第五章

线性变换

上一章中介绍了线性空间的概念, 本章将讨论线性空间之间的联系. 它们之间的联系主要反映为线性空间之间的映射, 所以研究定义域和值域都是线性(子)空间的映射是数学分析的基本目标之一, 其中最简单和最基本的一类映射是线性变换(Linear Transformation). 它也是线性代数中一个主要研究对象.

5.1 线性变换基本概念

先来讨论线性空间 V 到自身的映射, 常称为V的一个变换. 用符号σ表示, 即$\sigma : V \mapsto V$.

5.1.1 线性变换定义与性质

定义 5.1. *数域F上的n维线性空间V的一个变换σ, 对于V中的任意两个向量$\mathbf{x}, \mathbf{y}$ 和数$c \in F$, 若满足下列条件:*

(1) $\sigma(\mathbf{x}+\mathbf{y}) = \sigma(\mathbf{x}) + \sigma(\mathbf{y})$.

(2) $\sigma(c\mathbf{x}) = c\sigma(\mathbf{x})$.

*则称σ为***线性变换***.*

例 1. *设V是数域F上的线性空间, 给定$c \in F$定义变换$\sigma_c : V \mapsto V$, 其中 $\sigma_c(\mathbf{x}) = c\mathbf{x}$, 显然它满足线性变换定义中的两个条件. 若$c = 1$, 则称$\sigma_1$为$V$上的***恒等变换***, 若$c = 0$, 则称$\sigma_0$为$V$上的零变换.*

例 2. *数域F上的n次多项式空间$F[x]_n$, 变换$d/dx : F[x]_n \mapsto F[x]_n$, 是对$F[x]_n$中的任意多项式求它导数, 即*

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \overset{d/dx}{\longmapsto} a_1 + 2a_2x + \cdots + na_nx^{n-1}$$

显然d/dx满足线性变换的两个条件.

例 3. *设 $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$, 其中*

$$\begin{bmatrix} x \\ y \end{bmatrix} \overset{f}{\longmapsto} \begin{bmatrix} 2x \\ 3x-2y \end{bmatrix}$$

$\forall \begin{bmatrix} x_1 & y_1 \end{bmatrix}^T, \begin{bmatrix} x_2 & y_2 \end{bmatrix}^T$*和$c \in \mathbb{R}$, 有*

$$\begin{bmatrix} x_1+x_2 \\ y_1+y_2 \end{bmatrix} \overset{f}{\longmapsto} \begin{bmatrix} 2(x_1+x_2) \\ 3(x_1+x_2)-2(y_1+y_2) \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 3x_1-2y_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 3x_2-2y_2 \end{bmatrix}$$

$$c\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \overset{f}{\longmapsto} \begin{bmatrix} 2cx_1 \\ 3cx_1-2cy_1 \end{bmatrix} = c\begin{bmatrix} 2x_1 \\ 3x_1-2y_1 \end{bmatrix}$$

可见, f是线性变换.

例 4. *平面几何中的向量的旋转映射 σ_θ, 即将向量$\vec{a}$映射为绕原点逆时针旋转θ角后的向量$\vec{a}'$. 它满足对任意的$\vec{a}, \vec{b}$和常数c.*

$$\sigma_\theta\left(\vec{a}+\vec{b}\right) = \sigma_\theta(\vec{a}) + \sigma_\theta\left(\vec{b}\right), \quad \sigma_\theta(c\vec{a}) = c\sigma_\theta(\vec{a})$$

因此旋转变换是线性变换.

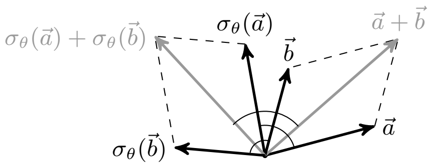


图 5.1: 旋转变换

例 5. *三维几何空间中关于平面的镜像变换σ, 就是以过原点的平面L为"镜面", 单位向量$\vec{n}$是镜面L的法向, 即 $\vec{n}\perp L$, 对任意向量$\vec{x}$ 它的镜像$\vec{x'}$为: $\vec{x}$在法向$\vec{n}$上的投影的反射$-(\vec{x}\cdot\vec{n})\vec{n}$与$\vec{x}$在镜面$L$上的投影向量$\vec{x}-(\vec{x}\cdot\vec{n})\vec{n}$之和, 即*

$$\sigma:\vec{x}\mapsto\vec{x'},\quad 其中\vec{x'}=\vec{x}-2\left(\vec{x}\cdot\vec{n}\right)\vec{n}$$

对任意向量 $\vec{x}$,$\vec{y}$ 和实数c, 有

$$\begin{aligned}\sigma\left(\vec{x}+\vec{y}\right)=&\quad\vec{x}+\vec{y}-2\left[\left(\vec{x}+\vec{y}\right)\cdot\vec{n}\right]\vec{n}\\ =&\quad\vec{x}+\vec{y}-2\left(\vec{x}\cdot\vec{n}\right)\vec{n}-2\left(\vec{y}\cdot\vec{n}\right)\vec{n}=\sigma\left(\vec{x}\right)+\sigma\left(\vec{y}\right)\\ \sigma\left(c\vec{x}\right)=&\quad c\vec{x}-2\left(c\vec{x}\cdot\vec{n}\right)\vec{n}=c\left(\vec{x}-2\left(\vec{x}\cdot\vec{n}\right)\vec{n}\right)=c\sigma\left(\vec{x}\right)\end{aligned}$$

σ是线性变换.

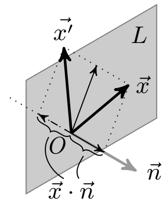


图5.2: 镜像变换

设σ是线性空间V的任意线性变换, 则具有下列性质:

(1) 对任意的V中元素$\mathbf{x}$ 有 $\sigma\left(-\mathbf{x}\right)=-\sigma\left(\mathbf{x}\right)$, 及 $\sigma\left(\mathbf{0}\right)=\mathbf{0}$.

(2) 线性变换保持线性组合或线性关系式不变:

$$\sigma\left(c_1\mathbf{x}_1+c_1\mathbf{x}_2+\cdots+c_n\mathbf{x}_n\right)=c_1\sigma\left(\mathbf{x}_1\right)+c_2\sigma\left(\mathbf{x}_2\right)+\cdots+c_n\sigma\left(\mathbf{x}_n\right)$$

(3) 线性相关的向量组经线性变换后仍保持线性相关.

定义 5.2. *设σ 是数域F上线性空间V的线性变换, 则将下列两个集合*

$$\begin{aligned}Ker\left(\sigma\right)=\mathcal{N}\left(\sigma\right)=\left\{\mathbf{x}\middle|\mathbf{x}\in V\wedge\sigma\left(\mathbf{x}\right)=\mathbf{0}\right\}\\ Im\left(\sigma\right)=\mathcal{R}\left(\sigma\right)=\left\{\mathbf{y}\middle|\mathbf{y}=\sigma\left(\mathbf{x}\right),\forall\mathbf{x}\in V\right\}\end{aligned}\tag{5.1}$$

*分别称为变换σ的**核**和σ的**值域**或**像**.*

定理 5.1.1. *设σ是数域F上线性空间V的线性变换, 则$Ker\left(\sigma\right)$和$Im\left(\sigma\right)$是V的子空间.*

证: (1). $Ker\left(\sigma\right)$ 和 $Im\left(\sigma\right)$ 均包含了$\mathbf{0}$, 所以它们非空.

(2). 加法封闭性

$$\begin{aligned}&\forall\mathbf{x},\mathbf{y}\in Ker\left(\sigma\right),\quad\sigma\left(\mathbf{x}+\mathbf{y}\right)=\sigma\left(\mathbf{x}\right)+\sigma\left(\mathbf{y}\right)=\mathbf{0}\Rightarrow\mathbf{x}+\mathbf{y}\in Ker\left(\sigma\right)\\ &\forall\mathbf{y}_1,\mathbf{y}_2\in Im\left(\sigma\right),\quad\exists\mathbf{x_1},\mathbf{x_2}\in V,使得\mathbf{y}_1=\sigma\left(\mathbf{x}_1\right),\mathbf{y}_2=\sigma\left(\mathbf{x}_2\right),\\ &\therefore\mathbf{y}_1+\mathbf{y}_2=\sigma\left(\mathbf{x}_1\right)+\sigma\left(\mathbf{x}_2\right)=\sigma\left(\mathbf{x}_1+\mathbf{x}_2\right)\in Im\left(\sigma\right)\end{aligned}$$

(3). 数乘封闭性, 对$\forall c\in F$

$$\begin{aligned}&\forall\mathbf{x}\in Ker\left(\sigma\right),\quad\sigma\left(c\mathbf{x}\right)=c\sigma\left(\mathbf{x}\right)=c\mathbf{0}=\mathbf{0}\Rightarrow c\mathbf{x}\in Ker\left(\sigma\right)\\ &\forall\mathbf{y}\in Im\left(\sigma\right),\quad\exists\mathbf{x}\in V使得\mathbf{y}=\sigma\left(\mathbf{x}\right),则c\mathbf{y}=c\sigma\left(\mathbf{x}\right)=\sigma\left(c\mathbf{x}\right)\in Im\left(\sigma\right)\end{aligned}$$

例 6. *数域F上的n次多项式空间$F[x]_n$, 设 $\sigma=d/dx$, 则$Ker\left(\sigma\right)=F$, $Im\left(\sigma\right)=F[x]_{n-1}$.*

定理 5.1.2. *数域F上的线性空间V的线性变换σ是一一映射, 当且仅当$Ker\left(\sigma\right)=\{\mathbf{0}\}$.*

证: 若 σ是一一映射, 设 $\sigma(\mathbf{x})=\mathbf{0}$, 根据线性变换性质(1)$\mathbf{x}=\mathbf{0}$.

反之, 若 $Ker(\sigma)=\{\mathbf{0}\}$, 假设存在 $\mathbf{x}_1,\mathbf{x}_2\in V$ 使得 $\sigma(\mathbf{x}_1)=\sigma(\mathbf{x}_2)$, 则根据线性变换的条件, 有

$$\mathbf{0}=\sigma(\mathbf{x}_1)-\sigma(\mathbf{x}_2)=\sigma(\mathbf{x}_1-\mathbf{x}_2)$$

因 $\mathbf{x}_1-\mathbf{x}_2\in Ker(\sigma)$ 即 $\mathbf{x}_1=\mathbf{x}_2$, 说明 σ 是单射的, 同时$\sigma: V\mapsto V$, 又是满射, 因此σ是一一映射.

5.1.2 线性变换的运算

定义 5.3. *设 σ,τ是数域F上线性空间V的任意线性变换, 定义 $\sigma+\tau$ 为 $V\mapsto V$的映射, 且*

$$\forall\mathbf{x}\in V,\quad(\sigma+\tau)(\mathbf{x})=\sigma(\mathbf{x})+\tau(\mathbf{x})$$

同时, 对任意的$c\in F$, 定义 $c\sigma$ 是 $V\mapsto V$的映射:

$$\forall\mathbf{x}\in V,\quad(c\sigma)(\mathbf{x})=c\sigma(\mathbf{x})$$

定义5.3给出了线性变换间的加法和数乘运算, 根据定义易证线性变换的和与数乘也是线性变换,即满足线性变换的线性条件:

对任意的$\mathbf{x},\mathbf{y}\in V$和$c,d\in F$, 有

$$\begin{aligned}(\sigma+\tau)(\mathbf{x}+\mathbf{y}) &= \sigma(\mathbf{x}+\mathbf{y})+\tau(\mathbf{x}+\mathbf{y})\\ &= \sigma(\mathbf{x})+\sigma(\mathbf{y})+\tau(\mathbf{x})+\tau(\mathbf{y})\\ &= (\sigma+\tau)(\mathbf{x})+(\sigma+\tau)(\mathbf{y})\\ (c\sigma)(d\mathbf{x}) &= c\sigma(d\mathbf{x})=cd\sigma(\mathbf{x})=dc\sigma(\mathbf{x})=d(c\sigma)(\mathbf{x})\end{aligned}$$

关于线性空间V上线性变换定义集合$L(V)=\left\{\sigma\middle|\sigma\text{是}V\text{的线性变换}\right\}$, 则有

定理 5.1.3. *$L(V)$关于定义5.3中的线性变换间的加法和数乘构成线性空间.*

(留给读者证明)

定义 5.4. *设A是数域F上的线性空间, 在A上定义"乘法"运算, 用符号"$\cdot$"表示, 事实上"$\cdot$": $A\times A\mapsto A$. 它使得对A中的任意元素α,β,γ 和数$c\in F$, 满足以下条件:*

(1) "乘法"成立结合律, 即$\alpha\cdot(\beta\cdot\gamma)=(\alpha\cdot\beta)\cdot\gamma$;

(2) A中存在元素e, 使得$e\cdot\alpha=\alpha\cdot e=\alpha$.

(3) "乘法"成立分配律, 即

$$\begin{aligned}&\alpha\cdot(\beta+\gamma)=\alpha\cdot\beta+\alpha\cdot\gamma, &&(\text{左分配律})\\ &(\beta+\gamma)\cdot\alpha=\beta\cdot\alpha+\gamma\cdot\alpha, &&(\text{右分配率})\\ &(c\alpha)\cdot\beta=c(\alpha\cdot\beta)=\alpha\cdot(c\beta)\end{aligned}$$

*则称A是数域F上的**代数**, 而元素e称为A的**恒等元**.*

有时常用"**1**"表示恒等元, 但注意与数域F中的1的区别.

定义 5.5. *对线性空间$L(V)$ 定义线性变换的"乘法", 即$\forall\sigma,\tau\in L(V)$,*

$$\sigma\cdot\tau=\sigma(\tau(\bullet))$$

定理 5.1.4. *设V是数域F上的线性空间, 则$L(V)$是F上的代数.*

证: 验证$L(V)$上关于线性变换的乘法满足定义5.4中的三个条件:

(1) 对 $\forall \sigma,\tau,\pi \in L(V)$, 有

$$(\sigma\cdot\tau)\cdot\pi=(\sigma\cdot\tau)(\pi(\bullet))=\sigma(\tau(\pi(\bullet)))=\sigma((\tau\cdot\pi)(\bullet))=\sigma\cdot(\tau\cdot\pi)$$

(2) $L(V)$中元素V上的恒等变换"$\mathbf{1}_V$"即为e, 且对$\forall\sigma\in V$, 满足 $\mathbf{1}_V\cdot\sigma=\sigma\cdot\mathbf{1}_V=\sigma$, 因此恒等变换是$L(V)$的恒等元.

(3) 对$\forall\sigma,\tau,\pi\in L(V)$, 有

$$\begin{aligned} & [\sigma\cdot(\tau+\pi)](\bullet)=\sigma((\tau+\pi)(\bullet))=\sigma(\tau(\bullet)+\pi(\bullet)) \\ = \ & \sigma(\tau(\bullet))+\sigma(\pi(\bullet))=(\sigma\cdot\tau)(\bullet)+(\sigma\cdot\pi)(\bullet) \end{aligned}$$

由此左分配律成立,即 $\sigma\cdot(\tau+\pi)=\sigma\cdot\tau+\sigma\cdot\pi$. 同理可证明右分配律成立.

对$\forall c\in F,\sigma,\tau\in L(V)$, 有

$$[(c\sigma)\cdot\tau](\bullet)=(c\sigma)(\tau(\bullet))=c\sigma(\tau(\bullet))=c(\sigma\cdot\tau)(\bullet)$$

从而, $(c\sigma)\cdot\tau=c(\sigma\cdot\tau)$成立. 同理可证 $\sigma\cdot(c\tau)=c(\sigma\cdot)$.

综上所述, $L(V)$是F上的代数.

例 7. *设σ,τ为$\mathbb{R}^2$空间上的线性变换, 分别定义如下:*

$$\forall\begin{bmatrix}x\\y\end{bmatrix}\in\mathbb{R}^2,\quad \sigma\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}x\\x-y\end{bmatrix},\quad \tau\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}y\\x\end{bmatrix}$$

求 $\alpha=\begin{bmatrix}-3 & 2\end{bmatrix}^T$ 在变换 $\sigma\cdot\tau$ 和$\tau\cdot\sigma$ 下的像.

解: 根据变换的定义先求出变换乘积,

$$\begin{aligned}\sigma\cdot\tau=\sigma\left(\tau\left(\begin{bmatrix}x\\y\end{bmatrix}\right)\right)=\sigma\left(\begin{bmatrix}y\\x\end{bmatrix}\right)=\begin{bmatrix}y\\y-x\end{bmatrix}\\ \tau\cdot\sigma=\tau\left(\sigma\left(\begin{bmatrix}x\\y\end{bmatrix}\right)\right)=\tau\left(\begin{bmatrix}x\\x-y\end{bmatrix}\right)=\begin{bmatrix}x-y\\x\end{bmatrix}\end{aligned}$$

将α代入, 得变换$\sigma\cdot\tau$下的像为$\begin{bmatrix}2 & 5\end{bmatrix}^T$, 在$\tau\cdot\sigma$下的像为 $\begin{bmatrix}-5 & 3\end{bmatrix}^T$.

