第三章 矩阵分解

3.7 谱分解与幂等阵

定义3.7.1(单纯矩阵谱分解) 设$\lambda_1,\cdots,\lambda_m$是单纯矩阵$A$的$m$个互异特征值,其代数重数分别为$d_1,\cdots,d_m$,则矩阵$A$的谱分解式定义为

$$A=\sum_{j=1}^{m}\lambda_j E_j$$

其中,$E_j,j=1,\cdots,m$,称为A的谱阵.

注1:单纯矩阵特征值的代数重数等于其几何重数.

思考:单纯矩阵的谱分解存在吗?

例3.7.1 求矩阵$A=\begin{bmatrix}4 & 6 & 0\\ -3 & -5 & 0\\ -3 & -6 & 1\end{bmatrix}$的谱分解.

分析:矩阵A的特征值为

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

对应的特征向量分别为

$$\boldsymbol{\alpha}_1=[2,-1,0]^T,\boldsymbol{\alpha}_2=[0,0,1]^T,\boldsymbol{\alpha}_3=[-1,1,1]^T$$

故A是单纯矩阵.

例3.7.1(续)

分析: 对A作对角化分解$A=P\Lambda P^{-1}$, 其中

$$\Lambda=\mathrm{diag}(1,1,-2)$$

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

$$P\Lambda=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}1&0&0\\0&1&0\\0&0&-2\end{bmatrix}$$

$$=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,-2\boldsymbol{\alpha}_3]$$

例3.7.1(续)

分析: 定义$P^{-1}=[\boldsymbol{\beta}_1^H,\boldsymbol{\beta}_2^H,\boldsymbol{\beta}_3^H]^H$,

$$P\Lambda P^{-1}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,-2\boldsymbol{\alpha}_3]\begin{bmatrix}\boldsymbol{\beta}_1^H\\ \boldsymbol{\beta}_2^H\\ \boldsymbol{\beta}_3^H\end{bmatrix}$$

$$=\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H+\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H-2\boldsymbol{\alpha}_3\boldsymbol{\beta}_3^H$$

$$\triangleq E_1-2E_2$$

注2: 上述过程对任意单纯阵都适应, 故单纯阵的谱分解总存在.

考查谱阵性质: 以例3.7.1中E_1和E_2为例.

$$E_1=\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H+\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H, E_2=\boldsymbol{\alpha}_3\boldsymbol{\beta}_3^H$$

首先计算$\boldsymbol{\beta}_i^H\boldsymbol{\alpha}_j, i,j=1,2,3.$

根据$P^{-1}P=I$知,

$$\begin{bmatrix}\boldsymbol{\beta}_1^H\\ \boldsymbol{\beta}_2^H\\ \boldsymbol{\beta}_3^H\end{bmatrix}[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]=I.$$

$$\boldsymbol{\beta}_i^H\boldsymbol{\alpha}_j=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}$$

考查谱阵性质: 分析E_iE_j

$$E_1^2=(\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H+\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H)(\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H+\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H)$$
$$=\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H+\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H$$
$$+\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H+\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H$$
$$=\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H+\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H=E_1$$

同理可得

$$E_2^2=E_2$$

$$E_1E_2=E_2E_1=0.$$

定义3.7.2(幂等矩阵) 设$E \in \mathbb{C}^{n \times n}$，若$E^2 = E$，则称$E$为**幂等矩阵**.

单纯矩阵的谱阵性质：设$E_1, \cdots, E_m$是单纯矩阵A的m个谱阵，则$\forall i, j = 1, \cdots, m$且$i \neq j$，有

（1）$E_j = (E_j)^2$；

（2）$E_i E_j = 0$；

考查矩阵AE_i:

$$AE_i = \left(\sum_{k=1}^{3} \lambda_k E_k\right) E_i$$

$$= \sum_{k=1}^{3} \lambda_k E_k E_i = \lambda_i E_i$$

同理可得$E_i A = AE_i = \lambda_i E_i$.

