pmatrix}0&\frac{1}{\sqrt{3}}\\0&-\frac{1}{\sqrt{3}}\\1&0\\0&\frac{1}{\sqrt{3}}\end{pmatrix}$，于是可得

$V=(V_1|V_2)=\begin{pmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}&0&\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}&0&-\frac{1}{\sqrt{3}}\\0&0&1&0\\0&\frac{2}{\sqrt{6}}&0&\frac{1}{\sqrt{3}}\end{pmatrix}$，$\therefore A=U\begin{pmatrix}1&0&0&0\\0&\sqrt{3}&0&0\\0&0&0&0\end{pmatrix}V^H$

本章测试题

一、设 $A \in C_r^{m\times n}$，且 A 的奇异值分解为 $A = U\begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}V^H$，试求矩阵 $B = \begin{pmatrix} A \\ A \end{pmatrix}$ 的一个奇异值分解.

二、求矩阵 $A = \begin{pmatrix} 5 & 2 & -4 & 0 \\ 2 & 1 & -2 & 1 \\ -4 & -2 & 5 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix}$ 的 *Doolittle* 分解与 *Crout* 分解.

三、已知 $A=\begin{pmatrix}3&2&1\\2&2&0\\1&0&3\end{pmatrix},\vec{b}=\begin{pmatrix}5\\3\\4\end{pmatrix}$，求对称正定矩阵 A 的 $Cholesky$ 分解，并利用该分解求解线性方程组 $A\vec{x}=\vec{b}$.

四、设 n 阶矩阵 D 可逆，证明：矩阵 $\begin{pmatrix}A&B\\C&D\end{pmatrix}$ 可逆的充要条件是 $A-BD^{-1}C$ 可逆．当矩阵 $\begin{pmatrix}A&B\\C&D\end{pmatrix}$ 可逆时，推导其逆矩阵的表达式．

五、分别利用 *Householder* 矩阵和 *Schmidt* 正交化方法求矩阵 $A=\begin{pmatrix}0&4&1\\1&1&1\\0&3&2\end{pmatrix}$ 的 QR 分解.

六、利用 *Givens* 矩阵求矩阵 $A=\begin{pmatrix}2&2&1\\0&2&2\\2&1&2\end{pmatrix}$ 的 QR 分解.

七、分别用 $Householder$ 矩阵和 $Givens$ 矩阵使矩阵 $A=\begin{pmatrix} 0 & 12 & 16 \\ 12 & 288 & 309 \\ 16 & 309 & 312 \end{pmatrix}$ 正交相似于三队角矩阵.

八、求下列矩阵的 $Hermite$ 标准形和满秩分解：

（1）$A=\begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 2 & 1 & -1 \\ 1 & 0 & 2 & 1 \end{pmatrix}$；（2）$B=\begin{pmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{pmatrix}$

九、求下列矩阵的奇异分解：

（1）$A=\begin{pmatrix}1 & 0\\ 0 & 1\\ 1 & 1\end{pmatrix}$；（2）$B=\begin{pmatrix}1 & 0 & 0\\ 2 & 0 & 0\end{pmatrix}$

第五章　线性赋范空间与矩阵范数

§1　线性赋范空间

1.1　向量的范数

定义 1-1　如果 V 是数域 P 上的线性空间，且 $\forall \vec{x} \in V$，对应着一个实值函数 $\|\vec{x}\|$，它满足以下三个条件：1．非负性：当 $\vec{x} \neq \vec{0}$ 时，$\|\vec{x}\| > 0$；当且仅当 $\vec{x} = \vec{0}$ 时，$\|\vec{x}\| = 0$；

2．齐次性：$\|k\vec{x}\| = |k|\|\vec{x}\|, k \in P$；

3．三角不等式：$\|\vec{x}+\vec{y}\| \le \|\vec{x}\| + \|\vec{y}\|, \vec{x}, \vec{y} \in V$．

则称 $\|\vec{x}\|$ 为 V 上向量 $\vec{x}$ 的**范数**（norm）．

★　几种常见的向量范数：

设 $\vec{x} = (x_1, x_2, \cdots, x_n)^T \in R^n$

① 2-范数　$\|\vec{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = (\vec{x}^T \vec{x})^{\frac{1}{2}}$；

② ∞-范数　$\|\vec{x}\|_\infty = \max\limits_{1 \le i \le n} |x_i|$；

③ 1-范数　$\|\vec{x}\|_1 = \sum\limits_{i=1}^{n} |x_i|$；

④ p-范数　$\|\vec{x}\|_p = \left(\sum\limits_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}, 1 \le p < +\infty$．

（霍尔德 $H\ddot{o}lder$ 不等式：$\sum\limits_{i=1}^{n} |a_i b_i| \le \left(\sum\limits_{i=1}^{n} |a_i|^p \right)^{\frac{1}{p}} \left(\sum\limits_{i=1}^{n} |b_i|^q \right)^{\frac{1}{q}}$，其中 $p, q > 1$，且 $\frac{1}{p} + \frac{1}{q} = 1$）

定义 1-2　定义了向量范数 $\|\bullet\|$ 的线性空间 V^n，称为**线性赋范空间**．其中 $\|\bullet\|$ 表示泛指的任何一种范数．

例：　设 $A, B \in C^{n \times n}$，且 B 可逆．对于 C^n 中的列向量 $\vec{\alpha}$，定义实数 $\|\vec{\alpha}\| = \|A\vec{\alpha}\|_1 + 3\|B\vec{\alpha}\|_3$，

验证$\left\|\vec{\alpha}\right\|$是$C^n$中的向量范数.

解：当$\vec{\alpha}=\vec{0}$时，$A\vec{\alpha}=\vec{0},B\vec{\alpha}=\vec{0}$，从而$\left\|\vec{\alpha}\right\|=0$；当$\vec{\alpha}\neq\vec{0}$时，$\left\|A\vec{\alpha}\right\|_1\geq 0$，由$B$可逆知$B\vec{\alpha}\neq\vec{0}$，从而$\left\|B\vec{\alpha}\right\|_3>0$，故$\left\|\vec{\alpha}\right\|>0$.

$\forall k\in C$，有$\left\|k\vec{\alpha}\right\|=\left\|A(k\vec{\alpha})\right\|_1+3\left\|B(k\vec{\alpha})\right\|_3=|k|\left\|A\vec{\alpha}\right\|_1+3|k|\left\|B\vec{\alpha}\right\|_3=|k|\left\|\vec{\alpha}\right\|$

$\forall\vec{\beta}\in C^n$，有
$$\begin{aligned}\left\|\vec{\alpha}+\vec{\beta}\right\|&=\left\|A(\vec{\alpha}+\vec{\beta})\right\|_1+3\left\|B(\vec{\alpha}+\vec{\beta})\right\|_3\leq\left\|A\vec{\alpha}\right\|_1+\left\|A\vec{\beta}\right\|_1+3\left\|B\vec{\alpha}\right\|_3+3\left\|B\vec{\beta}\right\|_3\\&=(\left\|A\vec{\alpha}\right\|_1+3\left\|B\vec{\alpha}\right\|_3)+(\left\|A\vec{\beta}\right\|_1+3\left\|B\vec{\beta}\right\|_3)=\left\|\vec{\alpha}\right\|+\left\|\vec{\beta}\right\|\end{aligned}$$

故$\left\|\vec{\alpha}\right\|$是$C^n$中的向量范数.

例： 设A是任一n阶对称正定矩阵，$\vec{x}\in R^n$，证明：$\left\|\vec{x}\right\|_a=(\vec{x}^T A\vec{x})^{\frac{1}{2}}$是一种向量范数.

证：当$\vec{x}=\vec{0}$时，$\left\|\vec{x}\right\|_a=0$；当$\vec{x}\neq\vec{0}$时，由$A$对称正定可知$\vec{x}^T A\vec{x}>0$，即$\left\|\vec{x}\right\|_a>0$.

$\forall k\in R$，$\left\|k\vec{x}\right\|_a=\sqrt{(k\vec{x})^T A(k\vec{x})}=|k|\sqrt{\vec{x}^T A\vec{x}}=|k|\left\|\vec{x}\right\|_a$.

再由A对称正定可知，存在正交矩阵Q，使得$Q^T AQ=diag(\lambda_1,\lambda_2,\cdots,\lambda_n),\lambda_i>0$，从而有

$A=L^T L$，$\therefore\left\|\vec{x}\right\|_a^2=\vec{x}^T L^T L\vec{x}=(L\vec{x})^T(L\vec{x})=\left\|L\vec{x}\right\|_2^2$，$\forall\vec{y}\in R^n$，有

$\left\|\vec{x}+\vec{y}\right\|_a=\left\|L(\vec{x}+\vec{y})\right\|_2=\left\|L\vec{x}+L\vec{y}\right\|_2\leq\left\|L\vec{x}\right\|_2+\left\|L\vec{y}\right\|_2=\left\|\vec{x}\right\|_a+\left\|\vec{y}\right\|_a$

$\therefore\left\|\vec{x}\right\|_a=(\vec{x}^T A\vec{x})^{\frac{1}{2}}$是一种向量范数.

1.2 向量范数的性质

定义 1-3 设$\left\|\vec{x}\right\|_a,\left\|\vec{x}\right\|_b$是$n$维线性空间$V^n$上定义的任意两种范数，若存在两个与$\vec{x}$无关的正常数$c_1,c_2$，使得$c_1\left\|\vec{x}\right\|_b\leq\left\|\vec{x}\right\|_a\leq c_2\left\|\vec{x}\right\|_b,\forall\vec{x}\in V^n$，则称$\left\|\vec{x}\right\|_a,\left\|\vec{x}\right\|_b$是**等价的**.

定理 1-1 有限维线性空间上的不同范数是等价的.

定理 1-2 R^n中$\lim\limits_{x\to\infty}\vec{x}^{(k)}=\vec{x}^{(*)}\Leftrightarrow\left\|\vec{x}^{(k)}-\vec{x}^{(*)}\right\|\to 0(k\to\infty)$，其中$\|\bullet\|$为向量的任一种范数.

例： 验证$\|\vec{\alpha}\|_1,\|\vec{\alpha}\|_2,\|\vec{\alpha}\|_\infty$两两等价.

证明：设$\vec{\alpha}=(\xi_1,\xi_2,\cdots,\xi_n)^T$，则有$\|\vec{\alpha}\|_1=\sum_{i=1}^{n}|\xi_i|\le n\cdot\max_{1\le j\le n}|\xi_j|=n\|\vec{\alpha}\|_\infty$，

$\|\vec{\alpha}\|_1\ge\max_{1\le j\le n}|\xi_j|=\|\vec{\alpha}\|_\infty$，即$1\cdot\|\vec{\alpha}\|_\infty\le\|\vec{\alpha}\|_1\le n\|\vec{\alpha}\|_\infty-(1)$；

又由于$\|\vec{\alpha}\|_2^2=\sum_{i=1}^{n}|\xi_i|^2\le n\cdot\max_{1\le j\le n}|\xi_j|^2=n\|\vec{\alpha}\|_2^\infty$，$\|\vec{\alpha}\|_2^2\ge\max_{1\le j\le n}|\xi_j|^2=\|\vec{\alpha}\|_2^\infty$，那么

$1\cdot\|\vec{\alpha}\|_\infty\le\|\vec{\alpha}\|_2\le\sqrt{n}\|\vec{\alpha}\|_\infty-(2)$；再由（1）、（2）式可得

$\frac{1}{n}\|\vec{\alpha}\|_1\le\|\vec{\alpha}\|_\infty\le\|\vec{\alpha}\|_2\le\sqrt{n}\|\vec{\alpha}\|_\infty\le\sqrt{n}\|\vec{\alpha}\|_1$，综上可得$\|\vec{\alpha}\|_1,\|\vec{\alpha}\|_2,\|\vec{\alpha}\|_\infty$两两等价.

例： 设数域R上的多项式空间$P[t]_{n-1}$的两个基为：

（Ⅰ）：$f_1(t)=1,f_2(t)=t,\cdots,f_n(t)=t^{n-1}$；（Ⅱ）：$(g_1,g_2,\cdots,g_n)=(f_1,f_2,\cdots,f_n)C$.

其中C为可逆矩阵，$f(t)\in P[t]_{n-1}$在基(Ⅰ)和基(Ⅱ)下的坐标分别为$\vec{\alpha}=(\xi_1,\xi_2,\cdots,\xi_n)^T$，$\vec{\beta}=(\eta_1,\eta_2,\cdots,\eta_n)^T$.

(1) 证明：$\forall f(t)\in P[t]_{n-1},\|\vec{\alpha}\|_2=\|\vec{\beta}\|_2\Leftrightarrow C^TC=I$；

(2) 取$n=3$，找出使(1)成立的基(Ⅱ)，而且$g_i(t)\ne\pm f_j(t)(i,j=1,2,3)$.

解：(1) 充分性．由$C^TC=I$及坐标变换公式$\vec{\alpha}=C\vec{\beta}$可得

$\|\vec{\alpha}\|_2^2=\vec{\alpha}^T\vec{\alpha}=(C\vec{\beta})^T(C\vec{\beta})=\vec{\beta}^T(C^TC)\vec{\beta}=\vec{\beta}^T\vec{\beta}=\|\vec{\beta}\|_2^2$，即$\|\vec{\alpha}\|_2=\|\vec{\beta}\|_2$.

必要性．由$\|\vec{\alpha}\|_2=\|\vec{\beta}\|_2$知$\vec{\alpha}^T\vec{\alpha}=\vec{\beta}^T\vec{\beta}$，由坐标变换公式

$\vec{\alpha}=C\vec{\beta}\Rightarrow(C\vec{\beta})^T(C\vec{\beta})=\vec{\beta}^T(C^TC)\vec{\beta}=\vec{\beta}^T\vec{\beta}$，令$B=C^TC=(b_{ij})_{n\times n}$，则$B$是是对称矩阵．取$f(t)=g_i(t)$时，$\vec{\beta}=\vec{e_i}$，于是可得$b_{ii}=1(i=1,2,\cdots,n)$；取$f(t)=g_i(t)+g_j(t)$时，$\vec{\beta}=\vec{e_i}+\vec{e_j}(i\ne j)$，于是$b_{ii}+b_{ij}+b_{ji}+b_{jj}=2$，即$b_{ij}=0(i\ne j;i,j=1,2,\cdots,n)$，故$B=I$，即$C^TC=I$.

(2) 易知，满足（1）的 C 为正交矩阵．取 $C=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{pmatrix}$，则使(1)成立的基(Ⅱ)

为：$g_1(t)=\frac{1}{\sqrt{2}}(1-t), g_2(t)=\frac{1}{\sqrt{3}}(1+t-t^2), g_3(t)=\frac{1}{\sqrt{6}}(1+t+2t^2)$

§2　矩阵的范数

2.1　矩阵范数的定义与性质

定义 2-1　设 $A\in C^{m\times n}$，按某一法则在 $C^{m\times n}$ 上规定 A 的一个实数值函数，记作 $\|A\|$，它满足下面四个条件：1．非负性：如果 $A\neq 0$，则 $\|A\|>0$；如果 $A=0$，则 $\|A\|=0$；

2．齐次性：$\forall k\in C, \|kA\|=|k|\|A\|$；

3．三角不等式：$\forall A,B\in C^{m\times n}, \|A+B\|\leq\|A\|+\|B\|$；

4．相容性：当矩阵乘积有意义时，还有 $\|AB\|\leq\|A\|\|B\|$．

则称 $\|A\|$ 为**矩阵范数**．

★ 设 $A=(a_{ij})_{m\times n}$，矩阵空间 $C^{m\times n}$ 中的常见矩阵范数如下：

① $\|A\|_{m_1}=\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$；② $\|A\|_{m_\infty}=n\cdot\max_{i,j}|a_{ij}|$；

③ $\|A\|_{m_2}=\|A\|_F=\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}}=\left(tr(AA^H)\right)^{\frac{1}{2}}=\left(tr(A^HA)\right)^{\frac{1}{2}}$．

定理 2-1　设 $A\in C^{m\times n}$，$\|A\|$ 是 $C^{m\times n}$ 上的矩阵范数，则 $C^{m\times n}$ 上的任意两个矩阵范数等价．

例：　设 $S\in C^{n\times n}$ 可逆，给定 $C^{n\times n}$ 中的矩阵范数 $\|\bullet\|_M$，对于 $A\in C^{n\times n}$，定义实数

$\|A\|=\|S^{-1}AS\|_M$，验证$\|A\|$是$C^{n\times n}$上的矩阵范数.