下面定义线性变换的幂运算, 即对$\forall\sigma\in L(V)$

$$\begin{aligned}&\sigma^0=\mathbf{1}_V(V\text{上的恒等变换})\\ &\sigma^{n+1}=\sigma^n\cdot\sigma,\quad n\in\mathbb{N}^+\end{aligned}$$

根据$L(V)$上的乘法成立结合律, 可证明$L(V)$上的幂运算成立指数律, 即对$\forall n,m\in\mathbb{N}$, 有

$$\sigma^n\cdot\sigma^m=\sigma^{n+m},\quad (\sigma^n)^m=\sigma^{nm} \tag{5.2}$$

对于数域F上的任意多项式:

$$f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$$

根据它, 可定义线性变换的多项式:

$$f(\sigma)=a_0\mathbf{1}_V+a_1\sigma+a_2\sigma^2+\cdots+a_n\sigma^n$$

若σ是一一映射, 则变换σ可逆, 记为σ^{-1}, 又称它为σ的逆变换. 显然, 它们也是$V\mapsto V$的自同构映射. 因而在σ可逆的前提下, 可定义它的负数次幂:$\sigma^{-n}=\left(\sigma^{-1}\right)^n$.

因$L(V)$上的乘法一般不满足交换律, 所以 $\sigma,\tau\in L(V)$, 一般 $\sigma\cdot\tau\neq\tau\cdot\sigma$, 且 $(\sigma\cdot\tau)^n\neq\sigma^n\tau^n$; 若$\sigma,\tau$都可逆, 则有:$(\sigma\cdot\tau)^{-1}=\tau^{-1}\cdot\sigma^{-1}$; 对于非零数$k\in F$有: $(k\sigma)^{-1}=k^{-1}\sigma^{-1}$.(这些结论读者自行证明)

5.2 线性变换与矩阵

下面研究线性变换与矩阵之间的关系.

5.2.1 线性变换的矩阵表示

设V是数域F上的n维线性空间, $\mathfrak{B}=\{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\}$ 是V的基, 则对$\forall\mathbf{x}\in V$ 可表示成 $\mathbf{x}=\sum_{i=1}^{n}x_i\varepsilon_i$, 考察$V$上的线性变换 σ 有:

$$\sigma(\mathbf{x})=\sum_{i=1}^{n}x_i\sigma(\varepsilon_i)$$

可见, V中任意向量在线性变换下的像可表示为V中的基在线性变换下的像的线性组合. 因σ是V到V的线性变换, 因此$\sigma(\varepsilon_i)(i=1,2,\ldots,n)$ 在基$\mathfrak{B}$下可表示成

$$\begin{bmatrix}\sigma(\varepsilon_1) & \sigma(\varepsilon_2) & \cdots & \sigma(\varepsilon_n)\end{bmatrix}=\begin{bmatrix}\varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n\end{bmatrix}\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}\tag{5.3}$$

上式中矩阵$[a_{ij}]$的第j列向量$\mathbf{a}_j$事实上是$\sigma(\varepsilon_j)$在基 $\mathfrak{B}$下的坐标, 按上一章中坐标的记法, 如 向量$\mathbf{x}$在基$\mathfrak{B}$下的坐标表示为 $[\mathbf{x}]_{\mathfrak{B}}$, 可将矩阵$[a_{ij}]$ 记为 $[\sigma(\mathfrak{B})]_{\mathfrak{B}}$, 则$V$中任意向量$\mathbf{x}$ 经线性变换 σ后的像可表示成:

$$\sigma(\mathbf{x})=\mathfrak{B}\left[\sigma(\mathfrak{B})\right]_{\mathfrak{B}}[\mathbf{x}]_{\mathfrak{B}}\tag{5.4}$$

由式(5.4)可知, 当V选定基$\mathfrak{B}$以后, V上的线性变换与矩阵 $[\sigma(\mathfrak{B})]_{\mathfrak{B}}$对应. 由任意向量在基下表示的唯一性, 线性变换$\sigma$对应的矩阵$[\sigma(\mathfrak{B})]_{\mathfrak{B}}$是唯一的. 反之, 对于$n$阶方阵$\mathbf{A}\in F^{n\times n}$ 是否唯一对应一个线性变换? 下述定理说明结论是成立的.

定理 5.2.1. *数域F上的n维线性空间V, 对$\forall\mathbf{A}\in F^{n\times n}$ 都存在V上唯一的线性变换与之对应.*

证: 设 $\mathfrak{B}=\{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\}$ 是线性空间V的基, 对$\forall\mathbf{A}\in F^{n\times n}$ 按下列方式构造一组向量

$$\alpha_j=\mathfrak{B}\mathbf{a}_j,\quad j=1,2,\ldots,n\tag{5.5}$$

其中 $\mathbf{a}_j$为矩阵$\mathbf{A}$的第j列向量, 从向量构造形式看就是将矩阵$\mathbf{A}$的第 j 列向量视作 α_j 在基$\mathfrak{B}$下的坐标. 由坐标的唯一性可知, 向量组$\alpha_j(j=1,2,\ldots,n)$由矩阵$\mathbf{A}$唯一确定. 同时, 对V的任意向量$\mathbf{x}$, 在基$\mathfrak{B}$下的坐标表示为$[\mathbf{x}]_{\mathfrak{B}}$, 即$\mathbf{x}=\mathfrak{B}[\mathbf{x}]_{\mathfrak{B}}$. 现定义$V$上的变换$\sigma$, 它满足下列条件, 对$\forall\mathbf{x}\in V$ 有

$$\sigma(\mathbf{x})=\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_3\end{bmatrix}[\mathbf{x}]_{\mathfrak{B}}=\mathfrak{B}\mathbf{A}[\mathbf{x}]_{\mathfrak{B}}\tag{5.6}$$

利用$[\bullet]_{\mathfrak{B}}$是V上的同构映射, 对$\forall\mathbf{x},\mathbf{y}\in V$, 有

$$\begin{aligned}\sigma(\mathbf{x}+\mathbf{y}) &= \begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{bmatrix}[\mathbf{x}+\mathbf{y}]_{\mathfrak{B}}\\ &= \begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{bmatrix}\left([\mathbf{x}]_{\mathfrak{B}}+[\mathbf{y}]_{\mathfrak{B}}\right)=\sigma(\mathbf{x})+\sigma(\mathbf{y})\end{aligned}$$

同理, 对$\forall k\in F$,

$$\sigma(k\mathbf{x})=\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{bmatrix}[k\mathbf{x}]_{\mathfrak{B}}=k\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{bmatrix}[\mathbf{x}]_{\mathfrak{B}}=k\sigma(\mathbf{x})$$

这就证明了 σ 是V上的线性变换. 若令$\mathbf{x}=\varepsilon_j$, 则$[\mathbf{x}]_{\mathfrak{B}}=\mathbf{e}_j$, 根据定义式(5.6)有 $\sigma(\varepsilon_j)=\alpha_j$. 同样根据式(5.6)可知这种对应关系是唯一的.

定理 5.2.2. *数域F上n维线性空间V的所有线性变换构成的线性空间$L(V)$, 则 $L(V)$和$F^{n\times n}$ 同构.*

根据本节开始的叙述, 在选取V的基后, 存在一种$L(V)$到$F^{n\times n}$的一种一一对应关系, 只需证明这种对应关系同时对线性变换的加法、数乘、乘法和矩阵的加法、数乘、乘法之间满足线性条件. 详细证明留作练习.

例 8. *分别求例1—4中线性变换对应的矩阵.*

解: (1) 对例1定义的线性变换σ, 取V的基$\mathfrak{B}=\{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\}$, 则有

$$\begin{bmatrix} \sigma(\varepsilon_1) & \sigma(\varepsilon_2) & \cdots & \sigma(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} c\varepsilon_1 & c\varepsilon_2 & \cdots & c\varepsilon_n \end{bmatrix} = \mathfrak{B}c\mathbf{E}$$

对应矩阵为$c\mathbf{E}$.

(2). 选取空间$F[x]_n$的基为 $\mathfrak{B}=\left\{1,x,x^2,\cdots,x^n\right\}$,则

$$\begin{aligned}&\begin{bmatrix} \frac{d}{dx}(1) & \frac{d}{dx}(x) & \frac{d}{dx}\left(x^2\right) & \cdots & \frac{d}{dx}\left(x^n\right) \end{bmatrix}\\ =& \begin{bmatrix} 1 & x & x^2 & \cdots & x^n \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & n \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix} = \mathfrak{B}\mathbf{D}\end{aligned}$$

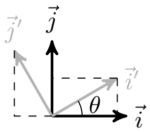
则$\mathbf{D}$为线性变换d/dx对应的矩阵.

(3). 取$\mathfrak{B}=\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$ 为 $\mathbb{R}^2$的基, 在线性变换f下有

$$\begin{bmatrix} f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) & f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 3 & -2 \end{bmatrix} = \mathfrak{B}\mathbf{f}$$

则矩阵$\mathbf{f}$为线性变换f对应的矩阵.

(4). 取二维几何空间中的标准正交向量$\vec{i},\vec{j}$为基, 实施旋转变换σ_θ后为$\vec{i}',\vec{j}'$(如图)



$$\begin{bmatrix} \sigma_\theta\left(\vec{i}\right) & \sigma_\theta\left(\vec{j}\right) \end{bmatrix} = \begin{bmatrix} \vec{i} & \vec{j} \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{5.7}$$

这小节建立了线性变换与矩阵之间的联系, 即线性变换与矩阵之间一一对应, 从而可通过矩阵工具来研究线性变换.

5.2.2 线性变换在不同基下的矩阵间的关系

上一小节讨论了线性空间V中的任意线性变换在选定基以后与矩阵唯一对应. 本节将讨论若选择不同的基, 则与线性变换相应的矩阵之间有怎样的关系?

定理 5.2.3. *设$\mathfrak{B}_1=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ 和 $\mathfrak{B}_2=\{\beta_1,\beta_2,\ldots,\beta_n\}$ 为数域F上的n维线性空间V的两个基, σ 为 V上的任意线性变换, 则σ在两个基下对应的矩阵$[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_1}$, $[\sigma(\mathfrak{B}_2)]_{\mathfrak{B}_2}$之间成立下列关系:*

$$[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_1} = \mathbf{M}^{-1}[\sigma(\mathfrak{B}_2)]_{\mathfrak{B}_2}\mathbf{M} \tag{5.8}$$

其中$\mathbf{M}$是基$\mathfrak{B}_2$到$\mathfrak{B}_1$的过渡矩阵.

证: $\forall\mathbf{x}\in V$, 根据式(5.4), 在基$\mathfrak{B}_1,\mathfrak{B}_2$ 下有

$$\sigma(\mathbf{x}) = \mathfrak{B}_1[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_1}[\mathbf{x}]_{\mathfrak{B}_1} = \mathfrak{B}_2[\sigma(\mathfrak{B}_2)]_{\mathfrak{B}_2}[\mathbf{x}]_{\mathfrak{B}_2}$$

由$\mathbf{M}$ 为$\mathfrak{B}_2$到$\mathfrak{B}_1$的过渡矩阵, 有$\mathfrak{B}_1\mathbf{M}^{-1}=\mathfrak{B}_2$ 及两个基下坐标之间转换关系 $[\mathbf{x}]_{\mathfrak{B}_2}=\mathbf{M}[\mathbf{x}]_{\mathfrak{B}_1}$, 将它们代入上式中, 得

$$\mathfrak{B}_1[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_1}[\mathbf{x}]_{\mathfrak{B}_1} = \mathfrak{B}_1\mathbf{M}^{-1}[\sigma(\mathfrak{B}_2)]_{\mathfrak{B}_2}\mathbf{M}[\mathbf{x}]_{\mathfrak{B}_1}$$

由$\mathfrak{B}_1$为V的基和向量$\mathbf{x}$的任意性, 得到

$$[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_1} = \mathbf{M}^{-1}[\sigma(\mathfrak{B}_2)]_{\mathfrak{B}_2}\mathbf{M}$$

定义 5.6. *设* $\mathbf{A},\mathbf{B}\in F^{n\times n}$, *若存在可逆矩阵* $\mathbf{P}\in F^{n\times n}$ *使得*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{B} \tag{5.9}$$

则称矩阵 $\mathbf{A}$ *相似于* $\mathbf{B}$, *记为* $\mathbf{A}\sim\mathbf{B}$

将相似视作 $F^{n\times n}$ 上矩阵之间的关系, 它具有下述性质: $\forall\mathbf{A},\mathbf{B},\mathbf{C}\in F^{n\times n}$

(1) 自反性, 设 $\mathbf{E}$ 为 F 上的单位矩阵, 使得 $\mathbf{E}^{-1}\mathbf{A}\mathbf{E}=\mathbf{A}$, 即 $\mathbf{A}\sim\mathbf{A}$.

(2) 对称性, 若 $\mathbf{A}\sim\mathbf{B}$, 则存在可逆矩阵 $\mathbf{P}$, 使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{B}$, 而 $\mathbf{A}=\left(\mathbf{P}^{-1}\right)^{-1}\mathbf{B}\mathbf{P}^{-1}$, 即 $\mathbf{B}\sim\mathbf{A}$.

(3) 传递性, 若 $\mathbf{A}\sim\mathbf{B}$, $\mathbf{B}\sim\mathbf{C}$, 则存在可逆矩阵 $\mathbf{P},\mathbf{Q}$, 使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{B}$ 和 $\mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}=\mathbf{C}$, 则 $(\mathbf{PQ})^{-1}\mathbf{A}(\mathbf{PQ})=\mathbf{C}$, 即 $\mathbf{A}\sim\mathbf{C}$.

因此矩阵的相似关系是 $F^{n\times n}$ 上的一种等价关系. 如果不考虑线性变换的基的选择, 则一个线性变换对应 $F^{n\times n}$ 中的一个相似等价类.

例 9. *设* $\mathfrak{B}_1=\{\varepsilon_1,\varepsilon_2,\varepsilon_3\}$ *和* $\mathfrak{B}_2=\{\eta_1,\eta_2,\eta_3\}$ *为线性空间* $\mathbb{R}^3$ *的两个基, 从* $\mathfrak{B}_1$ *到* $\mathfrak{B}_2$ *的过渡矩阵* $\mathbf{M}$ *为*

$$\mathbf{M}=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix},\quad \text{即 } \mathfrak{B}_2=\mathfrak{B}_1\mathbf{M}$$

$\mathbb{R}^3$ *上线性变换* σ *在基* $\mathfrak{B}_1$ *的矩阵为*

$$[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_1}=\begin{bmatrix}1&0&-1\\3&2&0\\-1&2&3\end{bmatrix}$$

求 σ *在基* $\mathfrak{B}_2$ *下对应的矩阵.*

解: 因 $[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_1}=\mathbf{M}[\sigma(\mathfrak{B}_2)]_{\mathfrak{B}_2}\mathbf{M}^{-1}$, 所以

$$[\sigma(\mathfrak{B}_2)]_{\mathfrak{B}_2}=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}^{-1}\begin{bmatrix}1&0&-1\\3&2&0\\-1&2&3\end{bmatrix}\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}=\begin{bmatrix}0&-1&-1\\5&3&1\\-1&3&3\end{bmatrix}$$

例 10. *平面上的旋转变换在标准正交基* $\mathbf{B}_1=\left\{\vec{i},\vec{j}\right\}$ *下的矩阵如式(5.7)所示, 基* $\mathfrak{B}_2=\left\{\vec{a},\vec{b}\right\}$, *其中* $\vec{a}=\vec{i}+\vec{j}$, $\vec{b}=\vec{i}-2\vec{j}$, *求旋转变换在基* $\mathfrak{B}_2$ *下的矩阵.*

解: 根据题意有: $\mathfrak{B}_2=\mathfrak{B}_1\mathbf{M}$, 其中

$$\mathbf{M}=\begin{bmatrix}1&1\\1&-2\end{bmatrix}$$

则

$$\begin{aligned}[\sigma_\theta(\mathfrak{B}_2)]_{\mathfrak{B}_2}&=\mathbf{M}^{-1}[\sigma_\theta(\mathfrak{B}_2)]_{\mathfrak{B}_2}\mathbf{M}=\frac{1}{3}\begin{bmatrix}2&1\\1&-1\end{bmatrix}\begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix}\begin{bmatrix}1&1\\1&-2\end{bmatrix}\\&=\frac{1}{3}\begin{bmatrix}3\cos\theta-\sin\theta&5\sin\theta\\-2\sin\theta&3\cos\theta+\sin\theta\end{bmatrix}\end{aligned}$$

上述建立了同一变换在不同基下的矩阵表示之间的相似关系. 反之, 两个相似的矩阵也可以理解为同一线性变换在不同基下的表示. 对于一个线性变换 σ, 如何寻找适当的基, 使得它在这组基下的矩阵表示具有间单的形式, 这将是下一节讨论的问题.