单纯矩阵的谱阵性质：设$E_1,\cdots,E_m$是单纯矩阵A的m个谱阵，则$\forall i,j=1,\cdots,m$且$i\neq j$，有

（1）$E_j=(E_j)^2$；

（2）$E_iE_j=0$；

（3）$E_iA=AE_i=\lambda_iE_i$；

根据$PP^{-1}=I$有

$$[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}\boldsymbol{\beta}_1^H\\\boldsymbol{\beta}_2^H\\\boldsymbol{\beta}_3^H\end{bmatrix}=I$$

$$\sum_{i=1}^{3}\boldsymbol{\alpha}_i\boldsymbol{\beta}_i^H=I$$

$$E_1+E_2=I$$

命题3.7.1(单纯矩阵的谱阵性质) 设$E_1,\cdots,E_m$是单纯矩阵A的m个谱阵,则$\forall i,j=1,\cdots,m$且$i\neq j$,有

（1）$E_j=(E_j)^2$;

（2）$E_iE_j=0$;

（3）$E_iA=AE_i=\lambda_iE_i$;

（4）$\sum_{k=1}^{m}E_k=I$;

（5）谱阵集$\{E_1,\cdots,E_m\}$唯一.

例3.7.2 设$A=\begin{bmatrix}2 & 1\\ 2 & 3\end{bmatrix}$, 求$A^{100}$.

解: 矩阵A的谱分解式为

$$A=\boldsymbol{\alpha}_1\boldsymbol{\beta}_1^H+4\boldsymbol{\alpha}_2\boldsymbol{\beta}_2^H\triangleq E_1+4E_2$$

$$\begin{aligned}A^2&=(E_1+4E_2)^2\\&=E_1^2+8E_1E_2+16E_2^2\\&=E_1+16E_2\end{aligned}$$

$$A^k=E_1+4^kE_2$$

当$k=100$时,

$$A^{100}=E_1+4^{100}E_2$$

定理3.7.1 设$f(\lambda)$为数域$\mathbb{C}$上的任一多项式，单纯阵$A \in \mathbb{C}^{n\times n}$的谱分解式为$A=\sum_{j=1}^{m}\lambda_j E_j$，则

$$f(A)=\sum_{j=1}^{m} f(\lambda_j)E_j$$

式中，$\lambda_1,\cdots,\lambda_m$为$A$阵的$m$个互异特征值，$E_j, j=1,\cdots,m$是$A$的谱阵. 特别地，当$f(\lambda)=\lambda^k, k\in\mathbb{N}$时，

$$A^k=\sum_{j=1}^{m}\lambda_j^k E_j$$

思考: 能否直接求解出矩阵的谱阵呢?

定理3.7.2 设单纯阵A的谱分解式为$A=\sum_{j=1}^{m}\lambda_j E_j$，

则$\forall i=1,\cdots,m$,

$$E_i=\frac{1}{\prod_{\substack{l=1,\\ l\neq i}}^{m}(\lambda_i-\lambda_l)}\prod_{\substack{l=1,\\ l\neq i}}^{m}(A-\lambda_l I)$$

注3: A的最小多项式为$m_A(\lambda)=\prod_{l=1}^{m}(\lambda-\lambda_l)$; 定义

$$f_i(\lambda)=\prod_{l=1,l\neq i}^{m}(\lambda-\lambda_l)$$

则$E_i=f_i(A)/f_i(\lambda_i)$.

例3.7.2(续) 求$A=\begin{bmatrix}2 & 1\\2 & 3\end{bmatrix}$的谱分解.

解: 矩阵A的两个特征值分别为$\lambda_1=1,\lambda_2=4$, 则其最小多项式为

$$m_A(\lambda)=(\lambda-1)(\lambda-4)\,.$$

定义

$$f_1(\lambda)=\lambda-4,\, f_2(\lambda)=\lambda-1$$

由此

$$f_1(A)=A-4I,\, f_2(A)=A-I$$

例3.7.2(续) 求$A=\begin{bmatrix}2 & 1\\ 2 & 3\end{bmatrix}$的谱分解.