证：当$A=0$时，$\|A\|=0$；当$A\neq 0$时，$S^{-1}AS\neq 0$，从而$\|A\|>0$.

$\forall k\in C$，$\|kA\|=\|S^{-1}(kA)S\|_M=\|k(S^{-1}AS)\|_M=|k|\|S^{-1}AS\|_M=|k|\|A\|$.

$\forall B\in C^{n\times n}$

$$\|A+B\|=\|S^{-1}(A+B)S\|_M=\|S^{-1}AS+S^{-1}BS\|_M\leq\|S^{-1}AS\|_M+\|S^{-1}BS\|_M=\|A\|+\|B\|$$

$$\|AB\|=\|S^{-1}(AB)S\|_M=\|(S^{-1}AS)(S^{-1}BS)\|_M\leq\|S^{-1}AS\|_M\|S^{-1}BS\|_M=\|A\|\|B\|$$

故$\|A\|$是$C^{n\times n}$上的矩阵范数.

2.2 算子范数

定义 2-2 设$A\in C^{m\times n},\vec{x}\in C^n$，取定的向量范数$\|\vec{x}\|$和矩阵范数$\|A\|$满足不等式$\|A\vec{x}\|\leq\|A\|\|\vec{x}\|$则称矩阵范数$\|A\|$与向量范数$\|\vec{x}\|$是**相容**的.

定理 2-2 设$A\in C^{m\times n},\vec{x}=(x_1,x_2,\cdots,x_n)^T\in C^n$，且在$C^n$中已规定了向量的某种范数$\|\vec{x}\|$，则与向量范数$\|\vec{x}\|$相容的矩阵范数可以取作向量$A\vec{x}$的范数的最大值，即：$\|A\|=\max\limits_{\|x\|=1}\|A\vec{x}\|$. 这里向量$\vec{x}$取遍范数为 1 的所有向量的集合. 由此定义的相容范数为**算子范数**，或称为向量范数的**从属范数**.

定理 2-3 设$A=(a_{ij})\in C^{m\times n},\vec{x}=(x_1,x_2,\cdots,x_n)^T\in C^n$，则从属于向量$\vec{x}$的三种范数$\|\vec{x}\|_1$、$\|\vec{x}\|_2$、$\|\vec{x}\|_\infty$的算子范数依次是：

1. $\|A\|_1=\max\limits_j\sum\limits_{i=1}^m|a_{ij}|$（称为**列范数**）；2. $\|A\|_2=\sqrt{\lambda_{\max}(A^HA)}$（称为**谱范数**）；

3. $\|A\|_\infty=\max\limits_i\sum\limits_{j=1}^n|a_{ij}|$（称为**行范数**）.

例： 设 $A=\begin{pmatrix}2&1&0\\-1&2&3\\0&-2&1\end{pmatrix}$，计算 $\|A\|_1$，$\|A\|_2$，$\|A\|_\infty$，$\|A\|_F$.

解：$\|A\|_1=5$，$\|A\|_\infty=6$，$\|A\|_F=2\sqrt{6}$

$A^HA=\begin{pmatrix}2&-1&0\\1&2&-2\\0&3&1\end{pmatrix}\begin{pmatrix}2&1&0\\-1&2&3\\0&-2&1\end{pmatrix}=\begin{pmatrix}5&0&-3\\0&9&4\\-3&4&10\end{pmatrix}$，求得 A^HA 的特征值为

$\lambda_1=2.9542$，$\lambda_2=6.9305,\lambda_3=14.1153$，$\therefore\|A\|_2=\sqrt{14.1153}=3.757$

★ $\|A\|_{m_1}$、$\|A\|_F$、$\|A\|_{m_\infty}$ 都不是算子范数；但 $\|A\|_{m_1}$ 和 $\|A\|_1$ 均与 $\|\vec{x}\|_1$ 相容，$\|A\|_F$ 和 $\|A\|_2$ 均与 $\|\vec{x}\|_2$ 相容，但算子范数 $\|A\|_1$、$\|A\|_2$ 分别是与 $\|\vec{x}\|_1$、$\|\vec{x}\|_2$ 相容的范数中值最小的一个.

定理 2-4 设 $A\in C^{m\times n}$，而 $U\in C^{m\times m}$ 与 $V\in C^{n\times n}$ 都是酉矩阵，则 $\|UA\|_F=\|A\|_F=\|AV\|_F$

推论 与 A 酉(或正交)相似的矩阵的 F-范数是相同的，即若 $B=U^HAU$，则 $\|B\|_F=\|A\|_F$，其中 $A\in C^{m\times n}$，U 为酉矩阵.

例： 设 $S\in R^{n\times n}$ 可逆，$\|\vec{\alpha}\|=\|S^{-1}\vec{\alpha}\|_1$ 是 R^n 中的向量范数，$\|A\|$ 是 $R^{n\times n}$ 中从属于向量范数 $\|\vec{\alpha}\|$ 的矩阵范数，试导出 $\|A\|$ 与矩阵的 1-范数之间的关系式.

解：根据从属范数的定义可得：

$$\|A\|=\max_{\vec{\alpha}\neq\vec{0}}\frac{\|A\vec{\alpha}\|}{\|\vec{\alpha}\|}=\max_{\vec{\alpha}\neq\vec{0}}\frac{\|S^{-1}(A\vec{\alpha})\|_1}{\|S^{-1}\vec{\alpha}\|_1}=\max_{\vec{\beta}\neq\vec{0}}\frac{\|(S^{-1}AS)\vec{\beta}\|_1}{\|\vec{\beta}\|_1}=\|S^{-1}AS\|_1$$

例： 设 $A=(a_{ij})_{n\times n}(n>1)$，判断实数 $\|A\|=\max\limits_{i,j}|a_{ij}|$ 是否构成 $C^{n\times n}$ 中的矩阵范数.

解：取 $A_0=\begin{pmatrix}1&1&\cdots&1\\0&0&\cdots&0\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&0\end{pmatrix},B_0=\begin{pmatrix}1&0&\cdots&0\\1&0&\cdots&0\\\cdots&\cdots&\cdots&\cdots\\1&0&\cdots&0\end{pmatrix}$，那么 $\|A_0\|=1,\|B_0\|=1$.

但是 $A_0B_0=\begin{pmatrix} n & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$，从而 $\|A_0B_0\|=n$． $\because n>1$， $\therefore \|A_0B_0\|>\|A_0\|\|B_0\|$，故矩阵乘法的相容性不成立．因此，$\|A\|$ 不能构成 $C^{n\times n}$ 中的矩阵范数．

例： 设列向量 $\vec{x},\vec{y}\in C^n, A\in C^{n\times n}$，证明：$\|A\|_2=\max\{|\vec{x}^T A\vec{y}|\,\big|\,\|\vec{x}\|_2=\|\vec{y}\|_2=1\}$．

证：当 $A=0$ 时，结论显然成立；当 $A\neq 0$ 时，若 $\|\vec{x}\|_2=\|\vec{y}\|_2=1$，则由 $Cauchy-Schwarz$ 不等式可得 $|\vec{x}^T A\vec{y}|=|\vec{x}^T(A\vec{y})|\le\|\vec{x}\|_2\|A\vec{y}\|_2\le\|\vec{x}\|_2\|A\|_2\|\vec{y}\|_2=\|A\|_2$

$\because \|A\|_2=\max\limits_{\|\vec{y}\|_2=1}\|A\vec{y}\|_2, \therefore \exists\vec{y_0}\in C^n, sb.\|\vec{y_0}\|_2=1$，且使 $\|A\vec{y_0}\|_2=\|A\|_2\neq 0$．令 $\vec{x_0}=\dfrac{A\vec{y_0}}{\|A\vec{y_0}\|_2}$，

则 $\|\vec{x_0}\|_2=1$，且有 $|\vec{x_0}^T A\vec{y_0}|=\left|\dfrac{(A\vec{y_0})^H}{\|A\vec{y_0}\|_2}(A\vec{y_0})\right|=\|A\vec{y_0}\|_2=\|A\|_2$，因此，结论成立．

例： 设 $A=(a_{ij})_{m\times n}$，列向量 $\vec{\alpha}\in C^n$，证明：矩阵范数 $\|A\|=\max\{m,n\}\cdot\max\limits_{i,j}|a_{ij}|$ 与向量的 2-范数和∞-范数都相容．

证：设 $\vec{\alpha}=(\xi_1,\xi_2,\cdots,\xi_n)^T$，则有 $\|A\vec{\alpha}\|_2^2=\sum\limits_{i=1}^m\left|\sum\limits_{j=1}^n a_{ij}\xi_j\right|^2\le\sum\limits_{i=1}^m(\sum\limits_{j=1}^n|a_{ij}|^2\cdot\sum\limits_{j=1}^n|\xi_j|^2)$

$=(\sum\limits_{i=1}^m\sum\limits_{j=1}^n|a_{ij}|^2)\sum\limits_{j=1}^n|\xi_j|^2\le mn\cdot\max\limits_{i,j}|a_{ij}|^2\cdot\|\vec{\alpha}\|_2^2\le(\max\{m,n\})^2(\max\limits_{i,j}|a_{ij}|)^2\|\vec{\alpha}\|_2^2=\|A\|^2\|\vec{\alpha}\|_2^2$

即有 $\|A\vec{\alpha}\|_2\le\|A\|\|\vec{\alpha}\|_2$．

$\|A\vec{\alpha}\|_\infty=\max\limits_i\left|\sum\limits_{k=1}^n a_{ik}\xi_k\right|\le\max\limits_i\sum\limits_{k=1}^n|a_{ik}||\xi_k|\le(\max\limits_i\sum\limits_{k=1}^n|a_{ik}|)\cdot\max\limits_j|\xi_j|\le(n\cdot\max\limits_{i,j}|a_{ij}|)\|\vec{\alpha}\|_\infty$

$\le(\max\{m,n\}\cdot\max\limits_{i,j}|a_{ij}|)\|\vec{\alpha}\|_\infty=\|A\|\|\vec{\alpha}\|_\infty$

例： 设 $A,B\in C^{n\times n}$，证明：$\|AB\|_F\le\min\{\|A\|_2\|B\|_F,\|A\|_F\|B\|_2\}$．

证：设 $B=(\vec{\beta_1},\vec{\beta_2},\cdots,\vec{\beta_n})$，则由 $\|A\vec{\beta_j}\|_2\le\|A\|_2\|\vec{\beta_j}\|_2$ 可得

$$\|AB\|_F^2=\|A\overrightarrow{\beta_1}\|_2^2+\|A\overrightarrow{\beta_2}\|_2^2+\cdots+\|A\overrightarrow{\beta_n}\|_2^2\le\|A\|_2^2(\|\overrightarrow{\beta_1}\|_2^2+\|\overrightarrow{\beta_2}\|_2^2+\cdots+\|\overrightarrow{\beta_n}\|_2^2)=\|A\|_2^2\|B\|_F^2$$

即有 $\|AB\|_F\le\|A\|_2\|B\|_F$，

另外 $\|AB\|_F=\|(AB)^H\|_F=\|B^HA^H\|_F\le\|B^H\|_2\|A^H\|_F=\|B\|_2\|A\|_F=\|A\|_F\|B\|_2$

综上可得 $\|AB\|_F\le\min\{\|A\|_2\|B\|_F,\|A\|_F\|B\|_2\}$

2.3　谱范数的性质和谱半径

定理 2-5　设 $A\in C^{m\times n}$，则 1. $\|A\|_2=\max\limits_{\|\vec{x}\|_2=\|\vec{y}\|_2=1}\left|\vec{y}^H A\vec{x}\right|,\vec{x}\in C^n,\vec{y}\in C^m$；2. $\|A^H\|_2=\|A\|_2$；3. $\|A^HA\|_2=\|A\|_2^2$.

定理 2-6　设 $A\in C^{m\times n},U\in C^{m\times m},V\in C^{n\times n}$ 且 $U^HU=I_m,V^HV=I_n$，则 $\|UAV\|_2=\|A\|_2$

定理 2-7　若 $\|A\|<1$，则 $I\pm A$ 为非奇异，且 $\|(I\pm A)^{-1}\|\le(1-\|A\|)^{-1}$.

定义 2-3　设 $A\in C^{n\times n},\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 A 的特征值，我们称 $\rho(A)=\max\limits_i|\lambda_i|$ 为 A 的**谱半径**.

定理 2-8　（特征值上界定理）$\forall A\in C^{n\times n}$，总有 $\rho(A)\le\|A\|$，即 A 的谱半径不会超过 A 的任何一种范数.

定理 2-9　如果 $A\in C^{n\times n}$，且 A 为正规矩阵，则 $\rho(A)=\|A\|_2$.

定理 2-10　对任意非奇异矩阵 $A\in C^{n\times n}$，则 A 的谱范数为 $\|A\|_2=\sqrt{\rho(A^HA)}=\sqrt{\rho(AA^H)}$.

例：　证明对向量 $\vec{x}=(x_1,x_2)^T$ 及 $h>0$，则这样构成的 $\|\vec{x}\|=\max(|x_1|,\dfrac{|x_2-x_1|}{h})$ 是 $\vec{x}$ 的范

数，并求从属于$\vec{x}$的矩阵范数.

解：当$\vec{x}=\vec{0}$时，$x_1=x_2=0$，则$\|\vec{x}\|=0$；当$\vec{x}\neq\vec{0}$时，则x_1,x_2中必有一个不为零，显然

$\max(|x_1|,\dfrac{|x_2-x_1|}{h})>0$，则$\|\vec{x}\|>0$.

$\forall k\in R$，$\|k\vec{x}\|=\max(|kx_1|,\dfrac{|kx_2-kx_1|}{h})=|k|\max(|x_1|,\dfrac{|x_2-x_1|}{h})=|k|\|\vec{x}\|$

$\forall\vec{y}=(y_1,y_2)^T$，

$$\|\vec{x}+\vec{y}\|=\max(|x_1+y_1|,\frac{|(x_2+y_2)-(x_1+y_1)|}{h})\le\max(|x_1|+|y_1|,\frac{|x_2-x_1|}{h}+\frac{|y_2-y_1|}{h})$$

$=\max(|x_1|,\dfrac{|x_2-x_1|}{h})+\max(|y_1|,\dfrac{|y_2-y_1|}{h})=\|\vec{x}\|+\|\vec{y}\|$，所以$\|\vec{x}\|$是$\vec{x}$的范数.

令$P=\begin{pmatrix}1&0\\-\dfrac{1}{h}&\dfrac{1}{h}\end{pmatrix},\vec{\beta}=P\vec{x}=\begin{pmatrix}1&0\\-\dfrac{1}{h}&\dfrac{1}{h}\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}x_1\\\dfrac{x_2-x_1}{h}\end{pmatrix}$，那么

$\|\vec{\beta}\|_\infty=\max(|x_1|,\left|\dfrac{x_2-x_1}{h}\right|)=\max(|x_1|,\dfrac{|x_2-x_1|}{h})=\|\vec{x}\|$，根据算子范数的定义可知

$$\|A\|=\max_{\vec{x}\neq\vec{0}}\frac{\|A\vec{x}\|}{\|\vec{x}\|}=\max_{\vec{x}\neq\vec{0}}\frac{\|PA\vec{x}\|_\infty}{\|P\vec{x}\|_\infty}=\max_{\vec{x}\neq\vec{0}}\frac{\|PAP^{-1}\vec{\beta}\|_\infty}{\|\vec{\beta}\|_\infty}=\|PAP^{-1}\|_\infty.$$

例： 设$A\in C^{n\times n}$，且有某种范数$\|\cdot\|$使得$\|A\|<1$，证明：

(1) 矩阵$I-A$可逆；(2) $\|(I-A)^{-1}\|\le\dfrac{\|I\|}{1-\|A\|}$；(3) $\|I-(I-A)^{-1}\|\le\dfrac{\|A\|}{1-\|A\|}$.

证：（1）假设$I-A$不可逆，即$\det(I-A)=0$，则$\lambda=1$是A的一个特征值，从而有$\rho(A)\ge1$. 这与$\rho(A)\le\|A\|<1$矛盾，故假设不成立，即矩阵$I-A$可逆.

（2）$\because$矩阵$I-A$可逆，$\therefore(I-A)^{-1}(I-A)=I$，$\Rightarrow(I-A)^{-1}=I+(I-A)^{-1}A\Rightarrow$

$\|(I-A)^{-1}\|=\|I+(I-A)^{-1}A\|\le\|I\|+\|(I-A)^{-1}\|\|A\|$，$\therefore\|(I-A)^{-1}\|\le\dfrac{\|I\|}{1-\|A\|}$.