5.3 特征值和特征向量

5.3.1 不变子空间

利用线性变换可生成子空间, 如线性变换的核和像. 反之, 与子空间相关的线性变换也会提供一些重要信息, 例如: 哪些子空间在线性变换下的 具有不变性. 下面先给出不变子空间的定义.

定义 5.7. *设σ是线性空间V上的线性变换, S 是V的子空间, 若S在变换σ下的像 $\sigma(S)\subseteq S$, 则称 S 是σ-***不变子空间***.*

例 11. *线性变换σ的核$Ker(\sigma)$ 和像$\mathrm{Im}(\sigma)$ 是 σ-不变子空间.*

例 12. *设σ是V上的数乘变换, 即存在常数c, $\sigma(\mathbf{x})=c\mathbf{x}$, 则$V$ 的任意子空间均是σ-不变子空间.*

定理 5.3.1. *S是n维线性空间V上线性变换σ的不变子空间, 设S的基为$\mathfrak{B}_S=\{\varepsilon_1,\varepsilon_2,\dots,\varepsilon_r\}(0<r<n)$, 将它扩充成$V$的基$\mathfrak{B}_V=\{\varepsilon_1,\varepsilon_2,\dots,\varepsilon_r,\varepsilon_{r+1},\dots,\varepsilon_n\}=\left[\begin{array}{c|c}\mathfrak{B}_S & \mathfrak{B}_E\end{array}\right]$, 则$\sigma$在基$\mathfrak{B}_V$下的矩阵具有下列形状:*

$$\left[\begin{array}{cccc|ccc} a_{11} & a_{12} & \cdots & a_{1r} & a_{1(r+1)} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} & a_{r(r+1)} & \cdots & a_{rn} \\ \hline 0 & 0 & \cdots & 0 & a_{(r+1)(r+1)} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & a_{n(r+1)} & \cdots & a_{nn} \end{array}\right] \tag{5.10}$$

证: 因S是σ的不变子空间, 则

$$\begin{aligned} &\left[\begin{array}{cccc}\sigma(\varepsilon_1) & \sigma(\varepsilon_2) & \cdots & \sigma(\varepsilon_r)\end{array}\right]=\mathfrak{B}_S\mathbf{A}_S \\ &\left[\begin{array}{ccc}\sigma(\varepsilon_{r+1}) & \cdots & \sigma(\varepsilon_r)\end{array}\right]=\left[\begin{array}{c|c}\mathfrak{B}_S & \mathfrak{B}_E\end{array}\right]\left[\begin{array}{c}\mathbf{A}_1 \\ \hline \mathbf{A}_2\end{array}\right] \end{aligned}$$

其中

$$\mathbf{A}_s=\left[\begin{array}{ccc} a_{11} & \cdots & a_{1r} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rr}\end{array}\right],\mathbf{A}_1=\left[\begin{array}{ccc} a_{1(r+1)} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{r(r+1)} & \cdots & a_{rn}\end{array}\right],\mathbf{A}_2=\left[\begin{array}{ccc} a_{(r+1)(r+1)} & \cdots & a_{(r+1)n} \\ \cdots & \cdots & \cdots \\ a_{n(r+1)} & \cdots & a_{nn}\end{array}\right],$$

则

$$\left[\begin{array}{cccccc}\sigma(\varepsilon_1) & \cdots & \sigma(\varepsilon_r) & \sigma(\varepsilon_{r+1}) & \cdots & \sigma(\varepsilon_n)\end{array}\right]=\left[\begin{array}{c|c}\mathfrak{B}_S & \mathfrak{B}_E\end{array}\right]\left[\begin{array}{cc}\mathbf{A}_S & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{A}_2\end{array}\right]$$

反之, 若矩阵具有式(5.10)形式, 因矩阵与线性变换之间一一对应, 则与该矩阵对应的线性变换必存在不变子空间.

推论 1. *设$V=V_1\oplus V_2$且V_1,V_2都是线性变换σ的不变子空间, 若V_1的基为$\mathfrak{B}_1=\{\alpha_1,\alpha_2,\dots,\alpha_r\}$, V_2的基为$\mathfrak{B}_2=\{\beta_1,\beta_2,\dots,\beta_{n-r}\}$, 则$\sigma$ 在基$\mathfrak{B}=\{\mathfrak{B}_1,\mathfrak{B}_2\}$ 下的矩阵形式为*

$$\left[\begin{array}{cc}\mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2\end{array}\right] \tag{5.11}$$

类似定理5.3.1即可证明.

例 13. *σ为$\mathbb{R}^3$的线性变换, 在基$\mathfrak{B}=\{\varepsilon_1,\varepsilon_2,\varepsilon_3\}$下的变换矩阵为$\left[\begin{array}{ccc}3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0\end{array}\right]$, 令$\beta_1=\varepsilon_3,\beta_2=\varepsilon_1+\varepsilon_2+2\varepsilon_3$, 且 $S=\mathrm{span}\{\beta_1,\beta_2\}$. 求证: S是变换σ的不变子空间.*

证: 易证, β_1,β_2线性无关, 构成S的一组基.

$$\begin{aligned} \sigma(\beta_1)=\mathfrak{B}[\sigma(\mathfrak{B})]_{\mathfrak{B}}[\beta_1]_{\mathfrak{B}}=\mathfrak{B}\left[\begin{array}{ccc}3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0\end{array}\right]\left[\begin{array}{c}0 \\ 0 \\ 1\end{array}\right]=\mathfrak{B}\left[\begin{array}{c}-1 \\ -1 \\ 0\end{array}\right]=2\beta_1-\beta_2\in S \\ \sigma(\beta_2)=\mathfrak{B}[\sigma(\mathfrak{B})]_{\mathfrak{B}}[\beta_2]_{\mathfrak{B}}=\mathfrak{B}\left[\begin{array}{ccc}3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0\end{array}\right]\left[\begin{array}{c}1 \\ 1 \\ 2\end{array}\right]=\mathfrak{B}\left[\begin{array}{c}2 \\ 2 \\ 4\end{array}\right]=2\beta_2\in S \end{aligned}$$

对$\forall\mathbf{x}\in S,\mathbf{x}=x_1\beta_1+x_2\beta_2$, 则

$$\sigma(\mathbf{x})=x_1\sigma(\beta_1)+x_2\sigma(\beta_2)=2x_1\beta_1+(2x_2-x_1)\beta_2\in S$$

因此S是σ-不变子空间.

下面讨论在给定线性变换对应的矩阵后, 如何寻找它的不变子空间.

5.3.2 特征值与特征向量

后续的讨论中若无特殊交代，数域F取实数域$\mathbb{R}$. 设n维欧式空间$\mathbb{R}^n$的某个线性变换在选定基下(如自然基$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$)对应的矩阵为$\mathbf{A}\in\mathbb{R}^{n\times n}$, 若$S$是$\sigma$-不变子空间, 则对$\forall\mathbf{x}\in S$, 有 $\sigma(\mathbf{x})\in S$. 特别地当$\dim S=1$时, 有 $\sigma(\mathbf{x})=\lambda\mathbf{x}$, 其中$\lambda\in\mathbb{R}$. 关于如何求线性变换的不变子空间有如下定义.

定义 5.8. *设σ是数域F上线性空间V的一个线性变换, 存在某个数$\lambda\in F$和$\mathbf{x}\in V$, 且$\mathbf{x}\neq\mathbf{0}$, 使得*

$$\sigma(\mathbf{x})=\lambda\mathbf{x} \tag{5.12}$$

*则称λ为线性变换σ的一个**特征值**, 而称$\mathbf{x}$为关于特征值λ的**特征向量**.*

将σ关于特征值λ的所有特征向量和零向量构成的集合记为V_λ, 即

$$V_\lambda=\left\{\mathbf{x}\middle|\mathbf{x}\in V\wedge\sigma(\mathbf{x})=\lambda\mathbf{x}\right\} \tag{5.13}$$

则关于V_λ有下列性质:

定理 5.3.2. *由式(5.13)定义的V_λ是V的线性子空间, 而且是σ-不变子空间.*

证: 考察线性空间V的加法与数乘在V_λ上的封闭性, 即对$\forall\mathbf{x},\mathbf{y}\in V_\lambda$和$c\in F$, 有

$$\begin{aligned}\sigma(\mathbf{x}+\mathbf{y})&=\sigma(\mathbf{x})+\sigma(\mathbf{y})=\lambda(\mathbf{x}+\mathbf{y})\in V_\lambda\\ \sigma(c\mathbf{x})&=c\sigma(\mathbf{x})=\lambda(c\mathbf{x})\in V_\lambda\end{aligned}$$

可见V_λ中元素关于加法和数乘是封闭的, 因此它是V的子空间. 另外, V_λ中任意向量在变换σ下的像还是属于V_λ, 即V_λ是σ-不变子空间.

通常将V_λ 称为线性变换σ的**特征子空间**.

例 14. *例5中的镜像变换σ, 其中平面L中的点构成σ对应特征值为1的不变子空间. 而过原点与法向$\vec{n}$平行的直线上的点构成特征值为-1的σ-不变子空间.*

接下来讨论: 给定一个线性变换, 如何求它的特征值和特征向量?

根据本章第2节的结论, 当n维线性空间V选定一个基后, V上的线性变换与矩阵一一对应, 线性变换σ 在基$\mathfrak{B}=\{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\}$下对应矩阵为$[\sigma(\mathfrak{B})]_{\mathfrak{B}}$, 根据特征值和特征向量的定义, 有

$$\mathfrak{B}\left[\sigma(\mathfrak{B})\right]_{\mathfrak{B}}[\mathbf{x}]_{\mathfrak{B}}=\lambda\mathfrak{B}[\mathbf{x}]_{\mathfrak{B}}$$

为方便起见, 令$\mathbf{A}=[\sigma(\mathfrak{B})]_{\mathfrak{B}}$, 而特征向量在基$\mathfrak{B}$下的坐标用向量$\mathbf{X}$表示, 即$\mathbf{X}=[\mathbf{x}]_{\mathfrak{B}}$, 由此得到特征值、特征向量坐标必须满足下列方程

$$\mathbf{AX}=\lambda\mathbf{X} \tag{5.14}$$

移项后得:

$$(\lambda\mathbf{E}-\mathbf{A})\mathbf{X}=\mathbf{0},\quad \mathbf{X}\neq\mathbf{0} \tag{5.15}$$

由齐次线性方程组的解理论知, 方程组(5.15)有非零解$\mathbf{x}\neq\mathbf{0}$, 当且仅当 $\operatorname{rank}(\lambda\mathbf{E}-\mathbf{A})<n$ 或

$$\det(\lambda\mathbf{E}-\mathbf{A})=0 \tag{5.16}$$

也就是若λ是线性变换的特征值, 则它必须满足方程(5.16). 将λ视作未知量, 用$f(\lambda)$ 表示行列式 $\det(\lambda\mathbf{E}-\mathbf{A})$, 则称$f(\lambda)$为矩阵$\mathbf{A}$的**特征多项式**, 而 $f(\lambda)=0$的根称为**特征根**, 它们就是矩阵$\mathbf{A}$的特征值.

关于特征多项式$f(\lambda)$有下述性质:

(1) 线性变换σ的特征多项式$f(\lambda)$与基的选择无关.

证: 因线性变换在不同基下对应矩阵之间的有相似关系(定理5.2.3), 设σ在两组不同基下对应的矩阵分别为$\mathbf{A}$ 和$\mathbf{B}$, 则存在一个可逆矩阵(两个基之间的过渡矩阵)$\mathbf{P}$使得 $\mathbf{P}^{-1}\mathbf{AP}=\mathbf{B}$, 则

$$\begin{aligned}f_{\mathbf{B}}(\lambda)&=\det(\lambda\mathbf{E}-\mathbf{B})=\det\left(\lambda\mathbf{P}^{-1}\mathbf{P}-\mathbf{P}^{-1}\mathbf{AP}\right)=\det\left[\mathbf{P}^{-1}(\lambda\mathbf{E}-\mathbf{A})\mathbf{P}\right]\\&=\det\left(\mathbf{P}^{-1}\right)\det(\mathbf{P})\det(\lambda\mathbf{E}-\mathbf{A})=\det(\lambda\mathbf{E}-\mathbf{A})=f_{\mathbf{A}}(\lambda)\end{aligned}$$

(2) 利用行列式展开法,可知特征多项式的展开式形如:

$$f(\lambda) = \lambda^n + (-1)^1 p_1 \lambda^{n-1} + \cdots + (-1)^{n-1} p_{n-1}\lambda + (-1)^n p_n \tag{5.17}$$

其中 p_i是矩阵$\mathbf{A}$的所有i阶主子式之和($i = 1, 2, \ldots, n$). 因此

$$p_1 = \sum_{i=1}^{n} a_{ii} = \mathrm{trace}\,(\mathbf{A})$$
$$p_n = \det\,(\mathbf{A})$$

根据多项式根相关理论可知, 在复数域内n次多项式恰含n个根(包括重根), 设$f(\lambda)$的n个根分别为: $\lambda_1, \lambda_2, \ldots, \lambda_n$, 则根据根与系数的关系有

(I) $\displaystyle\sum_{i=1}^{n} \lambda_i = p_1 = \mathbf{A}\text{的1阶主子式的和} = \mathrm{trace}\,(\mathbf{A})$

(II) $\displaystyle\prod_{i=1}^{n} \lambda_i = p_n = \det\,(\mathbf{A})$

所有这些特征值也称为矩阵$\mathbf{A}$的**特征谱**或**谱**, 假设特征多项式互不相同的根有$m(1 \le m \le n)$个, 为$\lambda_1, \lambda_2, \ldots, \lambda_m$, 则特征多项式$f(\lambda)$可分解成如下因式的乘积:

$$f\,(\lambda) = \prod_{i=1}^{m} (\lambda - \lambda_i)^{n_i}\,, \quad \sum_{i=1}^{m} n_i = n$$

其中$n_i(i = 1, 2, \ldots, m)$称为根λ_i的**代数重数**, 若$n_i = 1$相应的根称为**非退化特征值**, 若 $n_i > 1$则相应的根称为**退化特征值**.

在求得特征多项式的根(特征根或特征值)以后, 再将特征值λ_i代入方程(5.15), 得

$$(\lambda_i \mathbf{E} - \mathbf{A})\,\mathbf{X} = \mathbf{0}$$

此时, 因$\det\,(\lambda_i \mathbf{E} - \mathbf{A}) = 0$, 所以方程必有非零解, 它的基础解系的秩等于核空间 $Ker\,(\lambda_i \mathbf{E} - \mathbf{A})$的维数, 通常将$\lambda_i$对应特征方程组的基础解系的秩或 $\dim Ker\,(\lambda_i \mathbf{E} - \mathbf{A})$称为$\lambda_i$的**几何重数**. 特征值的几何重数和代数重数之间成立下列定理:

定理 5.3.3. *设λ_0是n阶方阵$\mathbf{A}$的任一特征值, 则λ_0的几何重数不超过它的代数重数. 即*

$$1 \le \dim Ker\,(\lambda_0 \mathbf{E} - \mathbf{A}) \le \lambda_0\text{的代数重数} \tag{5.18}$$

证: 设特征方程组$(\lambda_0 \mathbf{E} - \mathbf{A})\,\mathbf{X} = \mathbf{0}$的基础解系为 $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k$,即$\lambda_0$的几何重数为$k$, 不妨设它们相互正交且标准化, 即 $(\mathbf{X}_i, \mathbf{X}_j) = \begin{cases} 1, & i = j \\ 0, & i \ne j \end{cases}$, 否则可通过Gram-Schmidt过程将它们正交标准化, 将这组基础解系扩充成n维空间的一组标准正交基, 令矩阵 $\mathbf{\Psi}$ 由这些标准正交基为列向量构成:

$$\mathbf{\Psi} = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k, \mathbf{X}_{k+1}, \ldots, \mathbf{X}_n]$$

因此 $\mathbf{\Psi}$是正交矩阵(或酉矩阵), 即$\mathbf{\Psi}^{-1} = \mathbf{\Psi}^T$(酉矩阵时用共轭转置), 进一步将 $\mathbf{\Psi}$ 分块成:

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{\Psi}_1 & \mathbf{\Psi}_2 \end{bmatrix}$$
$$\text{其中} \quad \mathbf{\Psi}_1 = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_r \end{bmatrix}, \qquad \mathbf{\Psi}_2 = \begin{bmatrix} \mathbf{X}_{r+1} & \mathbf{X}_{r+2} & \cdots & \mathbf{X}_n \end{bmatrix}$$

则

$$\mathbf{B} = \mathbf{\Psi}^T \mathbf{A} \mathbf{\Psi} = \begin{bmatrix} \mathbf{\Psi}_1 & \mathbf{\Psi}_2 \end{bmatrix}^T \mathbf{A} \begin{bmatrix} \mathbf{\Psi}_1 & \mathbf{\Psi}_2 \end{bmatrix} = \begin{bmatrix} \lambda_0 \mathbf{E}_k & \mathbf{\Psi}_1^T \mathbf{A} \mathbf{\Psi}_2 \\ \mathbf{0} & \mathbf{\Psi}_2^T \mathbf{A} \mathbf{\Psi}_2 \end{bmatrix}$$

显然, 矩阵$\mathbf{A}$与$\mathbf{B}$相似, 它们的特征多项式相同, 即

$$\det\,(\lambda \mathbf{E} - \mathbf{A}) = \det\,(\lambda \mathbf{E} - \mathbf{B}) = (\lambda - \lambda_0)^k \det\left(\lambda \mathbf{E} - \mathbf{\Psi}_2^T \mathbf{A} \mathbf{\Psi}_2\right)$$

由上式可见λ_0的代数重数至少是k(λ_0的几何重数).