解:

$$E_1=\frac{f_1(A)}{f_1(\lambda_1)}=-\frac{1}{3}(A-4I)=\frac{1}{3}\begin{bmatrix}2 & -1\\ -2 & 1\end{bmatrix}$$

$$E_2=\frac{f_2(A)}{f_2(\lambda_2)}=\frac{1}{3}(A-I)=\frac{1}{3}\begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}$$

正规矩阵A的酉对角化分解式为:

$$U^H AU = \mathrm{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{d_1},\cdots,\underbrace{\lambda_m,\cdots,\lambda_m}_{d_m}).$$

式中, $\lambda_1,\cdots,\lambda_m$是矩阵$A$的$m$个互异特征值, 其代数重数分别为$d_1,\cdots,d_m$, 酉矩阵$U$定义为

$$U = [\boldsymbol{u}_{11},\cdots,\boldsymbol{u}_{1d_1},\cdots,\boldsymbol{u}_{m1},\cdots,\boldsymbol{u}_{md_m}] \in \mathbb{C}^{n\times n}$$

其中, $\boldsymbol{u}_{j1},\cdots,\boldsymbol{u}_{jd_j}$是属于特征值$\lambda_j$的$d_j$个单位正交的特征向量, $j=1,\cdots,m$.

正规矩阵A的谱分解式为:

$$
\begin{aligned}
A &= U\mathrm{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{d_1},\cdots,\underbrace{\lambda_m,\cdots,\lambda_m}_{d_m})\,U^H \\
&= \lambda_1 \boldsymbol{u}_{11}\boldsymbol{u}_{11}^H + \cdots + \lambda_m \boldsymbol{u}_{md_m}\boldsymbol{u}_{md_m}^H \\
&= \lambda_1\left(\sum_{i=1}^{d_1} \boldsymbol{u}_{1i}\boldsymbol{u}_{1i}^H\right) + \cdots + \lambda_m\left(\sum_{i=1}^{d_m} \boldsymbol{u}_{mi}\boldsymbol{u}_{mi}^H\right) \\
&\triangleq \lambda_1 E_1 + \cdots + \lambda_m E_m
\end{aligned}
$$

正规阵的谱阵性质：设复正规阵A有谱分解式$A=\sum_{j=1}^{m}\lambda_j E_j$，其中，$\lambda_1,\cdots,\lambda_m$是$A$的$m$个互异特征值，$E_1,\cdots,E_m$是对应谱阵，则$\forall i,j=1,\cdots,m$且$i\neq j$，有

（1）$E_j=(E_j)^2$；

（2）$E_iE_j=0$；

（3）$E_iA=AE_i=\lambda_iE_i$；

（4）$\sum_{k=1}^{m}E_k=I$；

（5）谱阵集合$\{E_1,\cdots,E_m\}$唯一.

考查谱阵

$$E_j = \sum_{i=1}^{d_j} \boldsymbol{u}_{ji}\boldsymbol{u}_{ji}^H$$

$$(E_j)^H = \sum_{i=1}^{d_j} \boldsymbol{u}_{ji}\boldsymbol{u}_{ji}^H = E_j$$

定义3.7.3(**正交投影矩阵**) 设$E \in \mathbb{C}^{n\times n}$, 若

$$E^2 = E = E^H$$

则称E为**正交投影矩阵**.

命题3.7.2(正规阵的谱阵性质) 设复正规阵A有谱分解式$A=\sum_{j=1}^{m}\lambda_j E_j$，其中，$\lambda_1,\cdots,\lambda_m$是$A$的$m$个互异特征值，$E_1,\cdots,E_m$是$A$的$m$个谱阵，则$\forall i,j=1,\cdots,m$且$i\neq j$，有

（1）$E_j=(E_j)^2=E_j^H$；

（2）$E_iE_j=0$；

（3）$E_iA=AE_i=\lambda_iE_i$；

（4）$\sum_{k=1}^{m}E_k=I$；

（5）谱阵集合$\{E_1,\cdots,E_m\}$唯一.

命题3.7.3(幂等阵性质) 若$E \in \mathbb{C}_r^{n\times n}$是幂等阵，则

（1）E^H, E^*和$I-E$都是幂等矩阵；

这表明幂等矩阵“诱导”出的矩阵仍是幂等矩阵.

命题3.7.3(幂等阵性质) 若$E \in \mathbb{C}_r^{n\times n}$是幂等阵，则

（2）E为单纯阵且相似于对角阵$\Lambda = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$;

分析: $\varphi(\lambda) = \lambda(\lambda - 1)$是$E$的零化多项式;

E有r个1特征值和$n-r$个零特征值.