（3）在 $(I-A)-I=-A$ 两端右乘 $(I-A)^{-1}$ 可得 $I-(I-A)^{-1}=-A(I-A)^{-1}$，两端再左乘 A 得 $A-A(I-A)^{-1}=-A[A(I-A)^{-1}]$，$\therefore A(I-A)^{-1}=A+A[A(I-A)^{-1}]\Rightarrow$

$$\|A(I-A)^{-1}\|=\|A+A[A(I-A)^{-1}]\|\le\|A\|+\|A\|\|A(I-A)^{-1}\|\Rightarrow\|A(I-A)^{-1}\|\le\frac{\|A\|}{1-\|A\|}$$

$$\Rightarrow\|I-(I-A)^{-1}\|=\|-A(I-A)^{-1}\|\le\frac{\|A\|}{1-\|A\|}$$

例： 设 $A\in R^{m\times n}$，且 $rank(A)=n$，证明：$\|A(A^TA)^{-1}A^T\|_2=1$

证：令 $B=A(A^TA)^{-1}A^T$，则

$B^T=(A(A^TA)^{-1}A^T)^T=A[(A^TA)^{-1}]^TA^T=A[(A^TA)^T]^{-1}A^T=A(A^TA)^{-1}A^T=B$，即 B 为正规矩阵，$\therefore\|B\|_2=\rho(B)$.

又 $B^2=[A(A^TA)^{-1}A^T][A(A^TA)^{-1}A^T]=[AA^{-1}(A^{-1})^TA^T][A(A^TA)^{-1}A^T]=IB=B$

$\therefore B(B-I)=0$，$\therefore|B||B-I|=0\Rightarrow|B|=0,or,|B-I|=0$，即 B 的特征值为 0 或 1．任取非零列向量 $\vec{x}\in R^n$，由 $rankA=n$ 知 $A\vec{x}\ne\vec{0}$．再由 $B(A\vec{x})=A(A^TA)^{-1}A^TA\vec{x}=1(A\vec{x})$，可得 1 是 B 的一个特征值，故而 $\|B\|_2=\rho(B)=1$.

例： 设单位向量 $\vec{y}\in R^n(n>1)$，令 $B=I-\vec{y}\vec{y}^T$，证明：(1) $\|B\|_2=1$；(2) $\forall\vec{x}\in R^n$，若 $B\vec{x}\ne\vec{x}$，则 $\|B\vec{x}\|_2<\|\vec{x}\|$.

证：（1）
$$\because B^2=(I-\vec{y}\vec{y}^T)(I-\vec{y}\vec{y}^T)=I-2I\vec{y}\vec{y}^T+\vec{y}\vec{y}^T\vec{y}\vec{y}^T=I-2\vec{y}\vec{y}^T+\vec{y}(\vec{y}^T\vec{y})\vec{y}^T$$
$$\overset{\vec{y}^T\vec{y}=1}{=}I-2\vec{y}\vec{y}^T+\vec{y}\vec{y}^T=I-\vec{y}\vec{y}^T=B$$

$\therefore B(B-I)=0$，$\therefore|B||B-I|=0\Rightarrow|B|=0,or,|B-I|=0$，即 B 的特征值为 0 或 1.

当 $n>1$ 时，$\det(I-B)=\det(\vec{y}\vec{y}^T)=0$，所以 1 是 B 的一个特征值，$\rho(B)=1$.

又 $\because B^T=(I-\vec{y}\vec{y}^T)^T=I-\vec{y}\vec{y}^T=B$，$\therefore\|B\|_2=\rho(B)=1$.

（2）由 $B\vec{x}\ne\vec{x}$ 知 $\vec{z}=\vec{x}-B\vec{x}\ne\vec{0}$，再由 $B^TB=B^2=B$ 可得

$$(\vec{z}, B\vec{x}) = (\vec{x}, B\vec{x}) - (B\vec{x}, B\vec{x}) = \vec{x}^T B\vec{x} - (B\vec{x})^T B\vec{x} = \vec{x}^T B\vec{x} - \vec{x}^T (B^T B)\vec{x} = 0$$

$$\therefore \left\|\vec{x}\right\|_2^2 = \left\|B\vec{x} + \vec{z}\right\|_2^2 \geq \left\|B\vec{x}\right\|_2^2 + \left\|\vec{z}\right\|_2^2 > \left\|B\vec{x}\right\|_2^2 \Rightarrow \left\|B\vec{x}\right\|_2 < \left\|\vec{x}\right\|$$

例： 设 $A \in C^{n\times n}$ 可逆，对于 $C^{n\times n}$ 中的某种矩阵范数 $\|\bullet\|$，选取 $X^{(0)} \in C^{n\times n}$ 使满足 $\left\|I - AX^{(0)}\right\| = q < 1$，证明：

(1) 迭代格式 $X^{(k+1)} = X^{(k)}(2I - AX^{(k)})$ 收敛；

(2) $\left\|A^{-1} - X^{(k)}\right\| \leq \dfrac{q^{2^k}}{1-q}\left\|X^{(0)}\right\|(k = 1, 2, \cdots)$.

证：不妨令 $E^{(0)} = I - AX^{(0)}$，则有

$$I - AX^{(k+1)} = I - AX^{(k)}(2I - AX^{(k)}) = I - 2IAX^{(k)} + (AX^{(k)})^2 = (I - AX^{(k)})^2$$

$$I - AX^{(k)} = (I - AX^{(k-1)})^2 = \cdots = (E^{(0)})^{2^k}$$

$$\Rightarrow A^{-1} - X^{(k)} = A^{-1}(E^{(0)})^{2^k} \Rightarrow A^{-1}(I - E^{(0)}) = X^{(0)} \Rightarrow A^{-1} = X^{(0)} + A^{-1}E^{(0)}$$

两边取范数可得：

$$\left\|A^{-1}\right\| = \left\|X^{(0)} + A^{-1}E^{(0)}\right\| \leq \left\|X^{(0)}\right\| + \left\|A^{-1}\right\|\left\|E^{(0)}\right\| = \left\|X^{(0)}\right\| + q\left\|A^{-1}\right\|$$

即 $\left\|A^{-1}\right\| \leq \dfrac{\left\|X^{(0)}\right\|}{1-q}$.

于是有 $\left\|A^{-1} - X^{(k)}\right\| = \left\|A^{-1}(E^{(0)})^{2^k}\right\| \leq \left\|A^{-1}\right\|\left\|E^{(0)}\right\|^{2^k} \leq \dfrac{q^{2^k}}{1-q}\left\|X^{(0)}\right\|(k = 1, 2, \cdots)$

因为 $q < 1$，所以 $\lim\limits_{k\to\infty}\left\|A^{-1} - X^{(k)}\right\| = 0$ 也就是 $\lim\limits_{k\to\infty} X^{(k)} = A^{-1}$

§3　矩阵的条件数

3.1　病态方程组与病态矩阵

定义 3-1　如果系数矩阵 A 或常数项 $\vec{b}$ 的微小变化，引起方程组 $A\vec{x} = \vec{b}$ 解的巨大变化，则称方程组为**病态方程组**，其系数矩阵 A 就叫做对于解方程组（或求逆）来说的**病态矩阵**；反之，方程组就称为**良态方程组**，A 称**良态矩阵**.

3.2 矩阵的条件数

定理 3-1 设 A 是非奇异矩阵，$A\vec{x}=\vec{b}\neq\vec{0}$，且 $A(\vec{x}+\delta\vec{x})=\vec{b}+\delta\vec{b}$，则

$$\frac{\|\delta\vec{x}\|}{\|\vec{x}\|}\leq\|A^{-1}\|\|A\|\frac{\|\delta\vec{b}\|}{\|\vec{b}\|}.$$

定义 3-2 设 A 是非奇异矩阵，称数 $cond(A)=\|A^{-1}\|_p\bullet\|A\|_p\ (p=1,2,\infty)$ 为矩阵 A 的**条件数**.

★ 通常使用的条件数有：

（1）$cond(A)_\infty=\|A^{-1}\|_\infty\bullet\|A\|_\infty$

（2）A 的谱条件数 $cond(A)_2=\|A^{-1}\|_2\bullet\|A\|_2=\sqrt{\dfrac{\lambda_{\max}(A^HA)}{\lambda_{\min}(A^HA)}}$，显然，当 A 是实对称矩阵时，$cond(A)_2=\dfrac{|\lambda_1|}{|\lambda_n|}$，其中 λ_1,λ_n 分别为矩阵 A 的按模最大和最小特征值.

★ 条件数的性质：

1° 对任何非奇异矩阵 A，都有 $cond(A)_p\geq1$；

2° 设 A 为非奇异矩阵，$k\neq0$（常数），则 $cond(kA)_p=cond(A)_p$；

3° 如果 A 为正交矩阵，则 $cond(A)_2=1$；如果 A 为非奇异矩阵，R 为正交矩阵，则 $cond(AR)_F=cond(RA)_F=cond(A)_F$.

本章测试题

一、设 $A=\begin{pmatrix}2\lambda & \lambda\\ 1 & 1\end{pmatrix}$，证明：当 $\lambda=-\dfrac{2}{3}$ 时，$cond(A)_{\infty}$ 有最小值.

二、设 $A\in C^{n\times n}$ 可逆，证明：对于 $C^{n\times n}$ 中的任何矩阵范数 $\|\bullet\|$，都有 $\left\|A^{-1}\right\|\geq\dfrac{1}{\|A\|}$.

三、设 $A,B\in C^{n\times n}$ 都是可逆矩阵，且其次线性方程组 $(A+B)\vec{x}=\vec{0}$ 有非零解，证明：对于 $C^{n\times n}$ 中的任何矩阵范数 $\|\bullet\|$，都有 $\left\|A^{-1}B\right\|\geq 1$ 及 $\left\|AB^{-1}\right\|\geq 1$.

四、设 $A\in C^{n\times n}$，证明：$\rho(A)<1$ 的充要条件是，存在某种矩阵范数 $\|\bullet\|$，使得 $\|A\|<1$.

五、举例说明 $C^{n\times n}(n>1)$ 中的矩阵范数 $\|A\|_1$ 与 C^n 中的向量范数 $\|\vec{\alpha}\|_2$ 不相容.

六、设列向量 $\vec{\alpha},\vec{\beta}\in C^n$，证明 $\left\|\vec{\alpha}\vec{\beta}^T\right\|_F=\|\vec{\alpha}\|_2\left\|\vec{\beta}\right\|_2$.

七、设 $A \in C^{n\times n}$，证明 $\|A\|_2 \le \|A\|_F$.

八、设列向量 $\vec{\alpha}, \vec{\beta} \in R^n$，证明：$\left\|\vec{\alpha}+\vec{\beta}\right\|_2 = \left\|\vec{\alpha}\right\|_2 + \left\|\vec{\beta}\right\|_2$ 的充要条件是 $\vec{\alpha}, \vec{\beta}$ 线性相关，且 $\vec{\alpha}^T \vec{\beta} \ge 0$.

九、已知 $A=\begin{pmatrix} 0 & 1 & 2 & 3 \\ 2 & 3 & 4 & 5 \\ 4 & 5 & 6 & 7 \\ 6 & 7 & 8 & 9 \end{pmatrix}$，则 $\begin{cases} \|A\|_1 = \_\_\_\_\_\_\_\_ \\ \|A\|_{m_\infty} = \_\_\_\_\_\_\_\_ \\ \|A\|_F = \_\_\_\_\_\_\_\_ \end{cases}$

十、设可逆矩阵 $S \in R^{n\times n}$，且知 $\left\|\vec{x}\right\|_S = \left\|S\vec{x}\right\|_2$ $(\vec{x} \in R^n)$ 是 R^n 中的向量范数.

1．若 $\left\|A\right\|_S$ 表示 $R^{n\times n}$ 中从属于向量范数 $\left\|\vec{x}\right\|_S$ 的矩阵范数，试导出 $\left\|A\right\|_S$ 与矩阵的 2-范数之间的关系式；

2．给定非零向量 $\vec{y} \in R^n$，证明：$\left\|\vec{x}\right\| = \left\|\vec{x}\vec{y}^T\right\|_S$ 是 R^n 中的向量范数.

十一、设 V^n 是实数域 R 上的线性空间，$\vec{x} \in V^n$ 在基（Ⅰ）：$\overrightarrow{x_1}, \overrightarrow{x_2}, \cdots, \overrightarrow{x_n}$ 下的坐标为 $\vec{\alpha} = (\xi_1, \xi_2, \cdots, \xi_n)^T$，$\left\|\bullet\right\|_2$ 表示 R^n 中向量的 2-范数.

1． 证明：$\left\|\vec{x}\right\| = \left\|\vec{\alpha}\right\|_2$ 是 V^n 中的向量范数；

2． 设 $\vec{x} \in V^n$ 在基（Ⅱ）：$\overrightarrow{y_1}, \overrightarrow{y_2}, \cdots, \overrightarrow{y_n}$ 下的坐标为 $\vec{\beta} = (\eta_1, \eta_2, \cdots, \eta_n)^T$，且由基（Ⅰ）改变为基（Ⅱ）的过渡矩阵为 C，证明：$\left\|\vec{\alpha}\right\|_2 = \left\|\vec{\beta}\right\|_2$ 的充要条件是 C 为正交矩阵.

第六章　矩阵分析

§1　向量序列和矩阵序列的极限

1.1　向量序列的极限

定义 1-1　设 $\overrightarrow{x^{(k)}},\vec{x}\in C^m,k=1,2\cdots\cdots$，若 $\left\|\overrightarrow{x^{(k)}}-\vec{x}\right\|\to 0(k\to+\infty)$，则称向量序列 $\{\overrightarrow{x^{(k)}}\}$ 收敛于 $\vec{x}$，或说向量 $\vec{x}$ 是向量序列 $\{\overrightarrow{x^{(k)}}\}$ 当 $k\to+\infty$ 时的**极限**，可记为 $\lim\limits_{k\to\infty}\overrightarrow{x^{(k)}}=\vec{x}$，或 $\overrightarrow{x^{(k)}}\to\vec{x}(k\to+\infty)$．由向量范数之间等价关系，在某一向量范数意义下收敛，在其它向量范数意义下也收敛．

定义 1-2　若线性赋范空间中任一收敛向量序列的极限均属于此线性赋范空间，则称此线性赋范空间为**完备的线性赋范空间**，或称为**巴拿赫** (Banach) **空间**．在巴拿赫空间中，柯西收敛原理成立．

柯西收敛原理：数列 $\{a_n\}$ 收敛的充分必要条件是：$\forall\varepsilon>0,\exists N\in Z^+$，当 $m,n>N$ 时，有 $|a_n-a_m|<\varepsilon$．

定理 1-1　设 $\overrightarrow{x^{(k)}}=(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)})^T,\vec{x}=(x_1,x_2,\cdots,x_n)\in C^n$，则向量序列 $\{\overrightarrow{x^{(k)}}\}$ 收敛于 $\vec{x}$ 的充要条件为：每一个坐标分量序列 $\{x_i^{(k)}\}$ 收敛于 x_i，即 $\lim\limits_{k\to\infty}x_i^{(k)}=x_i,(i=1,2,\cdots,n)$．

1.2　矩阵序列的极限

定义 1-3　设有矩阵序列 $\{A^{(k)}\}$，其中 $A^{(k)}=(a_{ij}^{(k)})\in C^{n\times n}$，且当 $k\to+\infty,a_{ij}^{(k)}\to a_{ij}$，则称 $\{A^{(k)}\}$ **收敛**，并把矩阵 $A=(a_{ij})$ 叫做 $\{A^{(k)}\}$ 的极限，或称 $\{A^{(k)}\}$ 收敛于 A，记为 $\lim\limits_{k\to+\infty}A^{(k)}=A$ 或 $A^{(k)}\to A$，不收敛的矩阵序列称为**发散**的．

$$\lim_{k\to+\infty} A^{(k)}=A \Leftrightarrow \lim_{k\to+\infty}\left\|A^{(k)}-A\right\|=0^{*}$$

★ 矩阵序列极限运算的性质：

1．设 $\lim_{k\to+\infty} A^{(k)}=A, \lim_{k\to+\infty} B^{(k)}=B$，则 $\lim_{k\to+\infty}(aA^{(k)}+bB^{(k)})=aA+bB,(a,b\in C)$；

2．设 $\lim_{k\to+\infty} A^{(k)}=A, \lim_{k\to+\infty} B^{(k)}=B, A^{(k)}, B^{(k)}\in C^{n\times n}$，则 $\lim_{k\to+\infty} A^{(k)}B^{(k)}=AB$；

$$\left\|A^{(k)}B^{(k)}-AB\right\|=\left\|A^{(k)}B^{(k)}-AB^{(k)}+AB^{(k)}-AB\right\|\le\left\|A^{(k)}-A\right\|\left\|B^{(k)}\right\|+\left\|A\right\|\left\|B^{(k)}-B\right\|$$

$A^{(k)}-A\to 0, B^{(k)}-B\to 0, B^{(k)}\to B$，$\therefore\left\|A^{(k)}-A\right\|\left\|B^{(k)}\right\|+\left\|A\right\|\left\|B^{(k)}-B\right\|\to 0$

$$\lim_{k\to+\infty}\left\|A^{(k)}B^{(k)}-AB\right\|=0$$

3．设 $\lim_{k\to+\infty} A^{(k)}=A$，且 $A^{(k)}(k=1,2,\cdots,n), A$ 均可逆，则 $\{(A^{(k)})^{-1}\}$ 也收敛，且 $\lim_{k\to+\infty}(A^{(k)})^{-1}=A^{-1}$.