例 15. *设σ为$\mathbb{R}^2 \mapsto \mathbb{R}^2$的线性变换, 定义为*

$$\sigma\left(\left[\begin{array}{c} x \\ y \end{array}\right]\right)=\left[\begin{array}{c} 2x-3y \\ x+6y \end{array}\right]$$

选择$\mathbb{R}^2$的两组基$\mathfrak{B}_1, \mathfrak{B}_2$, 其中

$$\mathfrak{B}_1=\left\{\left[\begin{array}{c} 1 \\ 0 \end{array}\right],\left[\begin{array}{c} 0 \\ 1 \end{array}\right]\right\}, \quad \mathfrak{B}_2=\left\{\left[\begin{array}{c} 1 \\ 1 \end{array}\right],\left[\begin{array}{c} 0 \\ 1 \end{array}\right]\right\}$$

(a) 分别求线性变换σ在基$\mathfrak{B}_1$,$\mathfrak{B}_2$下对应矩阵的特征值和特征向量;

(b) 求σ的特征子空间.

解: (1) 确定两组基下的变换矩阵

$$\mathbf{A}=\left[\sigma\left(\mathfrak{B}_1\right)\right]_{\mathfrak{B}_1}=\left[\begin{array}{cc} 2 & -3 \\ 1 & 6 \end{array}\right], \quad \mathbf{B}=\left[\sigma\left(\mathfrak{B}_2\right)\right]_{\mathfrak{B}_2}=\left[\begin{array}{cc} -1 & -3 \\ 8 & 9 \end{array}\right]$$

则特征多项式

$$f_{\mathbf{A}}(\lambda)=f_{\mathbf{B}}(\lambda)=\det(\lambda\mathbf{E}-\mathbf{A})=\lambda^2-8\lambda+15$$

特征根为$\lambda_1=3, \lambda_2=5$.

$$\begin{array}{llll}
\text{特征值} & \text{特征方程} & & \text{特征向量} \\
\lambda_1=3: & [\lambda_1\mathbf{E}-\mathbf{A}]\mathbf{X}=\mathbf{0} \Rightarrow & \left[\begin{array}{rr} 1 & 3 \\ -1 & -3 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right]=\mathbf{0} & \mathbf{X}_{\lambda_1}^{\mathbf{A}}=\left[\begin{array}{r} 3 \\ -1 \end{array}\right] \\
 & [\lambda_1\mathbf{E}-\mathbf{B}]\mathbf{X}=\mathbf{0} \Rightarrow & \left[\begin{array}{rr} 4 & 3 \\ -8 & -6 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right]=\mathbf{0} & \mathbf{X}_{\lambda_1}^{\mathbf{B}}=\left[\begin{array}{r} 3 \\ -4 \end{array}\right] \\
\lambda_2=5: & [\lambda_2\mathbf{E}-\mathbf{A}]\mathbf{X}=\mathbf{0} \Rightarrow & \left[\begin{array}{rr} 3 & 3 \\ -1 & -1 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right]=\mathbf{0} & \mathbf{X}_{\lambda_2}^{\mathbf{A}}=\left[\begin{array}{r} 1 \\ -1 \end{array}\right] \\
 & [\lambda_2\mathbf{E}-\mathbf{B}]\mathbf{X}=\mathbf{0} \Rightarrow & \left[\begin{array}{rr} 6 & 3 \\ -8 & -4 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right]=\mathbf{0} & \mathbf{X}_{\lambda_2}^{\mathbf{B}}=\left[\begin{array}{r} 1 \\ -2 \end{array}\right]
\end{array}$$

(2) σ的不变子空间:

$$V_{\lambda_1}=\left\{c\mathfrak{B}_1\mathbf{X}_{\lambda_1}^{\mathbf{A}} \middle| c\in\mathbb{R}\right\}=\left\{c\mathfrak{B}_2\mathbf{X}_{\lambda_1}^{\mathbf{B}} \middle| c\in\mathbb{R}\right\}=\left\{c\left[\begin{array}{r} 3 \\ -1 \end{array}\right] \middle| c\in\mathbb{R}\right\}$$

$$V_{\lambda_2}=\left\{c\mathfrak{B}_1\mathbf{X}_{\lambda_2}^{\mathbf{A}} \middle| c\in\mathbb{R}\right\}=\left\{c\mathfrak{B}_2\mathbf{X}_{\lambda_2}^{\mathbf{B}} \middle| c\in\mathbb{R}\right\}=\left\{c\left[\begin{array}{r} 1 \\ -1 \end{array}\right] \middle| c\in\mathbb{R}\right\}$$

由此例可见, 特征多项式、特征值及变换的不变子空间与基的选择无关, 但不同基下(坐标)特征向量与基的选择有关.

例 16. *求矩阵*

$$\mathbf{A}=\left[\begin{array}{rrr} 2 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 3 & 1 \end{array}\right]$$

的特征值和特征向量.

解:特征多项式: $f(\lambda)=\det\begin{bmatrix}\lambda-2 & -1 & 0\\ 1 & \lambda & -1\\ -1 & -3 & \lambda-1\end{bmatrix}=\lambda^3-3\lambda^2+4=(\lambda+1)(\lambda-2)^2$, 求得特征根为 $\lambda_1=-1,\lambda_2=\lambda_3=2$. 将$\lambda_1$代入特征方程得

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

求得基础解系(特征向量)为:$\begin{bmatrix}1 & -3 & 4\end{bmatrix}^T$. 将$\lambda_{2,3}$代入特征方程组得

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

求得基础解系(特征向量)为:$\begin{bmatrix}1 & 0 & 1\end{bmatrix}^T$. 特征值2的代数重数为2, 但它的几何重数为1, 验证了不等式(5.18).

定理 5.3.4. *设σ是n维线性空间V的线性变换, λ_1,λ_2是σ的两个不同特征值, 则σ的两个特征子空间$V_{\lambda_1},V_{\lambda_2}$的交为$\{\mathbf{0}\}$, 即$V_{\lambda_1}\cap V_{\lambda_2}=\{\mathbf{0}\}$.*

证: 设向量$\mathbf{x}\in V_{\lambda_1}\cap V_{\lambda_2}$, 则

$$\sigma(\mathbf{x})=\lambda_1\mathbf{x}\in V_{\lambda_1},\quad 且\ \sigma(\mathbf{x})=\lambda_2\mathbf{x}\in V_{\lambda_2}$$
$$\mathbf{0}=\sigma(\mathbf{0})=\sigma(\mathbf{x}-\mathbf{x})=\lambda_1\mathbf{x}-\lambda_2\mathbf{x}=(\lambda_1-\lambda_2)\mathbf{x}$$

因$\lambda_1\neq\lambda_2$, 所以$\lambda_1-\lambda_2\neq 0$, 只有 $\mathbf{x}=\mathbf{0}$.

定理结论可推广到k个不同特征值的特征子空间的两两交集均为$\{\mathbf{0}\}$. 即

定理 5.3.5. *设σ是n维线性空间V的线性变换, $\lambda_1,\lambda_2,\dots,\lambda_k$是$\sigma$的$k$个不同特征值, 则$\sigma$的特征子空间$V_{\lambda_1},V_{\lambda_2},\dots,V_{\lambda_k}$之间满足关系*

$$V_{\lambda_i}\cap V_{\lambda_j}=\{\mathbf{0}\},\quad i\neq j=1,2,\dots,k$$

证明同定理5.3.4的.

推论 2. *属于不同特征值的特征向量线性无关.*

推论 3. *设$\lambda_1,\lambda_2,\dots,\lambda_k$为线性变换$\sigma$不同的特征值, 且$\lambda_i$的几何重数为$r_i$ ($i=1,2,\dots,k$), 对应特征值λ_i的特征向量$\mathbf{X}_{i1},\mathbf{X}_{i2},\dots,\mathbf{X}_{ir_i}$ 线性无关(基础解系性质), 则特征向量*

$$\mathbf{X}_{11},\mathbf{X}_{12},\dots,\mathbf{X}_{1r_1},\mathbf{X}_{21},\dots,\mathbf{X}_{2r_2},\dots,\mathbf{X}_{k1},\dots,\mathbf{X}_{kr_k}$$

线性无关.

5.3.3 Hamilton Cayley定理

设σ为n维线性空间V上的线性变换, $\mathbf{x}$为V中任一取定的非零向量, 对它实施多次变换后得$\mathbf{x},\sigma(\mathbf{x}),\sigma^2(\mathbf{x}),\sigma^3(\mathbf{x}),\dots,\sigma^n(\mathbf{x})$ 共$n+1$个向量, 这组向量一定线性相关, 即存在不全为零的一组数使得下式成立:

$$c_0\mathbf{x}+c_1\sigma(\mathbf{x})+\cdots+c_n\sigma^n(\mathbf{x})=\mathbf{0}$$

或即写成

$$(c_0\mathbf{1}_V+c_1\sigma+\cdots+c_n\sigma^n)(\mathbf{x})=\mathbf{0}$$

其中$\mathbf{1}_V=\sigma^0$表示V上的恒等变换. 设多项式为

$$f(\lambda)=c_0+c_1\lambda+\cdots+\lambda^n$$

则 $f(\sigma)(\mathbf{x})=\mathbf{0}$, 这说明对取定的向量$\mathbf{x}$总存在某个多项式 $f(\lambda)$, 使得$f(\sigma)(\mathbf{x})=0$, 若将$\mathbf{x}$依次取V 的一个基$\{\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n\}$的每个基向量, 则得到关于每个基向量 ε_i 的一个多项式$f_i(\lambda)$ 使得 $f_i(\sigma)(\varepsilon_i)=\mathbf{0}$, 即

$$f_1(\lambda),f_2(\lambda),\dots,f_n(\lambda)$$

且 $f_i(\sigma) f_j(\sigma) = f_j(\sigma) f_i(\sigma)$, 若令

$$g(\sigma) = \prod_{i=1}^{n} f_i(\sigma)$$

则

$$g(\sigma)(\varepsilon_i) = \mathbf{0}, \quad i = 1, 2, \ldots, n$$

上式中$g(\sigma)$ 对V中一切向量作用后均为零, 这说明对V上的任意线性变换σ总存在一个多项式 $g(\lambda)$使得$g(\sigma)$等价于$\mathbf{0}_V$. 因线性变换与矩阵之间一一对应, 所以对任意n阶矩阵$\mathbf{A}$ 同样存在一个多项式$g(\lambda)$使得$g(\mathbf{A}) = \mathbf{0}$, 通常称使得$g(\mathbf{A}) = \mathbf{0}$的多项式$g(\lambda)$为$\mathbf{A}$的**化零多项式**.

定义 5.9. *对于任意n阶矩阵$\mathbf{A}$, 存在一个首项系数为1 且次数最小的$\mathbf{A}$化零多项式 $m(\lambda)$, 称为$\mathbf{A}$的***最小多项式**.

定理 5.3.6. *若$f(\lambda)$是矩阵$\mathbf{A}$的化零多项式, 则$\mathbf{A}$的最小多项式$m(\lambda)$是$f(\lambda)$的因子.*

证: 用$\deg f(\lambda)$表示多项式中λ的最高次幂的次数,显然$\deg f(\lambda) \geq \deg m(\lambda)$, 因此对$f(\lambda)$可唯一分解为

$$f(\lambda) = m(\lambda) q(\lambda) + r(\lambda)$$

其中$\deg r(\lambda) < \deg m(\lambda)$, 若$r(\lambda) \neq 0$ 则与$m(\lambda)$是最小多项式矛盾.

定理 5.3.7. *任意n阶矩阵的最小多项式唯一.*

证: 假设存在两个最小多项式, 根据定理5.3.6它们相互为因子, 且它们又都是首一的多项式, 则它们必相等.

例 17. *设矩阵*$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$, *多项式 $f(\lambda) = \lambda^2 - 4\lambda + 4$ 为$\mathbf{A}$的化零多项式, 因$f(\mathbf{A}) = \mathbf{0}$, 而且不存在次数比 $f(\lambda)$更低的$\mathbf{A}$的(非零)化零多项式, 因此$f(\lambda)$是$\mathbf{A}$的最小多项式.*

定理 5.3.8. *相似矩阵具有相同的最小多项式.*

证: 设矩阵 $\mathbf{A}, \mathbf{B}$相似, 则存在可逆矩阵$\mathbf{P}$, 使得$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$. 同时, 设 $\mathbf{A}, \mathbf{B}$的最小多项式分别为$m(\lambda), n(\lambda)$, 则

$$\mathbf{0} = n(\mathbf{B}) = n\left(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right) = \mathbf{P}^{-1} n(\mathbf{A}) \mathbf{P} = n(\mathbf{A})$$

则 $m(\lambda) \big| n(\lambda)$. 同理, 由$\mathbf{A} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1}$ 可得 $n(\lambda) \big| m(\lambda)$, 因此$m(\lambda) = n(\lambda)$.

定理 5.3.9. *(Hamilton-Cayley定理) 设$\mathbf{A}$是数域F上的n阶方阵, $f(\lambda) = \det(\lambda\mathbf{E} - \mathbf{A})$为矩阵$\mathbf{A}$的特征多项式, 则$f(\lambda)$ 是$\mathbf{A}$的化零多项式, 即*

$$f(\mathbf{A}) = \mathbf{0} \tag{5.19}$$

证: 设$\mathbf{B}(\lambda)$ 为矩阵$\lambda\mathbf{E}-\mathbf{A}$的伴随矩阵, 根据伴随矩阵的性质有

$$\mathbf{B}(\lambda)(\lambda\mathbf{E}-\mathbf{A})=\det(\lambda\mathbf{E}-\mathbf{A})\mathbf{E} \tag{5.20}$$

将矩阵$\mathbf{B}(\lambda)$ 展开成关于λ系数为常数矩阵的多项式, 且由式(5.20)知 $\mathbf{B}(\lambda)$ 中λ最高次不超过$n-1$, 即

$$\mathbf{B}(\lambda)=\mathbf{B}_0+\mathbf{B}_1\lambda+\cdots+\mathbf{B}_{n-1}\lambda^{n-1}$$

利用特征多项式的展开式(5.17)代入等式(5.20), 分别比较λ的各次幂前的系数矩阵,有

$$\begin{array}{rrcl}
\lambda^n: & \mathbf{B}_{n-1}\mathbf{E} & = & \mathbf{E}\\
\lambda^{n-1}: & \mathbf{B}_{n-2}-\mathbf{B}_{n-1}\mathbf{A} & = & (-1)^1p_1\mathbf{E}\\
\lambda^{n-2}: & \mathbf{B}_{n-3}-\mathbf{B}_{n-2}\mathbf{A} & = & (-1)^2p_2\mathbf{E}\\
 & & \cdots\cdots\cdots & \\
\lambda^{j}: & \mathbf{B}_{j-1}-\mathbf{B}_{j}\mathbf{A} & = & (-1)^{n-j}p_{n-j}\mathbf{E}\\
\lambda^{j-1}: & \mathbf{B}_{j-2}-\mathbf{B}_{j-1}\mathbf{A} & = & (-1)^{n-j+1}p_{n-j+1}\mathbf{E}\\
 & & \cdots\cdots\cdots & \\
\lambda: & \mathbf{B}_0-\mathbf{B}_1\mathbf{A} & = & (-1)^{n-1}p_{n-1}\\
\lambda^0: & -\mathbf{B}_0\mathbf{A} & = & (-1)^np_n\mathbf{E}
\end{array}$$

按上式中前面λ的幂次, 在等式两端乘上$\mathbf{A}$的相应幂次($j=n,n-1,\ldots,2,1$) 后得

$$\begin{array}{rcrcl}
 & & \mathbf{B}_{n-1}\mathbf{A}^n & = & \mathbf{A}^n\\
\mathbf{B}_{n-2}\mathbf{A}^{n-1} & - & \mathbf{B}_{n-1}\mathbf{A}^n & = & (-1)^1p_1\mathbf{A}^{n-1}\\
\mathbf{B}_{n-3}\mathbf{A}^{n-2} & - & \mathbf{B}_{n-2}\mathbf{A}^{n-1} & = & (-1)^2p_2\mathbf{A}^{n-2}\\
 & \cdots\cdots\cdots & & & \\
\mathbf{B}_{j-1}\mathbf{A}^{j} & - & \mathbf{B}_{j}\mathbf{A}^{j+1} & = & (-1)^{n-j}p_{n-j}\mathbf{A}^j\\
\mathbf{B}_{j-2}\mathbf{A}^{j-1} & - & \mathbf{B}_{j-1}\mathbf{A}^{j} & = & (-1)^{n-j+1}p_{n-j+1}\mathbf{A}^{j-1}\\
 & \cdots\cdots\cdots & & & \\
\mathbf{B}_0\mathbf{A} & - & \mathbf{B}_1\mathbf{A}^2 & = & (-1)^{n-1}p_{n-1}\mathbf{A}\\
 & & -\mathbf{B}_0\mathbf{A} & = & (-1)^np_n\mathbf{E}
\end{array}$$

对上述这些等式的两端求和后得: $\mathbf{0}=f(\mathbf{A})$.