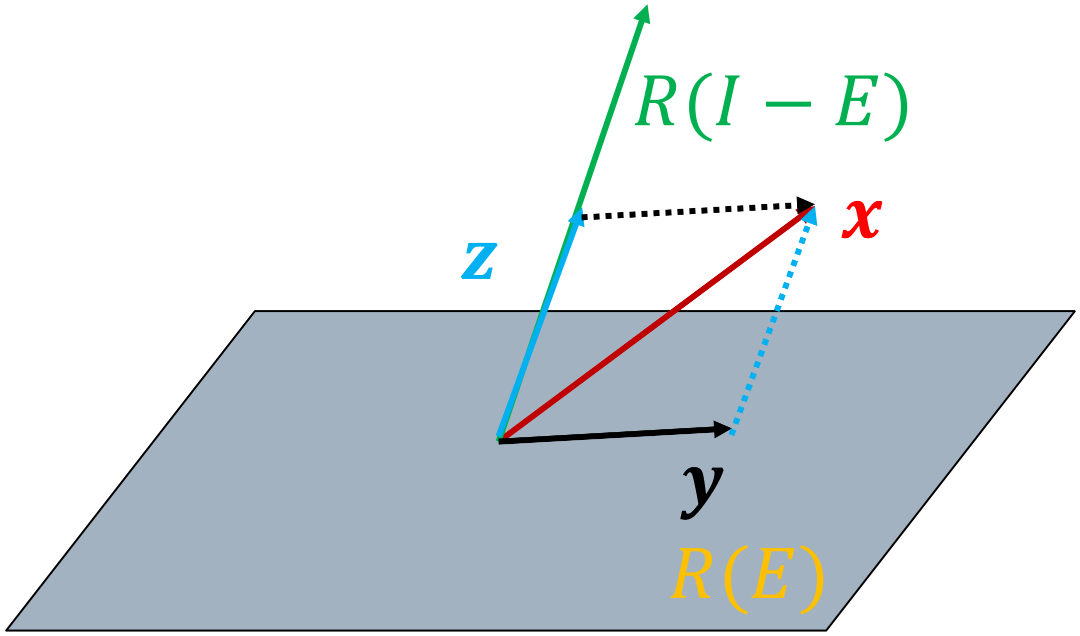
$$\mathrm{tr}(E) = \mathrm{tr}(\Lambda) = r = \mathrm{rank}(E)$$

当E为满秩矩阵时,

$$E = P^{-1} I_n P = I_n$$

命题3.7.3(幂等阵性质) 若$E \in \mathbb{C}_r^{n\times n}$是幂等阵，则

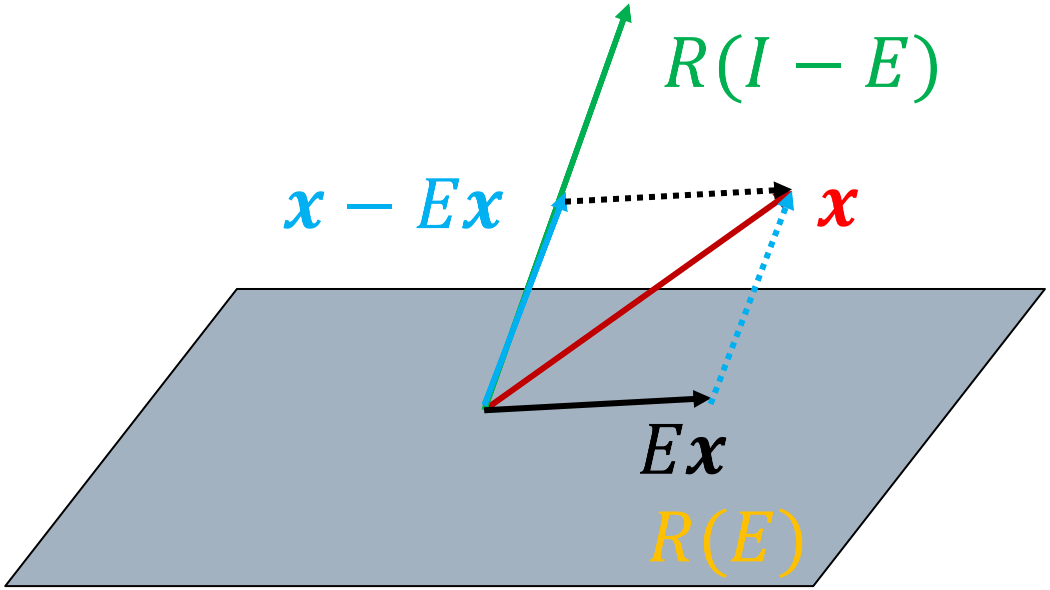
（3）$\mathbb{C}^n = R(E)\dot{+}N(E); N(E) = R(I-E)$.