由于 $A^{(k)}(A^{(k)})^{-1}=I$，两边取极限 $A\cdot\lim_{k\to+\infty}(A^{(k)})^{-1}=I$，$\therefore \lim_{k\to+\infty}(A^{(k)})^{-1}=A^{-1}$

定理 1-2　设有矩阵序列 $\{A^k\}: A, A^2,\cdots,A^k,\cdots$，则 $\lim_{k\to+\infty} A^k=0$ 的充分必要条件是矩阵 A 的所有特征值的模都小于 1，即 A 的谱半径小于 1：$\rho(A)<1$.

证：必要性．对 $C^{n\times n}$ 中的任意矩阵范数，有 $\rho^k(A)=\rho(A^k)\le\left\|A^k\right\|$，由于 $\lim_{k\to+\infty} A^k=0$，所以 $\lim_{k\to+\infty}\rho^k(A)=0$，故而 $\rho(A)<1$.

充分性．当 $\rho(A)<1$ 时，取足够小的 $\varepsilon>0$，使 $\rho(A)+\varepsilon<1$，则有 $\left\|A\right\|_\alpha\le\rho(A)+\varepsilon<1$，于是 $\left\|A^k\right\|_\alpha\le\left\|A\right\|_\alpha^k\to 0(k\to+\infty)$，故 $\lim_{k\to+\infty} A^k=0$

定理 1-3　若对于矩阵 A 的某一范数有 $\left\|A\right\|<1$，则 $\lim_{k\to+\infty} A^k=0$.

---

* 注意：$\lim_{k\to+\infty} A^{(k)}=A \to \lim_{k\to+\infty}\left\|A^{(k)}\right\|=\left\|A\right\|$，但反过来却不一定成立．例如：

$A^{(k)}=\begin{pmatrix}(-1)^k & \frac{1}{k}\\ 1 & 2\end{pmatrix}, A=\begin{pmatrix}1 & 0\\ 1 & 2\end{pmatrix}$，则有 $\lim_{k\to+\infty}\left\|A^{(k)}\right\|_F=\lim_{k\to+\infty}\sqrt{1+\frac{1}{k^2}+1+4}=\sqrt{6}=\left\|A\right\|_F$，

但此时，$\lim_{k\to+\infty} A^{(k)}$ 发散.

例： 设 $A^{(k)},B^{(k)}\in C^{m\times n},\alpha_k,\beta_k\in C$ 且 $\lim\limits_{k\to+\infty}A^{(k)}=A,\lim\limits_{k\to+\infty}B^{(k)}=B,\lim\limits_{k\to+\infty}\alpha_k=\alpha$，$\lim\limits_{k\to+\infty}\beta_k=\beta$，证明：$\lim\limits_{k\to+\infty}(\alpha_k A^{(k)}+\beta_k B^{(k)})=\alpha A+\beta B$

证：$\because \left\|(\alpha_k A^{(k)}+\beta_k B^{(k)})-(\alpha A+\beta B)\right\|=\left\|(\alpha_k A^{(k)}-\alpha A)+(\beta_k B^{(k)}-\beta B)\right\|$

$\le\left\|\alpha_k A^{(k)}-\alpha A\right\|+\left\|\beta_k B^{(k)}-\beta B\right\|=\left\|\alpha_k A^{(k)}-\alpha_k A+\alpha_k A-\alpha A\right\|+\left\|\beta_k B^{(k)}-\beta_k B+\beta_k B-\beta B\right\|$

$\le|\alpha_k|\left\|A^{(k)}-A\right\|+|\alpha_k-\alpha|\left\|A\right\|+|\beta_k|\left\|B^{(k)}-B\right\|+|\beta_k-\beta|\left\|A\right\|$

$\because \lim\limits_{k\to+\infty}A^{(k)}=A,\lim\limits_{k\to+\infty}B^{(k)}=B,\lim\limits_{k\to+\infty}\alpha_k=\alpha,\lim\limits_{k\to+\infty}\beta_k=\beta$

$\therefore \lim\limits_{k\to+\infty}\left\|A^{(k)}-A\right\|=0,\lim\limits_{k\to+\infty}\left\|B^{(k)}-B\right\|=0,\lim\limits_{k\to+\infty}\left\|\alpha_k-\alpha\right\|=0,\lim\limits_{k\to+\infty}\left\|\beta_k-\beta\right\|=0$

且$|\alpha_k|$和$|\beta_k|$有界，故 $\lim\limits_{k\to+\infty}\left\|(\alpha_k A^{(k)}+\beta_k B^{(k)})-(\alpha A+\beta B)\right\|=0$

$\therefore \lim\limits_{k\to+\infty}(\alpha_k A^{(k)}+\beta_k B^{(k)})=\alpha A+\beta B$

例： 下列矩阵是否收敛？为什么？*

$$(1)A=\begin{pmatrix}0.2&0.1&0.2\\0.5&0.5&0.4\\0.1&0.3&0.2\end{pmatrix};(2)A=\begin{pmatrix}\frac{1}{6}&-\frac{4}{3}\\-\frac{1}{3}&\frac{1}{6}\end{pmatrix}$$

解：（1）可求得$\left\|A\right\|_1=0.9<1$，从而 A 为收敛矩阵.

（2）可求得 A 的特征值为 $\lambda_1=\frac{5}{6},\lambda_2=-\frac{1}{2}$，于是 $\rho(A)=\frac{5}{6}<1$，从而 A 为收敛矩阵.

§2 矩阵级数与矩阵函数

2.1 矩阵级数

定义 2-1 设有矩阵序列 $A^{(1)},A^{(2)},\cdots,A^{(k)},\cdots$，其中 $A^{(k)}=(a_{ij}^{(k)})\in C^{n\times n}$，称无穷和 $A^{(0)}+A^{(1)}+A^{(2)}+\cdots+A^{(k)}+\cdots$ 为**矩阵级数**，记为$\sum\limits_{k=0}^{\infty}A^{(k)}$，$A^{(k)}$ 称为矩阵级数的**一般**

* 一般先验证是否有某一矩阵范数满足$\left\|A\right\|<1$，如果找不到这样的矩阵范数，则求 $\rho(A)$.

项，即有 $\sum_{k=0}^{+\infty} A^{(k)} = A^{(0)} + A^{(1)} + A^{(2)} + \cdots + A^{(k)} + \cdots$

定义 2-2 级数 $\sum_{k=0}^{\infty} A^{(k)}$ 前 $k+1$ 项的和 $S^{(k)} = A^{(0)} + A^{(1)} + A^{(2)} + \cdots + A^{(k)}$ 称为级数 $\sum_{k=0}^{\infty} A^{(k)}$ 的**部分和**，如果矩阵序列 $\{S^{(k)}\}$ 收敛，且有极限 S，即 $\lim_{k\to+\infty} S^{(k)} = S$，则称矩阵级数 $\sum_{k=0}^{\infty} A^{(k)}$ **收敛**，S 称为级数的**和**，记作 $S = \sum_{k=0}^{+\infty} A^{(k)}$．不收敛的矩阵级数称为**发散**的．

矩阵级数 $S = \sum_{k=0}^{+\infty} A^{(k)}$ 收敛的充要条件是 n^2 个数项级数 $\sum_{k=0}^{+\infty} a_{ij}^{(k)} (i, j = 1, 2, \cdots, n)$ 都收敛，且和为

$$S = \begin{pmatrix} \sum_{k=0}^{+\infty} a_{11}^{(k)} & \sum_{k=0}^{+\infty} a_{12}^{(k)} & \cdots & \sum_{k=0}^{+\infty} a_{1n}^{(k)} \\ \sum_{k=0}^{+\infty} a_{21}^{(k)} & \sum_{k=0}^{+\infty} a_{22}^{(k)} & \cdots & \sum_{k=0}^{+\infty} a_{2n}^{(k)} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{k=0}^{+\infty} a_{n1}^{(k)} & \sum_{k=0}^{+\infty} a_{n2}^{(k)} & \cdots & \sum_{k=0}^{+\infty} a_{nn}^{(k)} \end{pmatrix}.$$

★ 矩阵级数收敛的性质：

1．若 $\sum_{k=0}^{+\infty} A^{(k)}$ 收敛，则 $\lim_{k\to+\infty} A^{(k)} = 0$；

2．若 $\sum_{k=0}^{+\infty} A^{(k)} = S, \sum_{k=0}^{+\infty} B^{(k)} = S'$，则 $\sum_{k=0}^{+\infty} (A^{(k)} \pm B^{(k)}) = S \pm S'$；

3．若 $\sum_{k=0}^{+\infty} A^{(k)} = S$，则 $\sum_{k=0}^{+\infty} \mu A^{(k)} = \mu S, \mu \in C$．

定义 2-3 设矩阵级数 $\sum_{k=0}^{+\infty} A^{(k)} = A^{(0)} + A^{(1)} + A^{(2)} + \cdots + A^{(k)} + \cdots$，其中 $A^{(k)} = (a_{ij}^{(k)}) \in C^{n\times n}$，如果 n^2 个数项级数 $a_{ij}^{(0)} + a_{ij}^{(1)} + a_{ij}^{(2)} + \cdots + a_{ij}^{(k)} + \cdots (i, j = 1, 2, \cdots, n)$ 都绝对收敛，则称矩阵级数**绝对收敛**．

定理 2-1 矩阵级数 $\sum_{k=0}^{+\infty} A^{(k)}$ 为绝对收敛的充分必要条件是

$\sum_{k=0}^{+\infty}\left\|A^{(k)}\right\|=\left\|A^{(0)}\right\|+\left\|A^{(1)}\right\|+\left\|A^{(2)}\right\|+\cdots+\left\|A^{(k)}\right\|+\cdots$ 收敛，其中 $\left\|A^{(k)}\right\|$ 为 $A^{(k)}$ 的任何一种范数*.

定理 2-2 设有两个矩阵级数

$A^{(1)}+A^{(2)}+\cdots+A^{(k)}+\cdots, A^{(k)}\in C^{n\times n}$，$B^{(1)}+B^{(2)}+\cdots+B^{(k)}+\cdots, B^{(k)}\in C^{n\times n}$ 都绝对收敛，其和分别为 A,B，则将它们按项相乘后作成的矩阵级数、

$A^{(1)}B^{(1)}+(A^{(1)}B^{(2)}+A^{(2)}B^{(1)})+\cdots+(A^{(1)}B^{(k)}+A^{(2)}B^{(k-1)}+\cdots+A^{(k)}B^{(1)})+\cdots$ 绝对收敛，且具有和 AB.

性质 1° 设矩阵级数 $\sum_{k=0}^{+\infty}A^{(k)}$ 绝对收敛，则 (1) $\sum_{k=0}^{+\infty}A^{(k)}$ 收敛；(2) 级数 $\sum_{k=0}^{+\infty}A^{(k)}$ 在任意改变各项的次序后仍然收敛，且其和不变.

性质 2° 设 P,Q 为 n 阶非奇异矩阵，若级数 $\sum_{k=0}^{+\infty}A^{(k)}$ 收敛（或绝对收敛），则矩阵级数 $\sum_{k=0}^{+\infty}PA^{(k)}Q$ 收敛（或绝对收敛）.

定义 2-4 形如 $\sum_{k=0}^{+\infty}c_kA^k=c_0I+c_1A+c_2A^2+\cdots+c_kA^k+\cdots$ 的矩阵级数称为**矩阵幂级数**，其中 $c_i\in C, A\in C^{n\times n}$.

定理 2-3 若正项级数 $|c_0|\|I\|+\sum_{k=1}^{+\infty}|c_k|\|A\|^k$ 收敛，则矩阵幂级数

$c_0I+c_1A+c_2A^2+\cdots+c_kA^k+\cdots$ 绝对收敛，其中 $\|A\|$ 为矩阵 A 的某种范数.

推论 若矩阵 A 的某种范数 $\|A\|$ 在幂级数 $\sum_{k=0}^{+\infty}c_kz^k=c_0+c_1z+c_2z^2+\cdots+c_kz^k+\cdots$ 的收

---

* $\sum_{k=0}^{+\infty}A^{(k)}$ 绝对收敛 $\Leftrightarrow\sum_{k=0}^{+\infty}\left\|A^{(k)}\right\|$ 收敛.

敛圆内，则矩阵幂级数$\sum_{k=0}^{+\infty} c_k A^k$绝对收敛.

定理 2-4 设$A \in C^{n\times n}$，如果A的谱半径$\rho(A)$的值在纯量z的幂级数$\sum_{k=0}^{+\infty} c_k z^k$的收敛圆内，那么矩阵幂级数$\sum_{k=0}^{+\infty} c_k A^k$绝对收敛；如果$A$的特征值中有一个在幂级数$\sum_{k=0}^{+\infty} c_k z^k$的收敛圆外，则矩阵幂级数$\sum_{k=0}^{+\infty} c_k A^k$发散*.

例： 判断矩阵幂级数$\sum_{k=1}^{+\infty} \frac{1}{k^2}\begin{pmatrix} 1 & 7 \\ -1 & -3 \end{pmatrix}^k$的敛散性.

解：设$A=\begin{pmatrix} 1 & 7 \\ -1 & -3 \end{pmatrix}$，则$A$的特征值为$\lambda_1=\lambda_2=-2$，$\therefore \rho(A)=2$，又因为幂级数$\sum_{k=1}^{+\infty}\frac{x^k}{k^2}$的收敛半径为$r=\lim_{k\to+\infty}\left|\frac{c_k}{c_{k+1}}\right|=\lim_{k\to+\infty}\frac{(k+1)^2}{k^2}=1<\rho(A)$，$\therefore$矩阵幂级数$\sum_{k=1}^{+\infty} \frac{1}{k^2}\begin{pmatrix} 1 & 7 \\ -1 & -3 \end{pmatrix}^k$发散.

定理 2-5 矩阵幂级数$I+A+A^2+\cdots+A^k+\cdots$☆绝对收敛的充要条件是$A$的谱半径$\rho(A)<1$，且该级数的和为$(I-A)^{-1}$.

定理 2-6 设矩阵A的某种范数$\|A\|<1$，则对任何非负整数k，有

$$\left\|(I-A)^{-1}-(I+A+A^2+\cdots+A^k)\right\| \le \frac{\|A\|^{k+1}}{1-\|A\|}$$

例： 设$A\in C^{n\times n}$，且$\rho(A)<1$，证明：$\sum_{k=0}^{\infty} kA^k = A(I-A)^{-2}$

---

* $\rho(A)\le\|A\|<R=\lim_{k\to+\infty}\left|\frac{c_k}{c_{k+1}}\right|$.

☆ Neumann 级数$\sum_{k=0}^{+\infty} A^k$.

证明：由 $\rho(A)<1$ 知 $I-A$ 可逆.