定理5.3.9说明矩阵$\mathbf{A}$的特征多项式是$\mathbf{A}$的化零多项式, 但它不一定是$\mathbf{A}$的最小多项式, 如例17中的$\mathbf{A}$的特征多项式为 $f(\lambda)=(\lambda-2)^3$. 最小多项式 $m(\lambda)$一定是$f(\lambda)$的因式, 因此它的次数次数不会超过n, 即

推论 4. *设矩阵 $\mathbf{A}$的特征多项式为$f(\lambda)$, 而$m(\lambda)$是$\mathbf{A}$的最小多项式, 则*

$$m(\lambda)\,\big|\,f(\lambda)$$

5.3.4 矩阵对角化

同一个线性变换在不同的基下对应的矩阵不同, 这些矩阵之间有相似关系. 那么, 是否存在一个基, 使得线性变换在这个基下对应的矩阵具有简单的形式, 如对角阵或上三角阵等. 本节将研究矩阵的对角化问题.

设σ是n维线性空间V上的线性变换, 当选定基以后, 线性变换的特征值问题就转化为变换矩阵的特征值问题: $\mathbf{AX}=\lambda\mathbf{X}$, 若矩阵 $\mathbf{A}$ 所有特征值的几何重数与代数重数相等, 则$\mathbf{A}$有n个特征向量, 且根据定理5.3.5的推论3可知这些特征向量线性无关, 设$\mathbf{A}$的n个特征值为$\lambda_1,\lambda_2,\ldots,\lambda_n$, 将它们写成对角矩阵 $\Lambda=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$, 同理, 将特征向量按特征值的对应顺序排列成矩阵 $\mathbf{X}=\begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n\end{bmatrix}$, 则特征方程可写成

$$\mathbf{AX}=\mathbf{X}\Lambda$$

因n个特征向量线性无关, 所以$\mathbf{X}$可逆, 因此有

$$\mathbf{X}^{-1}\mathbf{AX}=\Lambda \tag{5.21}$$

式(5.21)说明矩阵$\mathbf{A}$与对角阵阵相似, 若矩阵$\mathbf{A}$是线性变换σ在基$\mathfrak{B}$下的对应矩阵, 选择基$\mathfrak{B}'=\mathfrak{B}\mathbf{X}$, 在基$\mathfrak{B}'$下与$\sigma$对应的矩阵为对角阵. 因而有下述定理

定理 5.3.10. *n阶方阵$\mathbf{A}$与对角阵相似的充要条件是$\mathbf{A}$有n个不同的特征向量.*

证: (必要性) 设$\mathbf{A}$相似于对角阵$\mathbf{D}=\operatorname{diag}(d_1,d_2,\ldots,d_n)$, 则存在可逆矩阵$\mathbf{P}$ 使得

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{D}\Rightarrow\mathbf{A}\mathbf{P}=\mathbf{P}\mathbf{D}$$

将矩阵$\mathbf{P}$的第j列向量记为$\mathbf{p}_j(j=1,2,\ldots,n)$. 则有

$$\mathbf{A}\mathbf{p}_j=d_j\mathbf{p}_j,\quad j=1,2,\ldots,n$$

显然, d_j和p_j是矩阵$\mathbf{A}$的特征值和相应的特征向量.

(充分性) 参见上述讨论.

例 18. *设矩阵*$\mathbf{A}=\begin{bmatrix}1&3&3\\-3&-5&-3\\3&3&1\end{bmatrix}$, *可能的话求可逆阵*$\mathbf{P}$ *使得*$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{D}$, *其中*$\mathbf{D}$*为对角阵.*

解: 通过求$\mathbf{A}$的特征值和特征向量判断$\mathbf{A}$是否可对角化.

$$f_{\mathbf{A}}(\lambda)=\det(\lambda\mathbf{E}-\mathbf{A})=\lambda^3+3\lambda^2-4=0$$

解得$\lambda_1=\lambda_2=-2,\lambda_3=1$.

当$\lambda_1=\lambda_2=-2$时,

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

求得基础解系(线性无关的特征向量)为: $\begin{bmatrix}1\\-1\\0\end{bmatrix},\begin{bmatrix}1\\0\\-1\end{bmatrix}$.

当$\lambda_3=1$时,

$$\begin{bmatrix}0&-3&-3\\3&6&3\\-3&-3&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\mathbf{0}$$

求得基础解系为: $\begin{bmatrix}1\\-1\\1\end{bmatrix}$ 因此

$$\Lambda=\begin{bmatrix}-2&&\\&-2&\\&&1\end{bmatrix},\quad 令\mathbf{P}=\mathbf{X}=\begin{bmatrix}1&1&1\\-1&0&-1\\0&-1&1\end{bmatrix}$$

$$则\quad\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\begin{bmatrix}-1&-2&-1\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}1&3&3\\-3&-5&-3\\3&3&1\end{bmatrix}\begin{bmatrix}1&1&1\\-1&0&-1\\0&-1&1\end{bmatrix}=\begin{bmatrix}-2&&\\&-2&\\&&1\end{bmatrix}$$

例 19. *例16中的矩阵*$\mathbf{A}$, *特征值2的几何重数小于代数重数, 最终只求得*$\mathbf{A}$*两个特征向量, 因此* $\mathbf{A}$ *不能对角化.*

若将(5.21)写成

$$\mathbf{A}=\mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-1}\tag{5.22}$$

则称式(5.22)为矩阵$\mathbf{A}$的**特征分解**.

判定矩阵 $\mathbf{A}$ 是否可对角化, 用求矩阵的特征问题来判断比较繁琐, 下面讨论一类一定能对角化的矩阵, 而且这类矩阵的特征向量之间 还具有正交性.

定理 5.3.11. *实对称矩阵的所有特征值都是实数.*

证: $\mathbf{A}$是n阶实对称矩阵, 设λ是任一个特征值, $\mathbf{X}=\left[\begin{array}{llll} x_1 & x_2 & \cdots & x_n \end{array}\right]^T$是对应$\lambda$的特征向量, 于是有

$$\mathbf{AX}=\lambda\mathbf{X} \tag{5.23}$$

对上式两边取共轭, 有

$$\overline{\mathbf{AX}}=\mathbf{A}\overline{\mathbf{X}}=\overline{\lambda}\,\overline{\mathbf{X}}$$

在对上式两边取转置,得

$$\overline{\mathbf{X}}^T\mathbf{A}=\overline{\lambda}\,\overline{\mathbf{X}}^T \quad (\mathbf{A}\text{是实对称矩阵}) \tag{5.24}$$

对式(5.23)两端左乘$\overline{\mathbf{X}}^T$, 式(5.24)两端右乘$\mathbf{X}$, 再相减得

$$0=\left(\lambda-\overline{\lambda}\right)\overline{\mathbf{X}}^T\mathbf{X}$$

因$\mathbf{X}\neq\mathbf{0}$, 所以$\overline{\mathbf{X}}^T\mathbf{X}\neq 0$, 上式又成立, 只有$\overline{\lambda}=\lambda$, 因此$\lambda$是实数.

通常将向量或矩阵取共轭后再取转置, 称为共轭转置, 用符号"$(\bullet)^H$"表示. 因实对称矩阵的所有特征值都是实数, 它对应的特征方程的系数矩阵全部是实数, 因此它的解也全部是实向量, 即特征向量也是实向量.

上述定理的结论可推广到复对称矩阵, 即

定理 5.3.12. *复对称矩阵的所有特征值都是实数.*

证明类似定理5.3.11.

定理 5.3.13. *设$\mathbf{A}$是n阶实对称矩阵, 则对应$\mathbf{A}$的不同特征值的特征向量正交.*

证: 设λ_1,λ_2是$\mathbf{A}$的任意两个不同特征值, $\mathbf{X}_1,\mathbf{X}_2$ 是 λ_1,λ_2 对应得特征向量, 对式 $\mathbf{AX}_1=\lambda_1\mathbf{X}_1$, $\mathbf{AX}_2=\lambda_2\mathbf{X}_2$ 两端分别左乘$\mathbf{X}_2^T$和$\mathbf{X}_1^T$有

$$\mathbf{X}_2^T\mathbf{AX}_1=\lambda_1\mathbf{X}_2^T\mathbf{X}_1,\quad \mathbf{X}_1^T\mathbf{AX}_2=\lambda_2\mathbf{X}_1^T\mathbf{X}_2$$

两式相减后得

$$0=(\lambda_1-\lambda_2)\,\mathbf{X}_1^T\mathbf{X}_2$$

因 $\lambda_1\neq\lambda_2$, 所以

$$\mathbf{X}_1^T\mathbf{X}_2=(\mathbf{X}_1,\mathbf{X}_2)=0$$

定理 5.3.14. *n阶实对称矩阵的任一特征值的代数重数等于它的几何重数.*

证: 设$\mathbf{A}$为n阶实对称矩阵, λ_0是它的任一特征值, 其对应的特征方程:

$$(\lambda_0\mathbf{E}-\mathbf{A})\,\mathbf{X}=\mathbf{0} \tag{5.25}$$

的基础解系为 $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_k$ (即 λ_0 的几何重数为 k), 不妨设它们正交标准化(否则通过Gram-Schmidt方法实现标准正交化), 即

$$(\mathbf{X}_i,\mathbf{X}_j)=\delta_{ij}=\left\{\begin{array}{ll} 1, & i=j \\ 0, & i\neq j \end{array}\right., \quad 1\leq i,j\leq k$$

将它们扩充成n个标准正交的向量组, 并将它们排列成矩阵Ψ, 即

$$\mathbf{\Psi}=[\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_k,\mathbf{X}_{k+1},\ldots,\mathbf{X}_n]$$

则$\mathbf{\Psi}$是正交矩阵, 即$\Psi^{-1}=\Psi^T$, 因$\mathbf{A}$是实对称矩阵, 所以$\Psi^T\mathbf{A\Psi}$还是实对称矩阵, 令

$$\mathbf{B}=\mathbf{\Psi}^T\mathbf{A\Psi}=\left[\begin{array}{cc} \lambda_0\mathbf{E}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_1 \end{array}\right]$$

且 $\mathbf{B}_1$ 是对称矩阵. 显然, 矩阵 $\mathbf{A}$ 与 $\mathbf{B}$ 相似, 它们的特征多项式相同,

$$\det(\lambda\mathbf{E}-\mathbf{A})=\det(\lambda\mathbf{E}-\mathbf{B})=(\lambda-\lambda_0)^k\det(\lambda\mathbf{E}-\mathbf{B}_1)$$

假设 λ_0 的代数重数大于 k, 即 $\det(\lambda\mathbf{E}-\mathbf{B}_1)$ 中还包含 $\lambda-\lambda_0$因子, 则对特征方程(5.25)中的 $\mathbf{X}$ 进行变量替换 $\mathbf{X}=\Psi\mathbf{Y}$, 则

$$(\lambda_0\mathbf{E}-\mathbf{A})\,\Psi\mathbf{Y}=\mathbf{0}$$

方程两端在左乘Ψ^T 不会改变方程的解, 即

$$\Psi^T(\lambda_0\mathbf{E}-\mathbf{A})\,\Psi\mathbf{Y}=\mathbf{0}$$

得

$$\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \lambda_0\mathbf{E}-\mathbf{B}_1\end{bmatrix}\mathbf{Y}=\mathbf{0}$$

其中左上角的分块为$k\times k$, 根据假设 $\lambda_0\mathbf{E}-\mathbf{B}_1$ 还是奇异阵, 系数矩阵的值小于 $n-k$, 基础解系的维数大于 k, 这与方程(5.25)的基础解系的秩为 k 矛盾. 因此λ_0 的代数重数等于几何重数 k.

定理5.3.14说明n阶任何实对称矩阵包含n个特征向量, 因此它们必可对角化.

对任意n阶实对称阵$\mathbf{A}$的对角化过程:

(1) 求出 $\mathbf{A}$ 的特征多项式的所有不同根, 设它们为: $\lambda_1,\lambda_2,\ldots,\lambda_m$.

(2) 将这些特征值依次代入特征方程, 求它的基础解系. 设将 λ_i 代入特征方程后为 $(\lambda_i\mathbf{E}-\mathbf{A})\,\mathbf{X}=\mathbf{0}$, 它的基础解系为 $\mathbf{X}_{i1},\mathbf{X}_{i2},\ldots,\mathbf{X}_{ir_i}$, 它们是对应特征值 λ_i 的所有特征向量.

(3) 用Gram-Schmidt方法对基础解系 $\mathbf{X}_{i1},\mathbf{X}_{i2},\ldots,\mathbf{X}_{ir_i}$ 进行正交标准化, 得到$\eta_{i1},\eta_{i2},\ldots,\eta_{ir_i}$, 它们为对应$\lambda_i$的标准正交化特征向量.

(4) 因特征值 $\lambda_1,\lambda_2,\ldots,\lambda_m$ 各不相同, 不同特征值对应的特征向量之间正交, 即特征向量 $\eta_{11},\eta_{12},\ldots,\eta_{1r_1},\eta_{21},\ldots,\eta_{2r_2},\eta_{31},\ldots,\eta_{m1},\ldots,\eta_{mr_m}$ 相互正交且标准化, 且$\sum\limits_{i=1}^{m}r_i=n$, 将这些特征向量向量排列成矩阵 $\mathbf{P}$, 则 $\mathbf{P}$是正交矩阵, 并将矩阵对角化, 即 $\mathbf{P}^T\mathbf{A}\mathbf{P}=\Lambda$.

例 20. *求正交矩阵$\mathbf{P}$, 将矩阵*

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

对角化.

解: 求矩阵$\mathbf{A}$的特征值,

$$f(\lambda)=\det(\lambda\mathbf{E}-\mathbf{A})=\det\begin{bmatrix}\lambda-1 & 2 & 2\\ 2 & \lambda-4 & 1\\ 2 & 1 & \lambda-4\end{bmatrix}=0$$

即: $(\lambda+1)(\lambda-5)^2=0$, 求得特征值为$\lambda_1=\lambda_2=5,\lambda_3=-1$.

求特征向量: 将$\lambda_1=\lambda_2=5$代入特征方程得

$$\begin{bmatrix}4 & 2 & 2\\ 2 & 1 & 1\\ 2 & 1 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\mathbf{0}$$

基础解系: $\begin{bmatrix}1\\ -2\\ 0\end{bmatrix},\begin{bmatrix}1\\ 0\\ -2\end{bmatrix}$

标准正交化:$\varepsilon_1 = \frac{\sqrt{5}}{5}\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \frac{\sqrt{30}}{30}\begin{bmatrix} 2 \\ 1 \\ -5 \end{bmatrix}$.

将$\lambda_3 = -1$代入特征方程,

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

基础解系:$\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, 标准正交化: $\varepsilon_3 = \frac{\sqrt{6}}{6}\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$.

因此,$\mathbf{P} = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{30} & 2/\sqrt{6} \\ -2/\sqrt{5} & 1/\sqrt{30} & 1/\sqrt{6} \\ 0 & -5/\sqrt{30} & 1/\sqrt{6} \end{bmatrix}$, 使得 $\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{bmatrix} 5 & & \\ & 5 & \\ & & -1 \end{bmatrix}$

5.4 常用的线性变换

5.4.1 正交变换

正交变换是欧式空间中一类重要的变换.

定义 5.10. *设 σ 是 n 维欧式空间 V 上的线性变换, 若对 V 中的任意向量 $\mathbf{x}, \mathbf{y}$ 满足*

$$(\sigma(\mathbf{x}), \sigma(\mathbf{y})) = (\mathbf{x}, \mathbf{y}) \tag{5.26}$$

*则称 σ 为**正交变换**.*

正交变换还具有下列性质:

(1) 对任意向量 $\mathbf{x}$, 因 $\|\sigma(\mathbf{x})\|^2 = (\sigma(\mathbf{x}), \sigma(\mathbf{x})) = (\mathbf{x}, \mathbf{x}) = \|\mathbf{x}\|^2$, 说明正交变换不改变向量的长度.

(2) 同样, 若 $\mathbf{x} \perp \mathbf{y}$, 则 $\sigma(\mathbf{x}) \perp \sigma(\mathbf{y})$, 即正交变换不会改变向量之间的正交性.

(3) 若 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是欧式空间 V 的一个标准正交基, 则由(1)和(2)可知 $\sigma(\varepsilon_1), \sigma(\varepsilon_2), \ldots, \sigma(\varepsilon_n)$ 每个向量长度不变还是 1, 而且相互正交性不变, 再由相互正交的一组非零向量必线性无关, 由此可知它们构成 V 的标准正交基.

(4) σ 在任意一组标准正交基下对应正交矩阵. 取 $\mathfrak{B} = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ 为 V的任意一组标准正交基, 设

$$\begin{gathered} \sigma(\varepsilon_1) = \mathfrak{B}\mathbf{a}_1 \\ \sigma(\varepsilon_2) = \mathfrak{B}\mathbf{a}_2 \\ \vdots \\ \sigma(\varepsilon_n) = \mathfrak{B}\mathbf{a}_n \end{gathered}$$

令矩阵 $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ 即为线性变换在此基下的矩阵$[\sigma(\mathfrak{B})]_{\mathfrak{B}}$. 由性质(1),(2)可知

$$\begin{aligned} & (\varepsilon_i, \varepsilon_j) = (\sigma(\varepsilon_i), \sigma(\varepsilon_j)) = (\mathfrak{B}\mathbf{a}_i, \mathfrak{B}\mathbf{a}_j) \\ = \quad & (\mathbf{a}_i, \mathbf{a}_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \end{aligned}$$

因此$[\sigma(\mathfrak{B})]_{\mathfrak{B}}(\mathbf{A})$是正交矩阵.

(5) 由(4)可知正交变换对应的矩阵$\mathbf{A}$是正交矩阵, 所以

$$\begin{aligned} & \mathbf{A}^T\mathbf{A} = \mathbf{E} \Rightarrow \det\left(\mathbf{A}^T\mathbf{A}\right) = 1 \\ \Rightarrow \quad & (\det \mathbf{A})^2 = 1 \Rightarrow \det \mathbf{A} = \pm 1 \end{aligned}$$

(6) 设τ是欧式空间V上的另一正交变换, 则$\sigma \cdot \tau$还是正交变换.
因为$\forall \mathbf{x}, \mathbf{y} \in V$, 有

$$\begin{aligned} & ((\sigma \cdot \tau)(\mathbf{x}), (\sigma \cdot \tau)(\mathbf{y})) = (\sigma(\tau(\mathbf{x})), \sigma(\tau(\mathbf{y}))) \\ = \quad & (\tau(\mathbf{x}), \tau(\mathbf{y})) = (\mathbf{x}, \mathbf{y}) \end{aligned}$$

由性质(5)知正交变换对应矩阵的行列式为正负1, 可将所有正交变换按齐对应矩阵的行列式等于1或-1进行分类, 将对应矩阵行列式为1的正交变换称为**第一类正交变换**, 而将对应矩阵的行列式等于-1的正交变换称为**第二类正交变换**.

例 21. *平面上的旋转变换σ_θ 在标准正交基下的矩阵为*

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

其行列式为1, 因此,它是第一类正交变换.

例 22. *例5中的镜像变换, 变换前后向量的内积不变, 因此它是正交变换. 推广到n维实欧式空间$\mathbb{R}^n$, 任选一个标准正交基, 在这组基下, 镜面的法方向表示为n元单位向量$\mathbf{n}$, 此时该变换对应的矩阵为*

$$\mathbf{M} = \mathbf{E} - 2\mathbf{n}\mathbf{n}^T$$

显然$\mathbf{M}$是对称矩阵, 且$\mathbf{M}^T\mathbf{M} = \mathbf{M}\mathbf{M} = \mathbf{E}$, 它有两个特征值1和$-1$, 其中1是$n-1$重特征值, 向量$n$是对应特征值$-1$的特征向量, 因为$\mathbf{M}\mathbf{n} = \mathbf{n} - 2\mathbf{n} = -\mathbf{n}$, 它构成特征子空间$V_{-1} = \{c\mathbf{n} | c \in \mathbb{R}\}$, 它的正交补空间为$V_1$. 所以$\det \mathbf{M} = -1$, 它是第二类正交变换.

5.4.2 投影变换

定义 5.11. *若n维欧式空间V表示成两个子空间的直和, 即$V = U \oplus W$, 则对$\forall \mathbf{x} \in V$, $\mathbf{x}$ 可唯一表示成: $\mathbf{x} = \mathbf{y} + \mathbf{z}$, 其中 $\mathbf{y} \in U$, $\mathbf{z} \in W$, 称 $\mathbf{y}$ 为 $\mathbf{x}$ 沿着W到U的**投影**. 若V上的变换能将任意 $\mathbf{x} \in V$ 变换成沿着W到U的投影, 则称该变换为沿着W到U的**投影变换**, 记为$P_{U,W}$. 即 $P_{U,W}(\mathbf{x}) = \mathbf{y}$.*

容易验证投影变换是线性变换, 即

(1) $\forall \mathbf{x}_1, \mathbf{x}_2 \in V$, $\mathbf{x}_i = \mathbf{y}_i + \mathbf{z}_i$ 其中 $\mathbf{y}_i \in U$,$\mathbf{z}_i \in W (i = 1, 2)$, 则 $P_{U,W}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{y}_1 + \mathbf{y}_2 = P_{U,W}(\mathbf{x}_1) + P_{U,W}(\mathbf{x}_2)$.

(2) c 是任意常数, $\forall \mathbf{x} \in V$, $\mathbf{x} = \mathbf{y} + \mathbf{z} (\mathbf{y} \in U, \mathbf{z} \in W)$, 则$P_{U,W}(c\mathbf{x}) = c\mathbf{y} = cP_{U,W}(\mathbf{x})$.

定义 5.12. *投影变换在n维欧式空间的基$\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$下对应的矩阵称为**投影矩阵**, 用符号$\mathbf{P}_{U,W}$表示.*

为了清楚投影变换, 先给出幂等阵的定义、性质及其与投影变换之间的关系.

定义 5.13. *n阶方阵$\mathbf{A}$满足$\mathbf{A}^2 = \mathbf{A}$, 则称$\mathbf{A}$为**幂等阵**.*

幂等阵的一个重要性质:

定理 5.4.1. *幂等矩阵 $\mathbf{A}$的核(零)空间等于矩阵 $\mathbf{E} - \mathbf{A}$的像空间.*

证: 因$\mathbf{A}$是幂等阵, 下列关系成立:

$$\mathbf{A}(\mathbf{E} - \mathbf{A}) = \mathbf{0}$$

对 $\forall \mathbf{x} \in Im(\mathbf{E} - \mathbf{A})$ 存在向量 $\mathbf{y}$, 使得$\mathbf{x} = (\mathbf{E} - \mathbf{A})\mathbf{y}$, 而 $\mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{E} - \mathbf{A})\mathbf{y} = \mathbf{0}$. 即

$$Im(\mathbf{E} - \mathbf{A}) \subseteq Ker(\mathbf{A}) \tag{5.27}$$

由上式可知

$$\dim Im(\mathbf{E} - \mathbf{A}) = \operatorname{rank}(\mathbf{E} - \mathbf{A}) \leq \dim Ker(\mathbf{A}) = n - \operatorname{rank}(\mathbf{A}) \tag{5.28}$$

另由 $\mathbf{E} = \mathbf{A} + (\mathbf{E} - \mathbf{A})$, 得

$$n = \operatorname{rank}(\mathbf{E}) \leq \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{E} - \mathbf{A})$$

即 $n - \operatorname{rank}(\mathbf{A}) \leq \operatorname{rank}(\mathbf{E} - \mathbf{A})$ 结合式(5.28), 有

$$n - \operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{E} - \mathbf{A})$$

再结合式(5.27)有 $Ker(\mathbf{A}) = Im(\mathbf{E} - \mathbf{A})$.

定理 5.4.2. *矩阵$\mathbf{P}$是投影矩阵的充要条件是$\mathbf{P}$为幂等阵.*

证: 设欧式空间$V=U\oplus W$, 且$\mathbf{P}=\mathbf{P}_{U,W}$ 是投影矩阵, 则$\forall\mathbf{x}\in V$,$\mathbf{x}=\mathbf{y}+\mathbf{z}$,$\mathbf{y}\in U$,$\mathbf{z}\in W$, 则有

$$\mathbf{P}_{U,W}\mathbf{x}=\mathbf{y}\Rightarrow\mathbf{P}_{U,W}\mathbf{P}_{U,W}\mathbf{x}=\mathbf{P}_{U,W}\mathbf{y}=\mathbf{y}$$

由$\mathbf{x}$的任意性得:$\mathbf{P}_{U,W}\mathbf{P}_{U,W}=\mathbf{P}_{U,W}$.

反之, 设$\mathbf{P}$是幂等阵, 则 $Im\left(\mathbf{P}\right)$和$Ker\left(\mathbf{P}\right)$是$V$的子空间, 则对 $\forall\mathbf{x}$, 有

$$\mathbf{x}=\mathbf{x}-\mathbf{P}\mathbf{x}+\mathbf{P}\mathbf{x}=\left(\mathbf{E}-\mathbf{P}\right)\mathbf{x}+\mathbf{P}\mathbf{x}$$

其中 $\mathbf{P}\mathbf{x}\in Im\left(\mathbf{P}\right)$, $\left(\mathbf{E}-\mathbf{P}\right)\mathbf{x}\in Ker\left(\mathbf{P}\right)$, 因此

$$Im\left(\mathbf{P}\right)+Ker\left(\mathbf{P}\right)=V \tag{5.29}$$

设$\forall\alpha\in Im\left(\mathbf{P}\right)\cap Ker\left(\mathbf{P}\right)$, 则存在$\beta,\gamma\in V$, 使得

$$\alpha=\mathbf{P}\beta=\left(\mathbf{E}-\mathbf{P}\right)\gamma$$

又因$\mathbf{P}$幂等阵, 则$\alpha=\mathbf{P}\beta=\mathbf{P}\left(\mathbf{P}\beta\right)=\mathbf{P}\left(\mathbf{E}-\mathbf{P}\right)\gamma=\mathbf{0}$ 即

$$Im\left(\mathbf{P}\right)\cap Ker\left(\mathbf{P}\right)=\mathbf{0}$$

所以 $Im\left(\mathbf{P}\right)\oplus Ker\left(\mathbf{P}\right)=V$. 因此 $\mathbf{P}$ 是沿着 $Ker\left(\mathbf{P}\right)$ 到 $Im\left(\mathbf{P}\right)$的投影.

如何求投影矩阵$\mathbf{P}_{U,W}$? 设n维欧式空间 $V=U\oplus W$, 且$\dim U=p$,$\dim W=n-p$. U的一个基为: $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_p$, W的基为$\eta_1,\eta_2,\ldots,\eta_{n-p}$. 则

$$\begin{aligned}\mathbf{P}_{U,W}\left(\varepsilon_i\right)=\varepsilon_i,\quad i=1,2,\ldots,p\\ \mathbf{P}_{U,W}\left(\eta_j\right)=\mathbf{0},\quad j=1,2,\ldots,n-p\end{aligned}$$

令 $\Upsilon=\begin{bmatrix}\varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_p\end{bmatrix}$,$\Xi=\begin{bmatrix}\eta_1 & \eta_2 & \cdots & \eta_{n-p}\end{bmatrix}$ 则有 $\mathbf{P}_{U,W}\begin{bmatrix}\Upsilon & \Xi\end{bmatrix}=\begin{bmatrix}\Upsilon & \mathbf{0}\end{bmatrix}$, 因此

$$\mathbf{P}_{U,W}=\begin{bmatrix}\Upsilon & \mathbf{0}\end{bmatrix}\begin{bmatrix}\Upsilon & \Xi\end{bmatrix}^{-1} \tag{5.30}$$

例 23. *设欧式空间$\mathbb{R}^3$ 的子空间$U=\operatorname{span}\left(\varepsilon_1,\varepsilon_2\right)$, $W=\operatorname{span}\left(\eta\right)$, 其中*

$$\varepsilon_1=\begin{bmatrix}1\\2\\-1\end{bmatrix},\quad\varepsilon_2=\begin{bmatrix}0\\-1\\1\end{bmatrix},\quad\eta=\begin{bmatrix}-2\\1\\1\end{bmatrix}$$

求投影矩阵$\mathbf{P}_{W,U}$ 及向量 $\mathbf{x}=\begin{bmatrix}1\\1\\1\end{bmatrix}$ 在子空间 W 的投影.

解: 投影矩阵$\mathbf{P}_{W,U}=\begin{bmatrix}-2&0&0\\1&0&0\\1&0&0\end{bmatrix}\begin{bmatrix}-2&1&0\\1&2&-1\\1&-1&1\end{bmatrix}^{-1}=\dfrac{1}{4}\begin{bmatrix}2&-2&-2\\-1&1&1\\-1&1&1\end{bmatrix}$

$$\mathbf{P}_{W,U}\mathbf{x}=\frac{1}{4}\begin{bmatrix}-2\\1\\1\end{bmatrix}$$

定义 5.14. *设n维欧式空间$V=S\oplus S^{\perp}$($S^{\perp}$为S的正交补空间), 则称沿着$S^{\perp}$到S的投影矩阵 $P_{S,S^{\perp}}$称为正交投影变换, 简记为P_S,它在基$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$ 下的矩阵$\mathbf{P}_S$ 称为**正交投影矩阵**.*

正交投影矩阵不仅是幂等阵, 而且还是Hermite阵(复对称阵).

定理 5.4.3. *矩阵 $\mathbf{P}$为n维欧式空间V的正交投影阵的充要条件是$\mathbf{P}$是幂等Hermite阵.*

证: 设$\mathbf{P}$为正交投影矩阵$\mathbf{P}_S$, 由定理5.4.2可知它是幂等阵, 对$\forall \mathbf{x} \in V$, 则 $\mathbf{P}_S\mathbf{x} = \mathbf{y} \in S$, $(\mathbf{E}-\mathbf{P}_S)\mathbf{x} = \mathbf{x}-\mathbf{y} = \mathbf{z} \in S^{\perp}$, 且$(\mathbf{y},\mathbf{z}) = 0$, 即

$$\begin{array}{l}(\mathbf{P}_S\mathbf{x})^H(\mathbf{E}-\mathbf{P}_S)\mathbf{x} = \mathbf{x}^H\left(\mathbf{P}_S^H - \mathbf{P}_S^H\mathbf{P}_S\right)\mathbf{x} = 0 \\ ((\mathbf{E}-\mathbf{P}_S)\mathbf{x})^H(\mathbf{P}_S\mathbf{x}) = \mathbf{x}^H\left(\mathbf{P}_S - \mathbf{P}_S^H\mathbf{P}_S\right)\mathbf{x} = 0\end{array}$$

上式对任意$\mathbf{x}$成立, 因此

$$\mathbf{P}_S^H = \mathbf{P}_S^H\mathbf{P}_S = \mathbf{P}_S$$

因此$\mathbf{P}_S$为幂等Hermite阵.

反之, 设$\mathbf{P}$是幂等Hermite阵, 结合定理5.4.2有

$$\mathbf{P} = \mathbf{P}_{Im(\mathbf{P}),Ker(\mathbf{P})} = \mathbf{P}_{Im(\mathbf{P}),Ker(\mathbf{P}^H)} = \mathbf{P}_{Im(\mathbf{P}),Ker(\mathbf{P})^{\perp}} = \mathbf{P}_{Im(\mathbf{P})}$$

说明$\mathbf{P}$是正交投影矩阵.