$\because$

$$(\sum_{k=0}^{N}kA^k)(I-A)=(A+2A^2+3A^3+\cdots+NA^N)(I-A)=A+2A^2+3A^3+\cdots+NA^N$$

$$-(A^2+2A^3+\cdots+(N-1)A^N+NA^{N+1})=A+A^2+A^3+\cdots+A^N-NA^{N+1}$$

$$=\sum_{k=0}^{N}A^k-I-NA^{N+1}$$

上式两边左乘 $(I-A)^{-1}$ 得：

$$\sum_{k=0}^{N}kA^k=\sum_{k=0}^{N}A^k(I-A)^{-1}-(I-A)^{-1}-NA^{N+1}(I-A)^{-1}$$

由于存在 $\varepsilon>0$ 和矩阵范数 $\|\bullet\|_a$ 使得 $\|A\|_a\le\rho(A)+\varepsilon<1$，于是可得

$\|NA^{N+1}\|_a\le N\|A\|_a^{N+1}\to 0(N\to+\infty)$，即 $\lim\limits_{N\to+\infty}NA^{N+1}=0$.

$$\therefore\sum_{k=0}^{\infty}kA^k=\lim_{N\to+\infty}\sum_{k=0}^{N}kA^k=\lim_{N\to+\infty}(\sum_{k=0}^{N}A^k)(I-A)^{-1}-(I-A)^{-1}-\lim_{N\to+\infty}NA^{N+1}(I-A)^{-1}$$

$$=(I-A)^{-2}-(I-A)^{-1}=(I-(I-A))(I-A)^{-2}=A(I-A)^{-2}.$$

2.2 矩阵函数

一、矩阵函数 $e^A,\sin A,\cos A$ 的定义与性质

$$e^A=I+A+\frac{A^2}{2!}+\cdots+\frac{A^k}{k!}+\cdots=\sum_{k=0}^{\infty}\frac{A^k}{k!};$$

$$\sin A=A-\frac{A^3}{3!}+\cdots+(-1)^k\frac{A^{2k+1}}{(2k+1)!}+\cdots=\sum_{k=0}^{\infty}(-1)^k\frac{A^{2k+1}}{(2k+1)!};$$

$$\cos A=I-\frac{A^2}{2!}+\cdots+(-1)^k\frac{A^{2k}}{(2k)!}+\cdots=\sum_{k=0}^{\infty}(-1)^k\frac{A^{2k}}{(2k)!}.$$

定理 2-7 如果 $AB=BA$，则有 $e^A\bullet e^B=e^B\bullet e^A=e^{A+B}$ ★.

★ 如果 $AB=BA$，还有以下结论：$\cos(A+B)=\cos A\cos B-\sin A\sin B$；$\cos 2A=\cos^2 A-\sin^2 A$；$\sin(A+B)=\sin A\cos B+\cos A\sin B$；

推论 1 对任意矩阵 $A\in C^{n\times n}$，e^A 总是可逆的（非奇异的）且 $(e^A)^{-1}=e^{-A}$*.

推论 2 $(e^A)^m=e^{mA}$（m 为整数）.

二、矩阵函数值的求法

定理 2-8（凯莱-哈密尔顿定理） 设有方阵 $A=(a_{ij})_{n\times n}$，其特征多项式为 $f(\lambda)=|\lambda I-A|=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$，则有 $f(A)=0$.

定理 2-9 设方阵 $A=(a_{ij})_{n\times n}$ 的 n 个特征值为 $\lambda_i(i=1,2,\cdots,n)$，$\varphi(z)$ 为一多项式，则方阵多项式 $\varphi(A)$ 的 n 个特征值为 $\varphi(\lambda_i)(i=1,2,\cdots,n)$.

★ 假定 A 与对角矩阵相似，即存在非奇异矩阵 C 使 $C^{-1}AC=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$，于是有

$$C^{-1}A^mC=diag(\lambda_1^m,\lambda_2^m,\cdots,\lambda_n^m)\Rightarrow\begin{cases}e^A=C\bullet diag(e^{\lambda_1},e^{\lambda_2},\cdots,e^{\lambda_n})\bullet C^{-1}\\ \sin A=C\bullet diag(\sin\lambda_1,\sin\lambda_2,\cdots,\sin\lambda_n)\bullet C^{-1}\\ \cos A=C\bullet diag(\cos\lambda_1,\cos\lambda_2,\cdots,\cos\lambda_n)\bullet C^{-1}\end{cases}$$

定义 2-5 设 $A\in C^{n\times n}$，显然存在 A 的约当标准形分解 $A=TJT^{-1}$，如果函数 $f(z)$ 在各个特征值 $\lambda_i(i=1,2,\cdots,r)$ 出具有直到 n_i-1 阶导数（n_i 为初等因子的指数，$i=1,2,\cdots,r$），则定义矩阵函数为 $f(At)=T\bullet f(Jt)\bullet T^{-1}$，其中

$$f(Jt)=diag(f(J_1(\lambda_1t)),f(J_2(\lambda_2t)),\cdots,f(J_r(\lambda_rt)))$$

$\sin 2A=2\sin A\cos A$

- 但是 $\sin A$ 和 $\cos A$ 却不一定可逆，如取 $A=\begin{pmatrix}\pi&0\\0&\frac{\pi}{2}\end{pmatrix}$，则 $\sin A=\begin{pmatrix}\sin\pi&0\\0&\sin\frac{\pi}{2}\end{pmatrix}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$，

  $\cos A=\begin{pmatrix}\cos\pi&0\\0&\cos\frac{\pi}{2}\end{pmatrix}=\begin{pmatrix}-1&0\\0&0\end{pmatrix}$ 均不可逆.

- $e^{iA}=\cos A+i\sin A$；$\cos A=\frac{1}{2}(e^{iA}+e^{-iA});\sin A=\frac{1}{2i}(e^{iA}-e^{-iA})$；

  $\cos(-A)=\cos A$；$\sin(-A)=-\sin A$.

$$f(J_i(\lambda_i t))=\begin{pmatrix} f(\lambda_i t) & tf^{'}(\lambda_i t) & \dfrac{t^2}{2!}f^{''}(\lambda_i t) & \cdots & \dfrac{t^{n_i-1}}{(n_i-1)!}f^{(n_i-1)}(\lambda_i t) \\ & f(\lambda_i t) & tf^{'}(\lambda_i t) & \cdots & \dfrac{t^{n_i-2}}{(n_i-2)!}f^{(n_i-2)}(\lambda_i t) \\ & & \cdots & \cdots & \cdots \\ & & & \cdots & tf^{'}(\lambda_i t) \\ & & & & f(\lambda_i t) \end{pmatrix}.$$

例： 设 $A=\begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$，求 $e^{At},\sin At$.

解：$\det(\lambda I-A)=(\lambda-2)(\lambda-1)(\lambda+1)$，对应 $\lambda_1=2$ 的特征向量为 $(1,0,0)^T$，而对应 $\lambda_2=1$ 的特征向量为 $(-1,1,1)^T$，对应 $\lambda_3=-1$ 的特征向量为 $(1,-3,3)^T$，所以 A 为单纯矩阵，可以化为对角矩阵，于是得 $T=\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -3 \\ 0 & 1 & 3 \end{pmatrix},T^{-1}=\begin{pmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{6} & \frac{1}{6} \end{pmatrix}$.

$$e^{At}=T\begin{pmatrix} e^{2t} & & \\ & e^{t} & \\ & & e^{-t} \end{pmatrix}T^{-1}=\begin{pmatrix} e^{2t} & \frac{1}{6}(4e^{2t}-3e^{t}-e^{-t}) & \frac{1}{6}(2e^{2t}-3e^{t}+e^{-t}) \\ 0 & cht & sht \\ 0 & sht & cht \end{pmatrix}$$

$$\sin At=T\begin{pmatrix} \sin(2t) & & \\ & \sin t & \\ & & \sin(-t) \end{pmatrix}T^{-1}=\frac{1}{6}\begin{pmatrix} \sin 2t & 4\sin 2t-2\sin t & 2\sin 2t-4\sin t \\ 0 & 0 & 6\sin t \\ 0 & 6\sin t & 0 \end{pmatrix}$$

§3　矩阵的微积分法

3.1　函数矩阵对实变量的导数

定义 3-1　若矩阵 $A=(a_{ij})$ 的诸元素 a_{ij} 均是变量 t 的函数，即

$$A(t)=\begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}(t) & a_{m2}(t) & \cdots & a_{mn}(t) \end{pmatrix}$$

则称 $A(t)$ 为**函数矩阵**，推而广之，变量 t 还可以是**向量**，也可以是**矩阵**.

如果所有元素 $a_{ij}(t)$ 在 $t=t_0$ 时存在极限，即 $\lim\limits_{t\to t_0} a_{ij}(t)=a_{ij}$（$a_{ij}$ 为一常数），则称函数矩阵 $A(t)$ 也有极限，且极限值为 A，即 $\lim\limits_{t\to t_0} A(t)=A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$（常量矩阵）.

若 $\lim\limits_{t\to t_0} A(t)=A, \lim\limits_{t\to t_0} B(t)=B$，则有 $\lim\limits_{t\to t_0}[A(t)+B(t)]=A+B$，$\lim\limits_{t\to t_0}[A(t)\bullet B(t)]=A\bullet B$，$\lim\limits_{t\to t_0} kA(t)=kA$（其中 A,B 为常量矩阵，k 为常数）.

定义 3-2 设 $A(t)=(a_{ij}(t))_{m\times n}$，若 $a_{ij}(t)(i=1,2,\cdots,m; j=1,2,\cdots,n)$ 在 $t=t_0$ 处（或 $[a,b]$ 上）可导，则称 $A(t)$ 在 $t=t_0$ 处（或 $[a,b]$ 上）可导，记为：

$$A^{'}(t_0)=\frac{dA(t)}{dt}\Big|_{t=t_0}=\lim_{\Delta t\to 0}\frac{A(t_0+\Delta t)-A(t_0)}{\Delta t}=\begin{pmatrix} a^{'}_{11}(t_0) & a^{'}_{12}(t_0) & \cdots & a^{'}_{1n}(t_0) \\ a^{'}_{21}(t_0) & a^{'}_{22}(t_0) & \cdots & a^{'}_{2n}(t_0) \\ \cdots & \cdots & \cdots & \cdots \\ a^{'}_{m1}(t_0) & a^{'}_{m2}(t_0) & \cdots & a^{'}_{mn}(t_0) \end{pmatrix}_{m\times n}$$

函数矩阵导数运算的性质：

1. $A(t)$ 是常数矩阵的充分必要条件是 $A^{'}(t)=0$；

2. 设 $A(t)=(a_{ij}(t))_{m\times n}$ 与 $B(t)=(b_{ij}(t))_{m\times n}$ 可导，则 $\frac{d}{dt}(A(t)\pm B(t))=A^{'}(t)\pm B^{'}(t)$；

3. 若 $k(t)$ 是可导的实函数，$A(t)$ 可导，则 $\frac{d}{dt}(k(t)A(t))=k^{'}(t)A(t)+k(t)A^{'}(t)$；

4. 设 $A(t),B(t)$ 都可导，则 $\frac{d}{dt}(A(t)B(t))=A^{'}(t)B(t)+A(t)B^{'}(t)$；

5. 若 $A(t),A^{-1}(t)$ 都有导数，则 $\frac{dA^{-1}(t)}{dt}=-A^{-1}(t)A^{'}(t)A^{-1}(t)$；

6. 设函数矩阵 $A(t)$ 是 t 的函数，而 $t=f(x)$ 是 x 的实值函数，且 $A(t)$ 与 $f(x)$ 均可导，则

有 $\frac{dA(t)}{dx}=\frac{dA(t)}{dt}f^{'}(x)=f^{'}(x)\frac{dA(t)}{dt}$；

7．函数矩阵对实变量的高阶导数：$\frac{d^k A(t)}{dt^k}=\frac{d}{dt}(\frac{d^{k-1}A(t)}{dt^{k-1}})(k=1,2,\cdots,n)$.

例： 设 $A(t)$ 是 m 阶可微矩阵，说明关系式 $\frac{d}{dt}[A(t)]^m=m[A(t)]^{m-1}\frac{d}{dt}A(t)$ 一般不成立. 问该式在什么条件下能成立？

解：$m=2$ 时，取 $A(t)=\begin{pmatrix} t^2 & t \\ 0 & t \end{pmatrix}$，则 $A^2(t)=\begin{pmatrix} t^4 & t^3+t^2 \\ 0 & t^2 \end{pmatrix}$，$\frac{d}{dt}A^2(t)=\begin{pmatrix} 4t^3 & 3t^2+2t \\ 0 & 2t \end{pmatrix}$

$2A(t)\frac{d}{dt}A(t)=\begin{pmatrix} 4t^3 & 2t^2+2t \\ 0 & 2t \end{pmatrix}$，可见 $\frac{d}{dt}[A(t)]^2\neq 2A(t)\frac{d}{dt}A(t)$.

当 $\frac{d}{dt}A(t)\bullet A(t)=A(t)\frac{d}{dt}A(t)$ 时，$\frac{d}{dt}[A(t)]^m=m[A(t)]^{m-1}\frac{d}{dt}A(t)$.

不论 A 是任何常量方阵，总有：$\frac{d}{dt}e^{At}=A\cdot e^{At}=e^{At}\cdot A$，

$\frac{d}{dt}\cos At=-A(\sin At)=-(\sin At)A$，$\frac{d}{dt}\sin At=A(\cos At)=(\cos At)A$

设 $A(t)=(a_{ij}(t))_{m\times n}$，则 $\frac{d}{dt}trA=tr\frac{dA}{dt}$，即迹 tr 和 $\frac{d}{dt}$ 两种运算，对方阵来说可以交换次序.

3.2 矩阵特殊的导数

一、数量函数对于向量的导数

定义 3-3 设 $\vec{x}=(x_1,x_2,\cdots,x_n)^T$，$f(\vec{x})=g(x_1,x_2,\cdots,x_n)$ 是以向量 $\vec{x}$ 为自变量的数量函数，即为 n 元函数，则规定数量函数 $f(\vec{x})$ 对于向量 $\vec{x}$ 的导数为 $\frac{df}{d\vec{x}}=(\frac{\partial g}{\partial x_1},\frac{\partial g}{\partial x_2},\cdots,\frac{\partial g}{\partial x_n})^T$，显然，若还有向量 $\vec{x}$ 的数量函数 $h(\vec{x})=h(x_1,x_2,\cdots,x_n)$，则下列导数法成立：

$$\frac{d[f(\vec{x})\pm h(\vec{x})]}{d\vec{x}}=\frac{df(\vec{x})}{d\vec{x}}\pm\frac{dh(\vec{x})}{d\vec{x}}$$

$$\frac{df(\vec{x})h(\vec{x})}{d\vec{x}}=\frac{df(\vec{x})}{d\vec{x}}h(\vec{x})+f(\vec{x})\frac{dh(\vec{x})}{d\vec{x}}$$

设 $A=(a_{ij}(t))_{n\times n}$ 为常量矩阵，$\vec{x}=(x_1,x_2,\cdots,x_n)^T$，，则数量函数 $f(\vec{x})=\vec{x}^T A\vec{x}$ 对于向量 $\vec{x}$ 的导数是 $\dfrac{df}{d\vec{x}}=(A+A^T)\vec{x}$.

1. A 是是对称矩阵时，$\dfrac{df}{d\vec{x}}=2A\vec{x}$；

2. 当 $A=I$ 时，$f(\vec{x})=x_1^2+x_2^2+\cdots+x_n^2$ 对 $\vec{x}$ 的导数是 $\dfrac{df}{d\vec{x}}=2\vec{x}$.

令 $\vec{x}=(\xi_1(t),\xi_2(t),\cdots,\xi_n(t))^T, f(\vec{x})=f(\xi_1,\xi_2,\cdots,\xi_n)$，则 $\dfrac{df}{dt}=(\dfrac{df}{d\vec{x}})^T\dfrac{d\vec{x}}{dt}$.

定义 3-4 设 $A\in R^{m\times n}, f(A)$ 为矩阵 A 的数量函数，即看成是 $m\times n$ 元函数，则规定数量函数 $f(A)$ 对于矩阵 A 的导数为 $\dfrac{df}{dA}=\left(\dfrac{\partial f}{\partial a_{ij}}\right)_{m\times n}=\begin{pmatrix}\dfrac{\partial f}{\partial a_{11}} & \cdots & \dfrac{\partial f}{\partial a_{1n}}\\ \cdots & \cdots & \cdots\\ \dfrac{\partial f}{\partial a_{m1}} & \cdots & \dfrac{\partial f}{\partial a_{mn}}\end{pmatrix}$.

设 $A\in R^{n\times n}$ 为实对称矩阵，则二次型 $\vec{x}^T A\vec{x}$ 对矩阵 A 的导数为 $\dfrac{d}{dA}(\vec{x}^T A\vec{x})=\vec{x}\vec{x}^T$.

设 $X=\begin{pmatrix}x & y\\ u & v\end{pmatrix}$，则 $\left|\dfrac{d|X|}{dX}\right|=|X|$.

设 $X=\begin{pmatrix}x_{11} & x_{12}\\ x_{21} & x_{22}\end{pmatrix}, f(X)=trX=x_{11}+x_{22}$，则 $\dfrac{df}{dX}=I$.

设 $X=(x_{ij})_{n\times n}$，则 $\dfrac{d}{dX}(tr(XX^T))=2X, \dfrac{d}{dX}(tr(X^2))=2X^T$.