设$\dim S = p$, 且$\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_p$为$S$的一个基, $\eta_1,\eta_2,\ldots,\eta_{n-p}$是$S^{\perp}$的一个基, 令$\Upsilon = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_p \end{bmatrix}$, $\Xi = \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_{n-p} \end{bmatrix}$, 则

$$\mathbf{P}_S \begin{bmatrix} \Upsilon & \Xi \end{bmatrix} = \begin{bmatrix} \Upsilon & \mathbf{0} \end{bmatrix}$$

因此

$$\begin{aligned}\mathbf{P}_S &= \begin{bmatrix} \Upsilon & \mathbf{0} \end{bmatrix}\begin{bmatrix} \Upsilon & \Xi \end{bmatrix}^{-1} = \begin{bmatrix} \Upsilon & \mathbf{0} \end{bmatrix}\left(\begin{bmatrix} \Upsilon & \Xi \end{bmatrix}^H\begin{bmatrix} \Upsilon & \Xi \end{bmatrix}\right)^{-1}\begin{bmatrix} \Upsilon & \Xi \end{bmatrix}^H \\ &= \begin{bmatrix} \Upsilon & \mathbf{0} \end{bmatrix}\begin{bmatrix} (\Upsilon^H\Upsilon)^{-1} & \mathbf{0} \\ \mathbf{0} & (\Xi^H\Xi)^{-1} \end{bmatrix}\begin{bmatrix} \Upsilon^H \\ \Xi^H \end{bmatrix} = \Upsilon\left(\Upsilon^H\Upsilon\right)^{-1}\Upsilon^H\end{aligned} \tag{5.31}$$

例 24. *设S是欧式空间$\mathbf{R}^3$的子空间, $S = \mathrm{span}(\varepsilon_1,\varepsilon_2)$, 其中*

$$\varepsilon_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \varepsilon_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

求正交投影矩阵$\mathbf{P}_S$和向量$\mathbf{x} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T$在$S$的投影.

解: 因

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

因此

$$\mathbf{P}_S = \Upsilon\left(\Upsilon^T\Upsilon\right)^{-1}\Upsilon^T = \frac{1}{6}\begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & 1 \\ -2 & 1 & 5 \end{bmatrix}$$

$$\mathbf{P}_S\mathbf{x} = \frac{1}{6}\begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & 1 \\ -2 & 1 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5/2 \\ 5/2 \end{bmatrix}$$

5.5 线性映射

5.5.1 线性映射定义及矩阵表示

上述讨论的是变换是线性空间V到自身的映射, 下面简单地讨论一下不同线性空间之间的映射.

定义 5.15. *设V,W是数域F上的线性空间, 若映射$\sigma : V \mapsto W$, 满足线性条件:*

(1) $\forall \mathbf{x},\mathbf{y} \in V, \sigma(\mathbf{x}+\mathbf{y}) = \sigma(\mathbf{x}) + \sigma(\mathbf{y})$.

(2) $\forall c \in F, \mathbf{x} \in V$, $\sigma(c\mathbf{x}) = c\sigma(\mathbf{x})$.

*则称σ为V到W的**线性映射**.*

例 25. *σ是$\mathbb{R}^2$到$\mathbb{R}^3$的映射:*

$$\sigma\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} ax \\ y \\ x+ay \end{bmatrix}$$

其中 a为给定常数, 很容易验证σ是线性映射.

数域F上线性空间V到W的线性映射σ具有下列性质:

(1) 对$\mathbf{0} \in V$, 则 $\sigma(\mathbf{0}) = \mathbf{0} \in W$.

(2) 对$\forall \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in V$ 和 $\forall c_1, c_2, \ldots, c_k \in F$, 线性映射保持线性组合形式不变, 即

$$\sigma\left(\sum_{i=1}^{k} c_i \mathbf{x}_i\right) = \sum_{i=1}^{k} c_i \sigma(\mathbf{x}_i)$$

(3) 若σ是V到W的双射(同构映射), 则σ^{-1}是W到V的同构映射.

请参照线性变换相应性质自行证明, 包括下列结论.

定义$V \mapsto W$ 的映射之间的加法和数乘运算, 当运算满足线性条件时, 它们是线性映射. 用$L(V, W)$表示$V \mapsto W$的线性映射全体, 可以证明$L(V, W)$是线性空间.

定义 5.16. *设σ是数域F上的线性空间V到W的映射, 令 $Im(\sigma) = \left\{\mathbf{y} \middle| \mathbf{y} = \sigma(\mathbf{x}), \forall \mathbf{x} \in V\right\}$, 则 $Im(\sigma)$ 称为σ的像, 显然$Im(\sigma) \subseteq W$. 令$Ker(\sigma) = \left\{\mathbf{x} \middle| \mathbf{x} \in V \wedge \sigma(\mathbf{x}) = \mathbf{0}\right\}$, 则称$Ker(\sigma)$为$\sigma$的**核**.*

定理 5.5.1. *$Im(\sigma)$是W的线性子空间, $Ker(\sigma)$是V的线性子空间.*

下面讨论线性映射与矩阵之间的关系. 设V和W是数域F上的线性空间, 若取定V的一个基, 则线性映射σ完全由对这个基的作用结果决定, 即

定理 5.5.2. *设$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是V的一个基, 则*

(1) 若 σ, τ都是V到W的线性映射, 且 $\sigma(\varepsilon_i) = \tau(\varepsilon_i)(i = 1, 2, \ldots, n)$, 则 $\sigma = \tau$.

(2) 给定W中的n个元素$\beta_1, \beta_2, \ldots, \beta_n$, 有且仅有一个线性映射$\sigma$满足$\sigma(\varepsilon_i) = \beta_i(i = 1, 2, \ldots, n)$.

证: (1). $\forall \mathbf{x} \in V$ 可唯一表示成

$$\mathbf{x} = \sum_{i=1}^{n} c_i \varepsilon_i$$

则

$$\sigma(\mathbf{x}) = \sum_{i=1}^{n} c_i \sigma(\varepsilon_i) = \sum_{i=1}^{n} c_i \tau(\varepsilon_i) = \sum_{i=1}^{n} \tau(c_i \varepsilon_i) = \tau(\mathbf{x})$$

因此$\sigma = \tau$.

(2). 定义映射$\sigma: V \mapsto W$, 它满足对$\forall \mathbf{x} = \sum_{i=1}^{n} c_i \varepsilon_i$ 有

$$\sigma(\mathbf{x}) = \sum_{i=1}^{n} c_i \beta_i$$

可以证明映射σ是V到W的线性映射, 且$\sigma(\varepsilon_i) = \beta_i$. 唯一性由(1)可得.

设数域F上的线性空间V和W的维数分别为:$\dim V = n$,$\dim W = m$. 并且$\mathfrak{B}_1 = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ 是V的一个基, $\mathfrak{B}_2 = \{\eta_1, \eta_2, \ldots, \eta_m\}$ 是W的一个基, σ是V到W的线性映射, $\sigma(\varepsilon_i)$ 可展开成 W的基的线性组合:

$$\sigma(\varepsilon_i) = \sum_{j=1}^{m} a_{ji} \eta_j, \quad i = 1, 2, \ldots, n$$

令$\mathbf{A}=[a_{ji}]_{m\times n}$, 若采用上面章节中的坐标表示形式$\mathbf{A}$等于 $[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_2}$. 对$\forall \mathbf{x}\in V$ 设$\mathbf{y}=\sigma(\mathbf{x})$, 将$\mathbf{x},\mathbf{y}$表示成基线性组合, 有 $\mathbf{x}=\sum_{i=1}^{n}c_i\varepsilon_i=\mathfrak{B}_1[\mathbf{x}]_{\mathfrak{B}_1}$, 和 $\mathbf{y}=\sum_{i=1}^{n}d_i\eta_i=\mathfrak{B}_2[\mathbf{y}]_{\mathfrak{B}_2}$, 且成立关系式:

$$\mathfrak{B}_2[\mathbf{y}]_{\mathfrak{B}_2}=\mathbf{y}=\sigma(\mathbf{x})=\sigma\left(\mathfrak{B}_1[\mathbf{x}]_{\mathfrak{B}_1}\right)=\mathfrak{B}_2[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_2}[\mathbf{x}]_{\mathfrak{B}_1}$$

因$\mathfrak{B}_2$是基, 上式亦可写成坐标之间的关系式:

$$[\mathbf{y}]_{\mathfrak{B}_2}=[\sigma(\mathfrak{B}_1)]_{\mathfrak{B}_2}[\mathbf{x}]_{\mathfrak{B}_1}=\mathbf{A}[\mathbf{x}]_{\mathfrak{B}_1} \tag{5.32}$$

则矩阵$\mathbf{A}$就是线性映射σ对应的矩阵. 当线性空间V和W的基选定后, 显然这种对应关系是一一对应, 即$L(V,W)$和$F^{m\times n}$同构.(读者参照线性变换与矩阵对应自行证明)

5.5.2 奇异值

在讨论线性变换时, 通过不变子空间概念导出方阵的特征值问题, 而且利用特征值、特征向量可实现矩阵的对角化. 但线性映射对应的矩阵不一定是方阵, 因此没有特征值、特征向量的概念, 但可以模仿特征值导出类似的概念帮助理解线性映射, 也可用它进行类似的矩阵“对角化”.

例 26. *如矩阵*$\mathbf{A}=\begin{bmatrix}4&11&14\\8&7&-2\end{bmatrix}$ *对应的是*$\mathbb{R}^3$*到*$\mathbb{R}^2$ *的一个线性映射, 对*$\mathbf{x}$ *经过映射后对应*$\mathbb{R}^2$ *中的向量*$\mathbf{A}\mathbf{x}$*. 对于*$\mathbb{R}^3$ *中的单位球面*$\{\mathbf{x}|\ \|\mathbf{x}\|=1\}$*, 变换后是*$\mathbf{R}^2$*中以原点为中心的椭圆, 它长轴的长度可通过求*$\|\mathbf{A}\mathbf{x}\|^2$*的最大值来确定, 也就是求*$(\mathbf{A}\mathbf{x},\mathbf{A}\mathbf{x})$*在*$\|\mathbf{x}\|=1$*约束下的最大化值, 即*

$$\max_{\|x\|=1}(\mathbf{A}\mathbf{x},\mathbf{A}\mathbf{x})=\mathbf{x}^T\left(\mathbf{A}^T\mathbf{A}\right)\mathbf{x}$$

通过求$\mathbf{A}^T\mathbf{A}$*的特征值* $\lambda_1=360$, $\lambda_2=90$,$\lambda_3=0$, *对应的特征向量为*

$$\mathbf{x}_1=\frac{1}{3}\begin{bmatrix}1\\2\\2\end{bmatrix},\quad \mathbf{x}_2=\frac{1}{3}\begin{bmatrix}-2\\-1\\2\end{bmatrix},\quad \mathbf{x}_3=\frac{1}{3}\begin{bmatrix}2\\-2\\1\end{bmatrix}$$

*可知最大值为*360, *对应长轴的长度为*$\sqrt{360}=6\sqrt{10}$, *再由*$\mathbf{A}\mathbf{x}_1$ *得到长轴对应的向量. 椭圆的短轴长度为*$\sqrt{90}=3\sqrt{10}$, *短轴对象向量有*$\mathbf{A}\mathbf{x}_2$*求得.*

对于矩阵$\mathbf{A}\in\mathbb{R}^{m\times n}$, $\mathbf{A}^T\mathbf{A}$ 为n阶对称矩阵, 设它对应的特征值和正交标准化的特征向量为$\lambda_1,\lambda_2,\ldots,\lambda_n$ 和 $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n$, 令$\mathbf{V}=\begin{bmatrix}\mathbf{v}_1&\mathbf{v}_2&\cdots&\mathbf{v}_n\end{bmatrix}$, 则$\mathbf{V}$是正交矩阵, 且有下列关系式:

$$\mathbf{A}^T\mathbf{A}\mathbf{v}_i=\lambda_i\mathbf{v}_i,\quad 或\mathbf{v}_i^T\mathbf{A}^T\mathbf{A}\mathbf{v}_i=\lambda_i$$

则$\lambda_i=\|\mathbf{A}\mathbf{v}_i\|^2(i=1,2,\ldots,n)$, 显然 $\lambda_i\geq 0$. 假设特征值按非升序排列, 即

$$\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$$

并设$\operatorname{rank}(\mathbf{A})=r\leq\min\{m,n\}$, 则 $\lambda_1\geq\lambda_2\geq\cdots\lambda_r>0$, 而$\lambda_{r+1}=\lambda_{r+2}=\ldots=\lambda_n=0$. 令

$$\omega_i=\sqrt{\lambda_i}=\|\mathbf{A}\mathbf{v}_i\|,\quad i=1,2,\ldots,r$$

称ω_i为矩阵$\mathbf{A}$的**奇异值**. 将对应非零奇异值的向量 $\mathbf{A}\mathbf{v}_i(1\leq i\leq r)$单位化, 即令

$$\mathbf{u}_i=\frac{\mathbf{A}\mathbf{v}_i}{\|\mathbf{A}\mathbf{v}_i\|}=\frac{1}{\omega_i}\mathbf{A}\mathbf{v}_i\Rightarrow\mathbf{A}\mathbf{v}_i=\omega_i\mathbf{u}_i,\quad i=1,2,\ldots,r$$

将标准正交向量组 $\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_r\}$ 扩充成$\mathbb{R}^m$的标准正交基 $\{\mathbf{u}_1,\ldots,\mathbf{u}_r,\mathbf{u}_{r+1},\ldots,\mathbf{u}_m\}$, 并令

$$\mathbf{U}=\begin{bmatrix}\mathbf{u}_1&\mathbf{u}_2&\cdots&\mathbf{u}_r&\mathbf{u}_{r+1}&\cdots&\mathbf{u}_n\end{bmatrix}$$

$$\Sigma=\begin{bmatrix}\omega_1&&&&\\&\omega_2&&\mathbf{0}&\\&\mathbf{0}&\ddots&&\mathbf{0}\\&&&\omega_r&\\&&\mathbf{0}&&\mathbf{0}\end{bmatrix}_{m\times n}$$

所以有
$$\mathbf{U}\Sigma = \mathbf{A}\mathbf{V} \tag{5.33}$$
而将$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ 称为矩阵$\mathbf{A}$的奇异值分解.

求$m \times n$阶矩阵$\mathbf{A}$奇异值分解的步骤:

(1) 求矩阵$\mathbf{A}^T\mathbf{A}$的特征值λ_i和标准正交的特征向量$v_i(i = 1, 2, \ldots, n)$.

(2) 将特征值按非降序排列后, 求非零奇异值 $\omega_i = \sqrt{\lambda_i}$, 按特征值(奇异值)和特征向量之间对应关系排列成矩阵形式$\Sigma_{m\times n}$ 和n阶正交矩阵 V.

(3) 求出非零奇异值对应的单位向量$\mathbf{u}_i = \dfrac{1}{\omega_i}\mathbf{A}\mathbf{v}_i$, 并将它们扩充成 $\mathbb{R}^m$空间的一个标准正交基. 按奇异值排列顺序构造相应矩阵$\mathbf{U}$.

例 27. *求矩阵*$\mathbf{A} = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$*的一个奇异值分解.*

解: (1). 求$\mathbf{A}^T\mathbf{A}$的特征值和特征向量: $\lambda_1 = 360, \lambda_2 = 90, \lambda_3 = 0$, 对应特征向量如例26中的$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

(2). 奇异值 $\omega_1 = 6\sqrt{10}, \omega_2 = 3\sqrt{10}$, 奇异值和特征向量矩阵为(注意对应关系)
$$\Sigma = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}, \quad \mathbf{V} = \frac{1}{3}\begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}$$

(3). 构造矩阵$\mathbf{U}$
$$\mathbf{u}_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ -3 \end{bmatrix}, \quad \mathbf{U} = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix}$$

最后得$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$.

习题五

1. 判断下列各变换是否为线性变换:

(a) 线性空间$\mathbb{R}^2$上的变换σ, 满足对$\forall\mathbf{x} \in \mathbb{R}^2$ 有
$$\sigma\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + ky \\ y \end{bmatrix}$$
k 为常数.

(b) 线性空间V中, 对$\forall\mathbf{x} \in V, \sigma(\mathbf{x}) = \mathbf{a}$, 其中 $\mathbf{a} \in V$ 为固定向量.

(c) 数域F上的多项式空间$F[x]_n$中,对$\forall f(x) \in F[x]_n$, 变换 $\sigma(f(x)) = f(x+1)$.

(d) 数域F上的多项式空间$F[x]_n$中,对$\forall f(x) \in F[x]_n$, 变换 $\sigma(f(x)) = f(x_0)$, 其中 $x_0 \in F$ 为固定常数.

(e) 数域F上的多项式空间$F[x]_n$中,对$\forall f(x) \in F[x]_n$, 变换 $\sigma(f(x)) = (x+a)\dfrac{d}{dx}f(x)$. 其中$a \in F$ 为固定常数.

(f) $F^{n\times n}$中, $\forall\mathbf{X} \in F^{n\times n}, \sigma(\mathbf{X}) = \mathbf{B}\mathbf{X}\mathbf{C}$, 其中 $\mathbf{B}, \mathbf{C} \in F^{n\times n}$为给定矩阵.