二、矩阵对矩阵的导数

定义 3-5 设矩阵 F 是以 $A\in C^{m\times n}$ 为自变量的 $p\times q$ 阶矩阵，即

$$F(A)=\begin{pmatrix}f_{11}(A) & f_{12}(A) & \cdots & f_{1q}(A)\\ f_{21}(A) & f_{22}(A) & \cdots & f_{2q}(A)\\ \cdots & \cdots & \cdots & \cdots\\ f_{p1}(A) & f_{p2}(A) & \cdots & f_{pq}(A)\end{pmatrix}_{p\times q}$$

其元素 $f_{ks}(A)$ 是以矩阵 $A=(a_{ij})_{m\times n}$ 的元素为自变量的 mn 元函数，则规定矩阵 $F(A)$ 对于

矩阵 A 的导数为 $\frac{dF}{dA}=\left(\frac{\partial F}{\partial a_{ij}}\right)_{pm\times qn}=\begin{pmatrix} \frac{\partial F}{\partial a_{11}} & \frac{\partial F}{\partial a_{12}} & \cdots & \frac{\partial F}{\partial a_{1n}} \\ \frac{\partial F}{\partial a_{21}} & \frac{\partial F}{\partial a_{22}} & \cdots & \frac{\partial F}{\partial a_{2n}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F}{\partial a_{m1}} & \frac{\partial F}{\partial a_{m2}} & \cdots & \frac{\partial F}{\partial a_{mn}} \end{pmatrix}$，其中

$$\frac{\partial F}{\partial a_{ij}}=\begin{pmatrix} \frac{\partial f_{11}}{\partial a_{ij}} & \frac{\partial f_{12}}{\partial a_{ij}} & \cdots & \frac{\partial f_{1q}}{\partial a_{ij}} \\ \frac{\partial f_{21}}{\partial a_{ij}} & \frac{\partial f_{22}}{\partial a_{ij}} & \cdots & \frac{\partial f_{q}}{\partial a_{ij}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_{p1}}{\partial a_{ij}} & \frac{\partial f_{p2}}{\partial a_{ij}} & \cdots & \frac{\partial f_{pq}}{\partial a_{ij}} \end{pmatrix} i=1,2,\cdots,m; j=1,2,\cdots,n .$$

设 $\vec{x}=(x_1,x_2,\cdots,x_n)^T$，则 $\frac{d\vec{x}^T}{d\vec{x}}=I$.

雅可比 Jacobi 行列式： 设 $\vec{x}=(x_1,x_2,\cdots,x_n),\vec{y}=(y_1,y_2,\cdots,y_n)^T=f(\vec{x})$，其中

$$\begin{cases} y_1=f_1(x_1,x_2,\cdots,x_n)=f_1(\vec{x}) \\ y_2=f_2(x_1,x_2,\cdots,x_n)=f_2(\vec{x}) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n=f_n(x_1,x_2,\cdots,x_n)=f_n(\vec{x}) \end{cases}$$

则有 $\frac{df}{d\vec{x}}=(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_n})=\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$，$\det(f^{'}(\vec{x}))$ 称为雅可比行列式.

设 $A\in C^{m\times n}$，$\vec{y}\in C^n$ 是常量列向量，则 $\frac{d(\vec{y}^T A^T A\vec{y})}{dA}=2A\vec{y}\vec{y}^T$.

设 $f(\vec{x})$ 是向量 $\vec{x}=(x_1,x_2,\cdots,x_n)^T$ 的函数，而 $x_i=x_i(\vec{u}),i=1,2,\cdots,n,\vec{u}=(u_1,u_2,\cdots,u_n)^T$，则有 $\frac{df}{d\vec{u}}=\frac{d\vec{x}^T}{d\vec{u}}\frac{df}{d\vec{x}}$.

3.3 矩阵的全微分

定义 3-6 设矩阵 $F=(f_{ij})_{m\times n}$，则规定矩阵 A 的全微分为 $dF=(df_{ij})_{m\times n}$.

矩阵全微分的性质：

1. $d(F\pm G)=dF\pm dG$；2. $d(kF)=kdF$；3. 当 A 是常量矩阵时，$dA=0$；

4. $d(X^T)=(dX)^T$；5. $d(trX)=tr(dX)$.

定理 3-1 设 $\vec{x}=(x_1,x_2,\cdots,x_n)^T$，矩阵 $F=(f_{ij})_{s\times m}$，其中 f_{ij} 都是 x_i 的实函数，那么

$$dF=\sum_{i=1}^{n}\frac{\partial F}{\partial x_i}dx_i.$$

定理 3-2 1. 设 $A=BC$，则 $dA=(dB)C+B(dC)$；

2. 设 $A=A_1A_2\cdots A_r$，则 $dF=(dA_1)A_2\cdots A_r+A_1(dA_2)\cdots A_r+\cdots+A_1A_2\cdots A_{r-1}(dA_r)$.

推论 1. $d(\vec{\alpha}^T\vec{x})=\vec{\alpha}^T d\vec{x}=(d\vec{x})^T\vec{\alpha}$，其中 $\vec{\alpha}$ 为 $n\times 1$ 常量矩阵，$\vec{x}$ 是 n 维列向量；

2. $d(A\vec{x})=Ad\vec{x}$；

3. $d(\vec{x}^T A\vec{x})=\vec{x}^T(A+A^T)d\vec{x}$. 其中 A 是 $m\times n$ 常量矩阵，$\vec{x}$ 是 n 维列向量.

3.4 函数矩阵的积分

定义 3-7 设函数矩阵 $A(t)=\begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{pmatrix}$，我们定义

$$\int A(t)dt=\begin{pmatrix} \int a_{11}(t)dt & \int a_{12}(t)dt & \cdots & \int a_{1n}(t)dt \\ \int a_{21}(t)dt & \int a_{22}(t)dt & \cdots & \int a_{2n}(t) \\ \cdots & \cdots & \cdots & \cdots \\ \int a_{n1}(t)dt & \int a_{n2}(t)dt & \cdots & \int a_{nn}(t)dt \end{pmatrix}$$

$$\int_a^b A(t)dt=\begin{pmatrix}\int_a^b a_{11}(t)dt & \int_a^b a_{12}(t)dt & \cdots & \int_a^b a_{1n}(t)dt\\ \int_a^b a_{21}(t)dt & \int_a^b a_{22}(t)dt & \cdots & \int_a^b a_{2n}(t)\\ \cdots & \cdots & \cdots & \cdots\\ \int_a^b a_{n1}(t)dt & \int_a^b a_{n2}(t)dt & \cdots & \int_a^b a_{nn}(t)dt\end{pmatrix}$$

假设积分 $\int a_{ij}(t)dt(i,j=1,2,\cdots,n)$ 存在.

§4 矩阵分析的应用

对于一阶常系数非齐次线性微分方程组的初值问题：

$$\begin{cases}\dfrac{d\overrightarrow{x(t)}}{dt}=A\overrightarrow{x(t)}+\overrightarrow{b(t)}\\ \overrightarrow{x(t_0)}=\overrightarrow{x^{(0)}}\end{cases}$$

其中 $A\in R^{n\times n},\overrightarrow{x(t)}=(x_1(t),x_2(t),\cdots,x_n(t),)^T,\overrightarrow{b(t)}=(b_1(t),b_2(t),\cdots,b_n(t),)^T$，为求此初值问题的解，可将方程组改写为 $\dfrac{d\overrightarrow{x(t)}}{dt}-A\overrightarrow{x(t)}=\overrightarrow{b(t)}$，两边左乘 e^{-At} 得 $e^{-At}\dfrac{d\overrightarrow{x(t)}}{dt}-e^{-At}A\overrightarrow{x(t)}=e^{-At}\overrightarrow{b(t)}$ 或 $\dfrac{d}{dt}\left(e^{-At}\overrightarrow{x(t)}\right)=e^{-At}\overrightarrow{b(t)}$，在 $[t_0,t]$ 上对上式积分得 $e^{-At}\overrightarrow{x(t)}-e^{-At_0}\overrightarrow{x(t_0)}=\int_{t_0}^t e^{-A\tau}\overrightarrow{b(\tau)}d\tau$，故 $\overrightarrow{x(t)}=e^{A(t-t_0)}\overrightarrow{x^{(0)}}+e^{At}\int_{t_0}^t e^{-A\tau}\overrightarrow{b(\tau)}d\tau$，可见，求解一阶常系数非齐次线性微分方程组的关键是求出 e^{At}.

例： 求解微分方程组 $\begin{cases}\dfrac{dx_1}{dt}=3x_1+8x_3\\ \dfrac{dx_2}{dt}=3x_1-x_2+6x_3\\ \dfrac{dx_3}{dt}=-2x_1-5x_3\end{cases}$ 满足初始条件 $x_1(0)=x_2(0)=x_3(0)=1$ 的解.

解：令 $A=\begin{pmatrix}3&0&8\\3&-1&6\\-2&0&-5\end{pmatrix},\overrightarrow{x(t)}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},\overrightarrow{x(0)}=\begin{pmatrix}1\\1\\1\end{pmatrix}$，$|\lambda I-A|=(\lambda+1)^3$，矩阵 A 的特征值为 $\lambda_1=\lambda_2=\lambda_3=-1$ 而 A 的最小多项式为 $m_A(\lambda)=(\lambda+1)^2$．设 $g(\lambda)=c_0+c_1\lambda$，由

$\begin{cases} g(-1)=e^{-t}=c_0-c_1 \\ g'(-1)=te^{-t}=c_1 \end{cases}$，解得 $c_0=(1+t)e^{-t}, c_1=te^{-t}$，于是有

$$e^{At}=c_0I+c_1A=e^{-t}\begin{pmatrix}1+t&0&0\\0&1+t&0\\0&0&1+t\end{pmatrix}+e^{-t}\begin{pmatrix}3t&0&8t\\3t&-t&6t\\-2t&0&-5t\end{pmatrix}=e^{-t}\begin{pmatrix}1+4t&0&8t\\3t&1&6t\\-2t&0&1-4t\end{pmatrix}$$

$$\overrightarrow{x(t)}=e^{A(t-0)}\overrightarrow{x^{(0)}}=e^{At}\begin{pmatrix}1\\1\\1\end{pmatrix}=e^{-t}\begin{pmatrix}1+12t\\9t\\1-6t\end{pmatrix}.$$

例： 求解微分方程组 $\begin{cases}\dfrac{dx_1}{dt}=-2x_1+x_2+1\\ \dfrac{dx_2}{dt}=-4x_1+2x_2+2\\ \dfrac{dx_3}{dt}=x_1+x_3+e^t-1\end{cases}$ 满足初始条件 $x_1(0)=1,\ x_2(0)=1,$ $x_3(0)=-1$ 的解.

解：令 $A=\begin{pmatrix}-2&1&0\\-4&2&0\\1&0&1\end{pmatrix},\overrightarrow{b(t)}=\begin{pmatrix}1\\2\\e^t-1\end{pmatrix}\overrightarrow{x(t)}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},\overrightarrow{x(0)}=\begin{pmatrix}1\\1\\-1\end{pmatrix}$，可求得

$\det(\lambda I-A)=\lambda^2(\lambda-1)$．设 $g(\lambda)=c_0+c_1\lambda+c_2\lambda^2$，由 $\begin{cases}g(0)=1=c_0\\ g'(0)=t=c_1\\ g(1)=e^t=c_0+c_1+c_2\end{cases}$，解得

$\begin{cases}c_0=1\\ c_1=t\\ c_2=e^t-t-1\end{cases}$，于是 $e^{At}=c_0I+c_1A+c_2A^2=\begin{pmatrix}1-2t&t&0\\-4t&1+2t&0\\1+2t-e^t&e^t-t-1&e^t\end{pmatrix}$，依次计算

$e^{At}\overrightarrow{x^{(0)}}=\begin{pmatrix}1-t\\1-2t\\t-e^t\end{pmatrix}$，$e^{-A\tau}\overrightarrow{b(\tau)}=\begin{pmatrix}1+2\tau&-\tau&0\\4\tau&1-2\tau&0\\1-2\tau-e^{-\tau}&e^{-\tau}+\tau-1&e^{-\tau}\end{pmatrix}\begin{pmatrix}1\\2\\e^\tau-1\end{pmatrix}=\begin{pmatrix}1\\2\\0\end{pmatrix}$，所以

$$e^{At}\int_{t_0}^{t}e^{-A\tau}\overrightarrow{b(\tau)}d\tau=e^{At}\begin{pmatrix}t\\2t\\0\end{pmatrix}=\begin{pmatrix}t\\2t\\te^t-t\end{pmatrix}$$

故 $\overrightarrow{x(t)}=e^{A(t-t_0)}\overrightarrow{x^{(0)}}+e^{At}\int_{t_0}^{t}e^{-A\tau}\overrightarrow{b(\tau)}d\tau=\begin{pmatrix}1\\1\\(t-1)e^t\end{pmatrix}.$

例： 给定 $A\in R^{m\times n}$ 且 $rankA=n,\vec{b}\in R^{n},B\in R^{k\times n},\vec{d}\in R^{k}$，且 $B\vec{x}=\vec{d}$ 有解，试求约束极小化问题 $\min\limits_{B\vec{x}=\vec{d}}\left\|A\vec{x}-\vec{b}\right\|_2^2$ 的解，也就是函数 $f(\vec{x})=\left\|A\vec{x}-\vec{b}\right\|_2^2$ 在约束条件 $B\vec{x}=\vec{d}$ 下的极小点和极小值.

解：引入 $Lagrange$ 乘子 $2\vec{u}\in R^{k}$，化成等价的无约束极限问题．令

$g(\vec{x},\vec{u})=\left\|A\vec{x}-\vec{b}\right\|_2^2+2\vec{u}^T(B\vec{x}-\vec{d})$，若 $(\overrightarrow{x^{(0)}},\overrightarrow{u^{(0)}})$ 为 $g(\vec{x},\vec{u})$ 的极值点，则应有：

$$\begin{cases}\left.\dfrac{\partial g}{\partial\vec{x}}\right|_{(\overrightarrow{x^{(0)}},\overrightarrow{u^{(0)}})}=2A^TA\overrightarrow{x^{(0)}}-2A^T\vec{b}+2B^T\overrightarrow{u^{(0)}}=\vec{0}\\[2ex]\left.\dfrac{\partial g}{\partial\vec{u}}\right|_{(\overrightarrow{x^{(0)}},\overrightarrow{u^{(0)}})}=2(B\overrightarrow{x^{(0)}}-\vec{d})=\vec{0}\end{cases}\ {}^{*}$$

这说明极值点 $(\overrightarrow{x^{(0)}},\overrightarrow{u^{(0)}})$ 应满足方程 $\begin{pmatrix}A^TA & B^T\\ B & O\end{pmatrix}\begin{pmatrix}\overrightarrow{x^{(0)}}\\ \overrightarrow{u^{(0)}}\end{pmatrix}=\begin{pmatrix}A^T\vec{b}\\ \vec{d}\end{pmatrix}$，注意到 $\dfrac{d^2f}{d\vec{x}^2}=2A^TA$ 为正定矩阵，故 $\overrightarrow{x^{(0)}}$ 必为 $f(\vec{x})$ 的极小点.

上式两端左乘矩阵 $\begin{pmatrix}I_n & O\\ -B(A^TA)^{-1} & I_k\end{pmatrix}$ 可得

$\begin{pmatrix}A^TA & B^T\\ O & -B(A^TA)^{-1}B^T\end{pmatrix}\begin{pmatrix}\overrightarrow{x^{(0)}}\\ \overrightarrow{u^{(0)}}\end{pmatrix}=\begin{pmatrix}A^T\vec{b}\\ \vec{d}-B(A^TA)^{-1}A^T\vec{b}\end{pmatrix}$，解该方程组便得：

$$\overrightarrow{u^{(0)}}=(B(A^TA)^{-1}B^T)^{(1)}(B(A^TA)^{-1}A^T\vec{b}-\vec{d})$$
$$\overrightarrow{x^{(0)}}=(A^TA)^{-1}(A^T\vec{b}-B^T\overrightarrow{u^{(0)}})$$
$$f(\overrightarrow{x^{(0)}})=f((A^TA)^{-1}(A^T\vec{b}-B^T\overrightarrow{u^{(0)}})$$

其中 $(B(A^TA)^{-1}B^T)^{(1)}$ 为矩阵 $B(A^TA)^{-1}B^T$ 的 $\{1\}$-逆.

---

* $f(\vec{x})=\left\|A\vec{x}-\vec{b}\right\|_2^2=(A\vec{x}-\vec{b})^T(A\vec{x}-\vec{b})=\vec{x}^TA^TA\vec{x}-\vec{x}^TA^T\vec{b}-\vec{b}^TA\vec{x}+\vec{b}^T\vec{b}$

$\dfrac{df}{d\vec{x}}=2A^TA\vec{x}+A^T\vec{b}-(\vec{b}^TA)^T=2A^TA\vec{x}-2A^T\vec{b}$

本章测试题

一、矩阵 $A=\frac{1}{10}\begin{pmatrix}1 & 2\\ 8 & 1\end{pmatrix}$ 是否为收敛矩阵？为什么？

二、计算矩阵幂级数 $\sum_{k=0}^{\infty}\frac{1}{2^k}\begin{pmatrix}2 & -\frac{1}{2}\\ 2 & 0\end{pmatrix}^k$.

三、设 $A\in R^{n\times n}$，证明：$\begin{matrix}\sin^2 A+\cos^2 A=I,\sin(A+2\pi iI)=\sin A\\ \cos(A+2\pi iI)=\cos A,e^{A\pm 2\pi iI}=e^A\end{matrix}$.

四、设 $A\in R^{n\times n}$，证明：(1) $f(A^T)=[f(A)]^T, f(A^H)=[f(A)]^H$；(2) $e^I=eI$；

(3) 若 A 为实反对称矩阵(即 $A^T=-A$)，则 e^A 为正交矩阵.

五、设矩阵 $A=\begin{pmatrix}2&2&1\\1&3&1\\1&2&2\end{pmatrix}$，试求 $e^{At},\sin At,\cos At$.

六、设矩阵 $A=\begin{pmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&0&0\\0&0&0&2&0\\0&0&0&1&2\end{pmatrix}$，试求 $e^{At},\sin At,\cos At$.

七、设 $C\in R^{n\times n}$，证明：$\begin{matrix}\dfrac{d}{dt}\sin At=A\cos At=(\cos At)A\\ \dfrac{d}{dt}\cos At=-A\sin At=-(\sin At)A\end{matrix}$.

八、设函数矩阵 $A(t)=\begin{pmatrix}1 & t^2\\ t & 0\end{pmatrix}$，试求 $\frac{d}{dt}[A^{-1}(t)](t\neq 0)$.

九、已知 $e^{At}=\begin{pmatrix}2e^{2t}-e^{t} & e^{2t}-e^{t} & e^{t}-e^{2t}\\ e^{2t}-e^{t} & 2e^{2t}-e^{t} & e^{t}-e^{2t}\\ 3e^{2t}-3e^{t} & 3e^{2t}-3e^{t} & 3e^{t}-2e^{2t}\end{pmatrix}$，求 A .

十、设 A 是 n 阶可逆矩阵，证明：$\int_0^t e^{A\tau}d\tau=A^{-1}e^{At}-A^{-1}$.

十一、设 A 是可逆矩阵，证明：$\overrightarrow{x(t)}=(\cos At)\overrightarrow{x^{(0)}}+(\sin At)A^{-1}\overrightarrow{x^{(1)}}$ 是微分方程组 $\dfrac{d^2\overrightarrow{x(t)}}{dt^2}+A^2\overrightarrow{x(t)}=\vec{0}$ 满足初始条件 $\overrightarrow{x(0)}=\overrightarrow{x^{(0)}},\dfrac{d\overrightarrow{x(0)}}{dt}=\overrightarrow{x^{(1)}}$ 的解.

十二、已知 $A=\begin{pmatrix}0.1 & 0.3\\0.7 & 0.6\end{pmatrix}$，判断矩阵幂级数 $\sum_{k=0}^{\infty}A^k$ 收敛的依据是____________________，幂级数的和是______________________________.

十三、已知 $A=\begin{pmatrix}2&1&0&0\\0&2&0&0\\0&0&0&0\\0&0&1&0\end{pmatrix},\overrightarrow{b(t)}=\begin{pmatrix}2\\0\\1\\t\end{pmatrix},\overrightarrow{x(0)}=\begin{pmatrix}-1\\0\\0\\1\end{pmatrix}$，1．求 e^{At}；2．应用矩阵函数方法求微分方程组 $\overrightarrow{x'(t)}=A\overrightarrow{x(t)}+\overrightarrow{b(t)}$ 满足初始条件 $\overrightarrow{x(0)}$ 的解．

第七章　广义逆矩阵及其应用

§1　矩阵的几种广义逆

1.1　广义逆矩阵的基本概念

对于任意的复数矩阵 $A_{m\times n}$，如果存在复矩阵 $X_{n\times m}$，满足 $\begin{cases} AXA=A \\ XAX=X \\ (XA)^H=XA \\ (AX)^H=AX \end{cases}$，则称 X 为 A 的一个 $Moore-Penrose$ 广义逆，并把上面四个方程叫做 $Moore-Penrose$ 方程，简称 $M-P$ 方程.

定义 1-1　设 $A\in C^{m\times n}$，若有某个 $X\in C^{n\times m}$，满足 $M-P$ 方程的全部或一部分，则称 X 为 A 的广义逆矩阵.

1.2　减号逆 A^-

定义 1-2　设有 $m\times n$ 实矩阵 $A(m\le n)$．若有一个 $n\times m$ 实矩阵（记为 A^-）存在，使下式成立，则称 A^- 为 A 的减号逆或 g 逆：$AA^-A=A$．

1．当 A^- 为 A 的一个减号逆时，$(A^-)^T$ 就是 A^T 的一个减号逆.

2．若 $A=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n}$ 则 $A^-=\begin{pmatrix} I_r & * \\ * & * \end{pmatrix}_{n\times m}$，其中*任意选取.

引理　设 $B_{m\times n}=P_{m\times m}A_{m\times n}Q_{n\times n}$，其中 P,Q 都是满秩方阵，如果 B^- 存在，则 $A^-=QB^-P$．

定理 1-1　（存在性）任给 $m\times n$ 阶矩阵 A，那么 A^- 一定存在，但不唯一.

定理 1-2　设 $A\in R^{m\times n}$，则 $rankA^-\ge rankA$．

例： 求矩阵 $A=\begin{pmatrix}1&2&0\\0&0&2\\2&4&0\end{pmatrix}$ 的广义逆矩阵 A^{-}.

解：
$$\begin{pmatrix}A&I_3\\I_3&0_3\end{pmatrix}=\begin{pmatrix}1&2&0&1&0&0\\0&0&2&0&1&0\\2&4&0&0&0&1\\1&0&0&&&\\0&1&0&&O_{3\times3}&\\0&0&1&&&\end{pmatrix}\to\begin{pmatrix}1&0&0&1&0&0\\0&0&1&0&\frac{1}{2}&0\\0&0&0&-2&0&1\\1&-2&0&&&\\0&1&0&&O_{3\times3}&\\0&0&1&&&\end{pmatrix}\to$$

$$\begin{pmatrix}1&0&0&1&0&0\\0&1&0&0&\frac{1}{2}&0\\0&0&0&-2&0&1\\1&0&-2&&&\\0&0&1&&O_{3\times3}&\\0&1&0&&&\end{pmatrix}$$，则可得 $\begin{pmatrix}1&0&0\\0&\frac{1}{2}&0\\-2&0&1\end{pmatrix}A\begin{pmatrix}1&0&-2\\0&0&1\\0&1&0\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}$，即

$PAQ=\begin{pmatrix}I_2&0_{2\times1}\\0_{1\times2}&0\end{pmatrix}_{3\times3}=B$，但标准形 B 的减号逆为 $B^{-}=\begin{pmatrix}1&0&*\\0&1&*\\*&*&*\end{pmatrix}$，故得

$A^{-}=Q\begin{pmatrix}1&0&*\\0&1&*\\*&*&*\end{pmatrix}P$，$*$为任意实数.

1.3 自反广义逆

定义 1-3 对于一个 $m\times n$ 阶实矩阵 A，使 $AXA=A$ 及 $XAX=X$ 同时成立的 $n\times m$ 阶实矩阵 X，称为是 A 的一个**自反广义逆**，用 A_r^{-} 表示，即有 $AA_r^{-}A=A$ 及 $A_r^{-}AA_r^{-}=A_r^{-}$，显然，**自反广义逆是减号逆的一个子集**.

一、最大秩矩阵的右逆和左逆

定义 1-4 设 A 是行最大秩的 $m\times n$ 阶实矩阵 $(m\le n)$，如果存在一个 $n\times m$ 阶矩阵 G，当 G 右乘 A 后得到一个 $m\times m$ 阶单位矩阵 I，即 $AG=I$，则 G 叫做 A 的**右逆**，记作 A_R^{-1}，这就是说有 $AA_R^{-1}=I$．$\underline{A_R^{-1}=A^T(AA^T)^{-1}}$.

定义 1-5 设 A 是列最大秩的 $m\times n$ 阶实矩阵 $(m\geq n)$，如果存在一个 $n\times m$ 阶矩阵 G，当 G 左乘 A 后得到一个 $n\times n$ 阶单位矩阵 I，即 $GA=I$，则 G 叫做 A 的**左逆**，记作 A_L^{-1}，这就是说有 $A_L^{-1}A=I$．$\underline{A_L^{-1}=(A^TA)^{-1}A^T}$．

对于行最大秩或列最大秩的矩阵 A，一般情况下，A_R^{-1} 和 A_L^{-1} 不可能同时存在.

二、求 A_r^- 的最大秩分解方法

定理 1-3 设 $A=BC$ 为矩阵 A 的最大秩分解，则 A 的自反广义逆的一般形式为 $\underline{A_r^-=C_R^{-1}B_L^{-1}}$.

1.4 最小范数广义逆 A_m^-

定义 1-6 设 $A\in R^{m\times n}(m\leq n)$，如果有一个 $n\times m$ 阶矩阵 X，满足 $AXA=A$ 及 $(XA)^T=XA$，则称 X 是 A 的一个**最小范数广义逆**，记为 A_m^-.

定理 1-4 设 $A\in R^{m\times n}(m\leq n)$，$A_m^-=A^T(AA^T)^-$ 为 A 的一个最小范数广义逆.

定理 1-5 条件 $\begin{cases}AXA=A\\(XA)^T=XA\end{cases}$ 与 $XAA^T=A^T$ 等价.

1.5 最小二乘广义逆 A_l^-

定义 1-7 设 $A\in R^{m\times n}(m\leq n)$，若有一个 $n\times m$ 阶矩阵 X，满足 $AXA=A$ 及 $(AX)^T=AX$，则称 X 是 A 的一个**最小二乘广义逆**，记为 A_l^-.

定理 1-6 设 $A\in R^{m\times n}(m\leq n)$，则 $A_l^-=(A^TA)^-A^T$.

1.6 加号逆 A^+

定义 1-8 设 $A\in R^{m\times n}$，若有存在 $n\times m$ 阶矩阵 X，它同时满足 $\begin{cases} AXA=A \\ XAX=X \\ (XA)^H=XA \\ (AX)^H=AX \end{cases}$，则称 X 为 A 的**加号逆**，或**伪逆**，或 $\boldsymbol{Moore-Penrose}$ **逆**，记为 A^+.

1. $(A^+)^+=A$；

2. AA^+, A^+A 都是对称矩阵；

3. 设 $0\in R^{m\times n}$，则 $0^+=0\in R^{n\times m}$；

4. 设 $A=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{n\times n}$，则 $A^+=A$；

5. 设对角阵 $\Lambda=\begin{pmatrix} \lambda_1 & & \\ & \cdots & \\ & & \lambda_n \end{pmatrix}, \Lambda^+=\begin{pmatrix} \lambda_1^+ & & \\ & \cdots & \\ & & \lambda_n^+ \end{pmatrix}$，其中 $\lambda_i^+=\begin{cases} \dfrac{1}{\lambda_i}, \lambda_i\neq 0 \\ 0, \lambda_i=0 \end{cases}$.

定理 1-7 若 $A\in R^{m\times n}$，且 $A=BC$ 是最大秩分解，则 $\underline{X=C^T(CC^T)^{-1}(B^TB)^{-1}B^T}$ 是 A 的加号逆.

推论 设 $A\in R^{m\times n}, rankA=r$，则

1. 当 $r=n$ 时，（即 A 列满秩）$\underline{A^+=(A^TA)^{-1}A^T}$（左逆）；

2. 当 $r=m$ 时，（即 A 行满秩）$\underline{A^+=A^T(AA^T)^{-1}}$（右逆）.

定理 1-8 加号逆 A^+ 是唯一的.

推论 若 A 是 n 阶满秩方阵，即 A^{-1} 存在，则 $A^+=A^{-1}=A^-$.

定理 1-9 $\begin{cases} (A^T)^+=(A^+)^T \\ A^+=(A^TA)^+A^T=A^T(AA^T)^+ \\ (A^TA)^+=A^+(A^T)^+ \end{cases}$.

★ A^+ <u>的各种常用算法：</u>

1. 若 A 是满秩方阵，则 $A^+=A^{-1}$；

2．若 A 是对角方阵，即 $A=diag(d_1,d_2,\cdots,d_n),(d_1,d_2,\cdots,d_n\in R)$，

则 $A^+=(d_1^+,d_2^+,\cdots,d_n^+)$，其中 $d_i^+=\begin{cases}\dfrac{1}{d_i},d_i\neq 0\\0,d_i=0\end{cases}$；

3． 若 A 是行最大秩矩阵，则 $A^+=A^T(AA^T)^{-1}$；特别的，对 $\vec{\alpha}^T=(a_1,a_2,\cdots,a_n)$ 有

$\left(\vec{\alpha}^T\right)^+=\left(\vec{\alpha}^T\right)^T\left(\vec{\alpha}^T\left(\vec{\alpha}^T\right)^T\right)^{-1}=\vec{\alpha}\left(\vec{\alpha}^T\vec{\alpha}\right)^{-1}=\dfrac{\vec{\alpha}}{\vec{\alpha}^T\vec{\alpha}}=\dfrac{\vec{\alpha}}{\sum\limits_{i=1}^n a_i^2}$，它是一个 n 维列向量；

4． 若 A 是列最大秩矩阵，则 $A^+=(A^TA)^{-1}A^T$；特别的，对 $\vec{\beta}=(b_1,b_2,\cdots,b_n)^T$ 有

$\left(\vec{\beta}\right)^+=\left(\vec{\beta}^T\vec{\beta}\right)^{-1}\vec{\beta}^T=\dfrac{\vec{\beta}^T}{\vec{\beta}^T\vec{\beta}}=\dfrac{\vec{\beta}^T}{\sum\limits_{i=1}^n b_i^2}$，它是一个 n 维行向量；

5． 如果 $A\in R^{m\times n},rankA=r<\min(m,n)$，可以用最大秩分解法求 A^+，即当 $A=BC$ 为最大秩分解时，则 $A^+=C^+B^+$；

6． 设 $A\in R^{m\times n}$ 的奇异值分解为 $A=UDV^T$，其中 U 和 V 分别是 m 阶和 n 阶正交矩阵，而

$$D=\begin{pmatrix}\sigma_1 & & & 0\\ & \cdots & & \\ & & \sigma_r & \\ 0 & & & 0\end{pmatrix}=\begin{pmatrix}\sum & 0\\0 & 0\end{pmatrix}_{m\times n},\sum=diag(\sigma_1\cdots\sigma_r),r=rankA,\sigma_i>0(i=1,\cdots,r)$$

是 A 的奇异值，则 $A^+=VD^+U^T$．

例： 已知矩阵 $A=\begin{pmatrix}1&0&0&1\\1&1&0&0\\0&1&1&0\\0&0&1&1\end{pmatrix}$．(1) 作出 A 的最大秩分解 $A=BC$；(2) 利用最大秩分解，求出一个自反广义逆 A_r^-；(3) 利用 $A^+=C^+B^+$ 求 A^+．

解：$A=\begin{pmatrix}1&0&0&1\\1&1&0&0\\0&1&1&0\\0&0&1&1\end{pmatrix}\simeq\begin{pmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&1\\0&0&0&0\end{pmatrix}$，取 A 的前三列组成矩阵 $B=\begin{pmatrix}1&0&0\\1&1&0\\0&1&1\\0&0&1\end{pmatrix}$，再

取 A 的行标准型的前三个非零行，组成矩阵 $C=\begin{pmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&1\end{pmatrix}$，$A=BC$.

$$C_R^{-1}=C^T(CC^T)^{-1}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\1&-1&1\end{pmatrix}\left(\begin{pmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\1&-1&1\end{pmatrix}\right)^{-1}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\1&-1&1\end{pmatrix}\begin{pmatrix}2&-1&1\\-1&2&-1\\1&-1&2\end{pmatrix}^{-1}$$

$$=\frac{1}{4}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\1&-1&1\end{pmatrix}\begin{pmatrix}3&1&-1\\1&3&1\\-1&1&3\end{pmatrix}=\frac{1}{4}\begin{pmatrix}3&1&-1\\1&3&1\\-1&1&3\\1&-1&1\end{pmatrix}$$

$$B_L^{-1}=(B^TB)^{-1}B^T=\left(\begin{pmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\end{pmatrix}\begin{pmatrix}1&0&0\\1&1&0\\0&1&1\\0&0&1\end{pmatrix}\right)^{-1}\begin{pmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\end{pmatrix}=\begin{pmatrix}2&1&0\\1&2&1\\0&1&2\end{pmatrix}^{-1}\begin{pmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\end{pmatrix}$$

$$=\frac{1}{4}\begin{pmatrix}3&-2&1\\-2&4&-2\\1&-2&3\end{pmatrix}\begin{pmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\end{pmatrix}=\frac{1}{4}\begin{pmatrix}3&1&-1&1\\-2&2&2&-2\\1&-1&1&3\end{pmatrix}$$

则 $A_r^-=C_R^{-1}B_L^{-1}=\frac{1}{16}\begin{pmatrix}3&1&-1\\1&3&1\\-1&1&3\\1&-1&1\end{pmatrix}\begin{pmatrix}3&1&-1&1\\-2&2&2&-2\\1&-1&1&3\end{pmatrix}=\frac{1}{16}\begin{pmatrix}6&6&-2&-2\\-2&6&6&-2\\-2&-2&6&6\\6&-2&-2&6\end{pmatrix}$.