(g) $C[0, 2\pi]$是由区间$[0, 2\pi]$上所有连续函数构成的线性空间, 对$\forall f(t) \in C[0, 2\pi]$, 变换$\sigma$ 定义如下:
$$\sigma(f(t)) = \int_0^{2\pi} f(x)\sin(t-x)\,dx$$

(h) $\mathbb{R}^n$ 中, $\forall\mathbf{x} \in \mathbb{R}^n$, 变换$\sigma$定义如下:
$$\sigma\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} x_n \\ x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix}$$

2. 线性空间$\mathbb{R}^n$ 上的变换σ, 对$\forall \mathbf{x} \in \mathbb{R}^n$,

$$\sigma(\mathbf{x}) = (\mathbf{x}, \mathbf{a})\, \mathbf{a}$$

其中$\mathbf{a} \in \mathbb{R}^n$是给定向量, $(\bullet, \bullet)$表示向量内积. 证明: σ是线性变换.

3. 设$C[a,b]$表示闭区间$[a,b]$上所有连续函数构成的线性空间, 对$\forall f(x) \in C[a,b]$, 变换σ

$$\sigma(f(x)) = \int_a^x f(t)dt$$

证明: σ是线性变换.

4. 在三维几何空间, 单位向量 $\vec{i}, \vec{j}, \vec{k}$ 相互正交, 设变换$\sigma_i, \sigma_j, \sigma_k$ 分别表示 绕$\vec{i}, \vec{j}, \vec{k}$旋转$90°$.
   (a) 求证 $\sigma_i^4 = \sigma_j^4 = \sigma_k^4 = \mathbf{1}$(恒等变换).
   (b) 求证 $\sigma_i\sigma_j \neq \sigma_j\sigma_i$
   (c) 求证 $\sigma_i^2 \cdot \sigma_j^2 = \sigma_j^2 \cdot \sigma_i^2$.
   (d) 判断 $\sigma_i \cdot \sigma_j \cdot \sigma_i \cdot \sigma_j = \sigma_i^2 \cdot \sigma_j^2$ 是否成立.

5. 数域F上在所有关于x的多项式构成的空间$F[x]$内, $\forall f(x) \in F[x]$, 变换

$$\sigma(f(x)) = \frac{d}{dx}f(x), \quad \tau(f(x)) = xf(x)$$

求变换$\sigma \cdot \tau - \tau \cdot \sigma$.

6. 设线性空间V上的线性变换σ, τ满足$\sigma \cdot \tau - \tau \cdot \sigma = \mathbf{1}_V$($V$上恒等变换), 证明: $\sigma^m \cdot \tau - \tau \cdot \sigma^m = m\sigma^{m-1}$.

7. 平面解析几何中的切变换τ_k定义为: 对任意向量$\mathbf{a} \in \mathbb{R}^2$, $\tau_k : \mathbf{a} \mapsto \mathbf{b}$(如图)

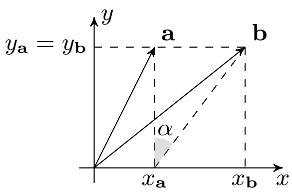


图: 切变换(习题7)

其中$k = \tan\alpha$为常数, 求变换对应的矩阵.

8. 线性空间$\mathbb{R}^3$上的线性变换σ, 它将三个线性无关的向量$\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, 分别变换成$\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$, 即 $\sigma(\mathbf{x}_i) = \mathbf{y}_i (i = 1,2,3)$. 向量$\mathbf{x}_i$在某组基$\mathfrak{B}$下的坐标为 $\mathbf{a}_i$, 向量$\mathbf{y}_i$在基$\mathfrak{B}$下坐标为 $\mathbf{b}_i (i = 1,2,3)$, 其中

$$\mathbf{a}_1 = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$
$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

   (a) 求线性变换σ在基$\mathfrak{B}$下对应的矩阵.
   (b) 设新的基$\mathfrak{B}' = \mathfrak{B}\mathbf{M}$, 其中

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

   坐标$\mathbf{a}_i$还是向量$\mathbf{x}_i$在旧基$\mathfrak{B}$ 下的坐标, 但坐标$\mathbf{b}_i$是向量$\mathbf{y}_i$ 在新基下的坐标. 求在新基$\mathfrak{B}'$下变换σ对应的矩阵.

9. 求线性空间$\mathbb{R}^3$上的两个线性变换, 它们将向量$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$变成$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$, 把$\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$变成$\begin{bmatrix} 1 & -1 & 0 \end{bmatrix}$.

10. 线性空间$\mathbb{R}^{2\times 2}$上的变换σ_1, σ_2定义如下: 对$\forall \mathbf{X} \in \mathbb{R}^{2\times 2}$

$$\sigma_1(\mathbf{X}) = \mathbf{AX}, \quad \sigma_2(\mathbf{X}) = \mathbf{XA}$$

其中$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$为给定矩阵. 求变换$\sigma_1, \sigma_2$在下列基下对应的矩阵.

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

11. 在$\mathbb{R}^2$中, 已知线性变换σ_1在基$\mathfrak{B}_1$下对应的矩阵为$\begin{bmatrix} 3 & 5 \\ 4 & 3 \end{bmatrix}$, 其中

$$\mathfrak{B}_1 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$$

线性变换σ_2在基$\mathfrak{B}_2$下的对应矩阵$\begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}$, 其中

$$\mathfrak{B}_2 = \left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \end{bmatrix} \right\}$$

(a) 求σ_1的像空间$Im(\sigma_1)$和σ_2的核空间$Ker(\sigma_2)$.

(b) 求线性变换$\sigma_1 + \sigma_2$在基$\mathfrak{B}_2$下对应的矩阵.

12. n维线性空间V上的所有线性变换关于变换的加法和数乘构成线性空间$L(V)$, 求空间$L(V)$的维数, 并给出它的一个基.

13. 线性空间$\mathbb{R}^{2\times 2}$上的线性变换σ定义如下: $\forall \mathbf{A} \in \mathbb{R}^{2\times 2}$,

$$\sigma(\mathbf{A}) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{A} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

求变换σ的秩和$\dim Ker(\sigma)$.

14. 设σ, τ是线性空间V上的线性变换, 并且$\sigma \cdot \tau = \tau \cdot \sigma$, 证明: $Im(\sigma)$和$Ker(\sigma)$都是τ-不变子空间.

15. 设S_1和S_2是线性空间V上线性变换σ的不变子空间, 证明: $S_1 + S_2$和$S_1 \cap S_2$也是σ-不变子空间.

16. 设σ是n维线性空间V上的幂等线性变换, 即$\sigma^2 = \sigma$,证明: 必定存在V的子空间W, 使得$\forall \mathbf{x} \in W, \sigma(\mathbf{x}) \in W$.

17. 线性空间$V = \mathbb{C}^n$上定义移位变换: $\sigma : \mathbb{C}^n \mapsto \mathbb{C}^n$,

$$\sigma\left(\begin{bmatrix} x_1 & x_2 & \cdots & x_{n-1} & x_n \end{bmatrix}^T\right) = \begin{bmatrix} x_2 & x_3 & \cdots & x_n & x_1 \end{bmatrix}^T$$

(a) 求证σ是线性变换, 并求它对应的矩阵$\mathbf{A}_\sigma$.

(b) 设$\omega_k = e^{j2k\pi/n}$, 用ω_k构造向量

$$\Omega_k = \begin{bmatrix} \omega_k^0 & \omega_k^1 & \omega_k^2 & \cdots & \omega_k^{n-2} & \omega_k^{n-1} \end{bmatrix}^T$$

其中$\omega_k^i (i = 0, 1, 2, \ldots, n-1)$是$\omega_k$的$i$次幂, 证明: ω_k和向量$\mathbf{\Omega}_k (k = 0, 1, 2, \ldots, n-1)$是移位变换的特征值和特征向量.

(c) 令变换$\tau = \sigma - \mathbf{1}_V$, 其中$\mathbf{1}_V$是$\mathbb{C}^n$上的恒等变换, 求$\tau$对应的矩阵.

(d) 向量$\Omega_k (k = 0, 1, 2, \ldots, n-1)$是否是变换$\tau$的特征向量? 如果是, 对应的特征值是什么?

(e) 设向量$\mathbf{a} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \end{bmatrix}^T$, 其中$a_i \in \mathbb{C}(i = 0, 1, 2, \ldots, n-1)$为常数, 考虑矩阵

$$\begin{bmatrix} \mathbf{a} & \sigma(\mathbf{a}) & \sigma^2(\mathbf{a}) & \cdots & \sigma^{n-1}(\mathbf{a}) \end{bmatrix}$$

试求它的特征值和特征向量.

18. 矩阵$\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$, 证明: $\mathbf{AB}$ 和$\mathbf{BA}$有相同的特征值.

19. 设$\mathbf{A} \in \mathbb{R}^{3\times 3}$, 若存在向量$\mathbf{x} \in \mathbb{R}^3$ 满足 $\mathbf{A}^3\mathbf{x} = 3\mathbf{A}\mathbf{x} - 2\mathbf{A}^2\mathbf{x}$, 且$\mathbf{x}, \mathbf{A}\mathbf{x}, \mathbf{A}^2\mathbf{x}$线性无关.

 (a) 令$\mathbf{P} = \begin{bmatrix} \mathbf{x} & \mathbf{A}\mathbf{x} & \mathbf{A}^2\mathbf{x} \end{bmatrix}$, 求矩阵$\mathbf{B}$ 使得$\mathbf{A} = \mathbf{PBP}^{-1}$.

 (b) 计算行列式$\det(\mathbf{A} + \mathbf{E})$.

20. 非零n元实向量组$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r (r < n)$两两正交,设 $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots, \mathbf{x}_r \end{bmatrix}$, n阶矩阵$\mathbf{A} = \mathbf{X}\mathbf{X}^T$, 求矩阵$\mathbf{A}$的特征值和特征向量.

21. 求下列矩阵的特征值和特征向量:

 (a) $\begin{bmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{bmatrix}$ (b) $\begin{bmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{bmatrix}$

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

22. 设n阶矩阵

$$\mathbf{A} = \begin{bmatrix} 1 & b & b & \cdots & b \\ b & 1 & b & \cdots & b \\ b & b & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & b \\ b & b & \cdots & b & 1 \end{bmatrix}$$

 (a) 求$\mathbf{A}$的特征值和特征向量.

 (b) 求可逆矩阵$\mathbf{P}$, 使得$\mathbf{P}^{-1}\mathbf{AP}$ 为对角阵.

23. 设λ为n阶可逆矩阵$\mathbf{A}$的特征值, 证明: λ^{-1} 是$\mathbf{A}^{-1}$的特征值.

24. 矩阵$\mathbf{A}_n$为如下的对称三对角阵:

$$\mathbf{A}_n = \begin{bmatrix} a & b & 0 & \cdots & \cdots & 0 \\ b & a & b & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & b & a & b \\ 0 & \cdots & \cdots & 0 & b & b \end{bmatrix}$$

 (a) 若$a = 0, b = 1$, 分别求$\mathbf{A}_2, \mathbf{A}_3$ 的特征值和特征向量.

 (b) 若$a = 0, b = 1$, 证明: $\lambda_k = 2\cos\left(\frac{k\pi}{n+1}\right)(k = 1, 2, \ldots, n)$ 是$\mathbf{A}_n$的特征值, 而向量

$$\begin{bmatrix} \sin\left(\frac{k\pi}{n+1}\right) & \sin\left(\frac{2k\pi}{n+1}\right) & \cdots & \sin\left(\frac{nk\pi}{n+1}\right) \end{bmatrix}^T$$

 是相应特征值λ_k的特征向量.

 (c) a, b为任意实数, 求$\mathbf{A}_n$的特征值特征向量.

25. 设$\mathbf{A} \in \mathbb{R}^{n\times n}$ 是对称矩阵, 证明: $\lambda = \max\limits_{\|\mathbf{x}\|=1} \mathbf{x}^T\mathbf{A}\mathbf{x}$ 是矩阵$\mathbf{A}$ 的最大特征值.

26. 设矩阵$\mathbf{A}=\begin{bmatrix}1&0&2\\0&-1&1\\0&1&0\end{bmatrix}$, 多项式为$g(\lambda)=2\lambda^8-3\lambda^5+\lambda^4+\lambda^2-4$.

(a) 求$g(\mathbf{A})$.

(b) 用$\mathbf{A}$的特征多项式表示$\mathbf{A}^{-1}$.

27. 若$\lambda_1,\lambda_2,\ldots,\lambda_n$ 是矩阵$\mathbf{A}$的特征值, $f(x)$是任意多项式, 证明: $f(\lambda_1),f(\lambda_2),\ldots,f(\lambda_n)$ 是矩阵 $f(\mathbf{A})$ 的特征值.

28. 分别在实数域和复数域判断下列矩阵是否相似于对角形:

$$(a)\quad\begin{bmatrix}5&2&-3\\4&5&-4\\6&4&-4\end{bmatrix}\qquad\qquad (b)\quad\begin{bmatrix}4&7&-5\\-4&5&0\\1&9&-4\end{bmatrix}$$

29. 已知矩阵$\mathbf{A}=\begin{bmatrix}2&0&0\\0&0&1\\0&1&x\end{bmatrix}$和 $\mathbf{B}=\begin{bmatrix}2&0&0\\0&y&1\\0&0&-1\end{bmatrix}$相似. 求$x$和$y$以及可逆矩阵$\mathbf{P}$, 它使得$\mathbf{P}^{-1}\mathbf{AP}=\mathbf{B}$.

30. 证明下列三个矩阵中任意两个都不相似:

$$\mathbf{A}=\begin{bmatrix}a&0&0\\0&a&0\\0&0&a\end{bmatrix}\quad \mathbf{B}=\begin{bmatrix}a&1&0\\0&a&1\\0&0&a\end{bmatrix}\quad \mathbf{C}=\begin{bmatrix}a&0&0\\0&a&1\\0&0&a\end{bmatrix}$$

31. 设n阶矩阵 $\mathbf{A}$ 与对角阵$\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$相似, 证明: 对任意一个正整数$m$, 矩阵 $(\lambda_i\mathbf{E}-\mathbf{A})^m$ 与矩阵 $(\lambda_i\mathbf{E}-\mathbf{A})$ 的秩相等.$(i=1,2,\ldots,n)$

32. 求正交矩阵$\mathbf{P}$, 使下列矩阵的对角化:

$$\text{(a)}\quad\begin{bmatrix}2&2&0\\-2&1&-2\\0&-2&0\end{bmatrix}\qquad \text{(b)}\quad\begin{bmatrix}3&1&0&-1\\1&3&-1&0\\0&-1&3&1\\-1&0&1&3\end{bmatrix}$$

33. 在$\mathbf{R}^3$中, 设向量$\mathbf{x}=\begin{bmatrix}2&-4&4\end{bmatrix}^T$,

(a) 分别采用旋转变换和镜像变换, 将向量$\mathbf{x}$ 变换成$\begin{bmatrix}6&0&0\end{bmatrix}^T$, 写出相应的变换矩阵.

(b) 方法是否可以推广到$\mathbb{R}^n$空间? 即对于非零向量 $\mathbf{y}=\begin{bmatrix}y_1&y_2&\cdots&y_n\end{bmatrix}^T$, 分别用上述两种正交变换, 将$\mathbf{y}$ 变换成 $\begin{bmatrix}\|\mathbf{y}\|&0&\cdots&0\end{bmatrix}^T$, 如果可行, 请叙述正交矩阵的设计步骤.

34. 设 $\mathbf{P}_1,\mathbf{P}_2$ 是线性空间V上的投影矩阵, 证明: $\mathbf{P}=\mathbf{P}_1-\mathbf{P}_2$ 是投影矩阵的充要条件是 $\mathbf{P}_1\mathbf{P}_2=\mathbf{P}_2\mathbf{P}_1=\mathbf{P}_2$.

35. 线性空间$\mathbb{R}^4$的子空间$S=\operatorname{span}(\varepsilon_1,\varepsilon_2)$, 其中

$$\varepsilon_1=\begin{bmatrix}1\\0\\2\\-1\end{bmatrix},\quad \varepsilon_2=\begin{bmatrix}0\\-2\\1\\1\end{bmatrix}$$

则$\mathbb{R}^4=S\oplus S^{\perp}$, 求将任意向量$\mathbf{x}\in\mathbb{R}^4$ 正交投影到$S^{\perp}$ 的正交投影矩阵$\mathbf{P}_{S^{\perp}}$, 并求向量$\begin{bmatrix}1&2&3&4\end{bmatrix}$在$S^{\perp}$ 上的投影.

36. 设$\mathbf{A}\in\mathbb{C}^{m\times n}$ 且$\operatorname{rank}(\mathbf{A})=r>0$, $\mathbf{A}$ 的奇异值分解为

$$\mathbf{A}=\mathbf{U}\Lambda\mathbf{V}^H,\quad 其中\Lambda=\begin{bmatrix}\Sigma_r&\mathbf{0}\\\mathbf{0}&\mathbf{0}\end{bmatrix}_{m\times n}$$

求矩阵$\mathbf{B}=\begin{bmatrix}\mathbf{A}\\\mathbf{A}\end{bmatrix}$的奇异值分解.

37. 求下列矩阵的奇异值分解:

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$