$rankC=3$，即 C 为行最大秩矩阵，$\therefore C^+=C_R^{-1}=\frac{1}{4}\begin{pmatrix}3&1&-1\\1&3&1\\-1&1&3\\1&-1&1\end{pmatrix}$

$B=\begin{pmatrix}1&0&0\\1&1&0\\0&1&1\\0&0&1\end{pmatrix}\sim\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\0&0&0\end{pmatrix}$，$rankB=3$，即 B 为列最大秩矩阵，$\therefore B^+=B_L^{-1}$

$$=\frac{1}{4}\begin{pmatrix} 3 & 1 & -1 & 1 \\ -2 & 2 & 2 & -2 \\ 1 & -1 & 1 & 3 \end{pmatrix}, \quad \therefore A^{+}=A_{r}^{-}=\frac{1}{16}\begin{pmatrix} 6 & 6 & -2 & -2 \\ -2 & 6 & 6 & -2 \\ -2 & -2 & 6 & 6 \\ 6 & -2 & -2 & 6 \end{pmatrix}.$$

§2　广义逆在解线性方程组中的应用

考虑非齐次线性方程组 $A\vec{X}=\vec{b}$，其中 $A\in C^{m\times n},\vec{b}\in C^{m}$ 给定，而 $\vec{X}\in C^{m}$ 为待定向量．若 $rank(A|\vec{b})=rankA$，则方程组 $A\vec{X}=\vec{b}$ 有解，或称方程组相容；若 $rank(A|\vec{b})\neq rankA$，则方程组 $A\vec{X}=\vec{b}$ 无解，或称方程组不相容或矛盾方程组.

定理 2-1　如果线性方程组 $A\vec{X}=\vec{b}$ 是相容的，A^{-} 是 A 的任意一个减号逆，则线性方程组 $A\vec{X}=\vec{b}$ 的一个特解可表示成 $\vec{X}=A^{-}\vec{b}$，而通解可以表示成 $\vec{X}=A^{-}\vec{b}+(I-A^{-}A)\vec{Z}$，其中 $\vec{Z}$ 是与 $\vec{X}$ 同维的任意向量.

定理 2-2　在相容线性方程组 $A\vec{X}=\vec{b}$ 的一切解中具有极小范数解的充要条件是 $\vec{X}=A_{m}^{-}\vec{b}$，A_{m}^{-} 是 A 的最小范数广义逆.

定理 2-3　相容线性方程组 $A\vec{X}=\vec{b}$ 具有唯一的极小范数解.

定理 2-4　不相容线性方程组 $A\vec{X}=\vec{b}$ 有最小二乘解的充要条件是 $\vec{X}=A_{l}^{-}\vec{b}$，A_{l}^{-} 是 A 的最小二乘广义逆.

加号逆 A^{+} 的应用

1．$A\vec{X}=\vec{b}$ 相容时，$\vec{X}=A^{+}\vec{b}+(I-A^{+}A)\vec{Z}$ 是通解，$\vec{X}=A^{+}\vec{b}$ 是极小范数解；

2．$A\vec{X}=\vec{b}$ 不相容时，$\vec{X}=A^{+}\vec{b}$ 是最小二乘解，$\vec{X}=A^{+}\vec{b}+(I-A^{+}A)\vec{Z}$ 是最小二乘解的通解；

3．矛盾方程组 $A\vec{X}=\vec{b}$ 的极小范数最小二乘解（即最佳逼近解）为 $\vec{X}=A^{+}\vec{b}$.

例： 已知 $A=\begin{pmatrix}1&1&0&1\\0&1&1&0\\1&2&1&1\end{pmatrix},\vec{b}=\begin{pmatrix}3\\1\\4\end{pmatrix}$．1．求 A 的满秩分解；2．求 A^{+}；3．判断线性方程组 $A\overrightarrow{X}=\vec{b}$ 是否有解，求解并指出解的类型.

解： $A=\begin{pmatrix}1&1&0&1\\0&1&1&0\\1&2&1&1\end{pmatrix}\simeq\begin{pmatrix}1&1&0&1\\0&1&1&0\\0&1&1&0\end{pmatrix}\simeq\begin{pmatrix}1&0&-1&1\\0&1&1&0\\0&0&0&0\end{pmatrix}$， $\therefore A$ 的满秩分解为：

$A=BC=\begin{pmatrix}1&1\\0&1\\1&2\end{pmatrix}\begin{pmatrix}1&0&-1&1\\0&1&1&0\end{pmatrix}$，其中 $B=\begin{pmatrix}1&1\\0&1\\1&2\end{pmatrix},C=\begin{pmatrix}1&0&-1&1\\0&1&1&0\end{pmatrix}$.

由于 $B=\begin{pmatrix}1&1\\0&1\\1&2\end{pmatrix}\simeq\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}$，$rankB=2$，即 B 为列满秩矩阵，那么 $B^{+}=B_L^{-1}=(B^TB)^{-1}B^T$

$$=\left(\begin{pmatrix}1&0&1\\1&1&2\end{pmatrix}\begin{pmatrix}1&1\\0&1\\1&2\end{pmatrix}\right)^{-1}\begin{pmatrix}1&0&1\\1&1&2\end{pmatrix}=\begin{pmatrix}2&3\\3&6\end{pmatrix}^{-1}\begin{pmatrix}1&0&1\\1&1&2\end{pmatrix}=\begin{pmatrix}2&-1\\-1&\frac{2}{3}\end{pmatrix}\begin{pmatrix}1&0&1\\1&1&2\end{pmatrix}=\begin{pmatrix}1&-1&0\\-\frac{1}{3}&\frac{2}{3}&\frac{1}{3}\end{pmatrix}$$

$$=\frac{1}{3}\begin{pmatrix}3&-3&0\\-1&2&1\end{pmatrix}$$

显然 C 为行满纸矩阵，那么

$$C^{+}=C_R^{-1}=C^T(CC^T)^{-1}=\begin{pmatrix}1&0\\0&1\\-1&1\\1&0\end{pmatrix}\left(\begin{pmatrix}1&0&-1&1\\0&1&1&0\end{pmatrix}\begin{pmatrix}1&0\\0&1\\-1&1\\1&0\end{pmatrix}\right)^{-1}=\begin{pmatrix}1&0\\0&1\\-1&1\\1&0\end{pmatrix}\begin{pmatrix}3&-1\\-1&2\end{pmatrix}^{-1}$$

$$=\begin{pmatrix}1&0\\0&1\\-1&1\\1&0\end{pmatrix}\begin{pmatrix}\frac{2}{5}&\frac{1}{5}\\\frac{1}{5}&\frac{3}{5}\end{pmatrix}=\frac{1}{5}\begin{pmatrix}1&0\\0&1\\-1&1\\1&0\end{pmatrix}\begin{pmatrix}2&1\\1&3\end{pmatrix}=\frac{1}{5}\begin{pmatrix}2&1\\1&3\\-1&2\\2&1\end{pmatrix}$$

$$\therefore A^{+}=C^{+}B^{+}=\frac{1}{15}\begin{pmatrix}2&1\\1&3\\-1&2\\2&1\end{pmatrix}\begin{pmatrix}3&-3&0\\-1&2&1\end{pmatrix}=\frac{1}{15}\begin{pmatrix}5&-4&1\\0&3&3\\-5&7&2\\5&-4&1\end{pmatrix}.$$

显然 $rankA=2$，而 $A\left|\vec{b}\right.=\begin{pmatrix}1&1&0&1&3\\0&1&1&0&1\\1&2&1&1&4\end{pmatrix}\simeq\begin{pmatrix}1&0&-1&1&2\\0&1&1&0&1\\0&0&0&0&0\end{pmatrix}$，则 $rank(A\left|\vec{b}\right.)=2$，

所以线性方程组是相容的，具有极小范数解 $\overrightarrow{X}=A^{+}\vec{b}=\dfrac{1}{15}\begin{pmatrix}5&-4&1\\0&3&3\\-5&7&2\\5&-4&1\end{pmatrix}\begin{pmatrix}3\\1\\4\end{pmatrix}=\begin{pmatrix}1\\1\\0\\1\end{pmatrix}$.

通解为：

$$\overrightarrow{X}=A^{+}\vec{b}+(I-A^{+}A)\overrightarrow{Z}=\begin{pmatrix}1\\1\\0\\1\end{pmatrix}+\frac{1}{15}\begin{pmatrix}6&3&-3&6\\3&9&6&3\\-3&6&9&-3\\6&3&-3&6\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\\z_4\end{pmatrix}=\begin{pmatrix}\frac{1}{5}(2z_1+z_2-z_3+2z_4)+1\\\frac{1}{5}(z_1+3z_2+2z_3+z_4)+1\\\frac{1}{5}(-z_1+2z_2+3z_3-z_4)\\\frac{1}{5}(2z_1+z_2-z_3+2z_4)+1\end{pmatrix}$$

其中 $\overrightarrow{Z}=\begin{pmatrix}z_1\\z_2\\z_3\\z_4\end{pmatrix}$ 为任意向量.

本章测试题

一、已知 $A=\begin{pmatrix}1&1&1&1\\1&1&1&1\end{pmatrix}$，则 $A_r^-=$____________________

二、设 A 是 n 阶可逆矩阵，O 是 n 阶零矩阵，则 $\begin{pmatrix}O&A\\O&O\end{pmatrix}^+=$________________

三、设 $\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_m}(m>1)$ 是 R^n 中两两正交的单位列向量，记 $A=(\overrightarrow{x_1},\overrightarrow{x_2},\cdots,\overrightarrow{x_m})$，则 $A^+=$________________

四、已知 $A=\begin{pmatrix}1&1&1&1\\1&2&3&4\\0&1&2&3\end{pmatrix},\vec{b}=\begin{pmatrix}1\\2\\1\end{pmatrix}$．1．求 A 的满秩分解；2．求 A^{+}；3．判断线性方程组 $A\overrightarrow{X}=\vec{b}$ 是否有解，求解并指出解的类型.

第八章 克罗内克Kronecker积及其应用

定义 1-1 设 $A=(a_{ij})\in C^{m\times n}, B=(b_{ij})\in C^{p\times q}$，则称如下的分块矩阵

$$A\otimes B=\begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}\in C^{mp\times nq}$$ 为 A 的克罗内克积，或称 A 与 B 的直积，

或张量积，简记为 $A\otimes B=(a_{ij}B)_{mp\times nq}$，即 $A\otimes B$ 是一个 $m\times n$ 块的分块矩阵，最后是一个 $mp\times nq$ 阶的矩阵.

Kronecker 内积满足下列运算规律：

1. $k(A\otimes B)=kA\otimes B=A\otimes kB, k\in C$；
2. $(A+B)\otimes C=A\otimes C+B\otimes C$；
3. $(A\otimes B)\otimes C=A(B\otimes C)$.

定理 1-1 设 $A=(a_{ij})_{m\times n}, B=(b_{ij})_{s\times r}, C=(c_{ij})_{n\times p}, D=(d_{ij})_{r\times l}$，则

$$(A\otimes B)\otimes(C\otimes D)=AC\otimes BD$$

推论 若 $A=(a_{ij})_{m\times m}, B=(b_{ij})_{n\times n}$，则 $A\otimes B=(A\otimes I_n)(I_m\otimes B)=(I_m\otimes B)(A\otimes I_n)$

定理 1-2 设 $A=(a_{ij})_{m\times n}, B=(b_{ij})_{p\times q}$，则 $(A\otimes B)^T=A^T\otimes B^T, (A\otimes B)^H=A^H\otimes B^H$

定理 1-3 设 A,B 分别为 m 阶和 n 阶可逆矩阵，则 $A\otimes B$ 也是可逆矩阵，且

$$(A\otimes B)^{-1}=A^{-1}\otimes B^{-1}$$

定理 1-4 设 $A=(a_{ij})_{m\times n}, B=(b_{ij})_{p\times q}$，则 $rank(A\otimes B)=rank(A)rank(B)$

定理 1-5 设 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 是 $A_{m\times m}$ 的 m 个特征值；$u_1,u_2,\cdots,u_p$ 是 $B_{p\times p}$ 的 p 个特征值，那么

$A \otimes B$ 的 mp 个特征值为 $\lambda_i u_j (i=1,2,\cdots,m; j=1,2,\cdots,p)$.

定理 1-6 设 A 为 m 阶矩阵，B 为 p 阶矩阵，则 $|A \otimes B| = |A|^p |B|^m$.

定理 1-7 设 A 为对角矩阵（对称矩阵、*Hermite* 矩阵、正交矩阵），则 $A \otimes B$ 也是对角矩阵（对称矩阵、*Hermite* 矩阵、正交矩阵）.

定理 1-8 设 A 为 m 阶矩阵，B 为 n 阶矩阵，则 $A \otimes B$ 相似于 $B \otimes A$.

定理 1-9 设 $A=(a_{ij}) \in C^{m\times n}, B=(b_{ij}) \in C^{p\times q}$，则 $(AB)^{(k)} = A^{(k)} B^{(k)}$

本章测试题

一、设 $\vec{x} \in R^n$ 是单位列向量，$A \in R^{n\times n}$ 是正交矩阵，则 $\left\| A \otimes \vec{x} \right\|_F =$ ________

解：$\left\| A \otimes \vec{x} \right\|_F = \sqrt{tr\left[\left(A \otimes \vec{x}\right)^H \left(A \otimes \vec{x}\right)\right]} = \sqrt{tr\left[\left(A^H \otimes \vec{x}^H\right)\left(A \otimes \vec{x}\right)\right]}$

$$= \sqrt{tr\left[\left(A^H A\right) \otimes \left(\vec{x}^H \vec{x}\right)\right]} = \sqrt{tr\left(I_n \otimes 1\right)} = \sqrt{trI_n} = \sqrt{n}$$

二、设 A 是 m 阶方阵，B 是 n 阶方阵，且 A 和 B 都是可对角化矩阵，证明：$A \otimes B$ 也可对角化.

解：由于 $P^{-1}AP = \Lambda_1 = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix}, Q^{-1}BQ = \Lambda_2 = \begin{pmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \cdots & \\ & & & \mu_n \end{pmatrix}$

则 $\left(P \otimes Q\right)^{-1}\left(A \otimes B\right)\left(P \otimes Q\right) = \left(P^{-1} \otimes Q^{-1}\right)\left(A \otimes B\right)\left(P \otimes Q\right) = \left(P^{-1}AP\right) \otimes \left(Q^{-1}BQ\right)$

$$= \Lambda_1 \otimes \Lambda_2 = \begin{pmatrix} \lambda_1\mu_1 & & & & & & \\ & \cdots & & & & & \\ & & \lambda_1\mu_n & & & & \\ & & & \cdots & & & \\ & & & & \lambda_m\mu_1 & & \\ & & & & & \cdots & \\ & & & & & & \lambda_m\mu_n \end{pmatrix}_{mn\times mn}.$$