$$x = y + z$$

$$\boldsymbol{x} = E\boldsymbol{x} + (I - E)\boldsymbol{x}$$

式中, $E\boldsymbol{x} \in R(E), (I - E)\boldsymbol{x} \in R(I - E).$



- 幂等矩阵也称为投影矩阵.

思考：如何构造幂等矩阵呢？

回顾例1.6.3 设$V = W_1 \dot{+} W_2$，其中

$$W_1 = \mathrm{span}\{[1,0,1]^T\}$$

$$W_2 = \{(x_1, x_2, x_3)^T | x_3 = 0, \forall x_1, x_2 \in \mathbb{R}\}$$

试求向量$\boldsymbol{x} = [1,1,1]^T$在$W_2$上的投影.

思路：

W_1的基$\boldsymbol{\varepsilon}_1$连同$W_2$的基$\boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3$当作某3阶矩阵的特征向量，其中$\boldsymbol{\varepsilon}_1$，$\boldsymbol{\varepsilon}_2$和$\boldsymbol{\varepsilon}_3$分别当作属于不同特征值的特征向量. 由此，分别构造谱阵.

思考:如何构造幂等矩阵呢?

$$P=[\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3]=\begin{bmatrix}1&1&0\\0&0&1\\1&0&0\end{bmatrix}$$

$$P^{-1}=\begin{bmatrix}0&0&1\\1&0&-1\\0&1&0\end{bmatrix}=\begin{bmatrix}\boldsymbol{\beta}_1^H\\\boldsymbol{\beta}_2^H\\\boldsymbol{\beta}_3^H\end{bmatrix}$$

$$E_1=\boldsymbol{\varepsilon}_1\boldsymbol{\beta}_1^H=\begin{bmatrix}0&0&1\\0&0&0\\0&0&1\end{bmatrix}$$

$$E_2=\boldsymbol{\varepsilon}_2\boldsymbol{\beta}_2^H+\boldsymbol{\varepsilon}_3\boldsymbol{\beta}_3^H=\begin{bmatrix}1&0&-1\\0&1&0\\0&0&0\end{bmatrix}$$

思考:如何构造幂等矩阵呢?

$$E_1\boldsymbol{x}=\begin{bmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1\\ 1\\ 1\end{bmatrix}=\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}$$

$$E_2\boldsymbol{x}=\begin{bmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1\\ 1\\ 1\end{bmatrix}=\begin{bmatrix}0\\ 1\\ 0\end{bmatrix}$$

命题3.7.3(幂等阵性质) 若$E \in \mathbb{C}_r^{n\times n}$是幂等阵,则

（4）$E\boldsymbol{x} = \boldsymbol{x} \Leftrightarrow \boldsymbol{x} \in R(E)$,其中$\boldsymbol{x} \in \mathbb{C}^n$.

证明: 由$E\boldsymbol{x} = \boldsymbol{x}$知,

$$\boldsymbol{x} \in R(E)$$

反之,若$\boldsymbol{x} \in R(E)$,则$\exists \boldsymbol{y} \in \mathbb{C}^n$满足$\boldsymbol{x} = E\boldsymbol{y}$,故

$$E\boldsymbol{x} = E(E\boldsymbol{y}) = E^2\boldsymbol{y} = E\boldsymbol{y} = \boldsymbol{x}.$$

注1: 性质(4)的几何解释为向量$\boldsymbol{x} \in R(E)$当且仅当向量$\boldsymbol{x}$在空间$R(E)$的投影恰为它本身.

思考：为什么Hermite幂等阵称为正交投影阵？

分析：考查$\mathbb{C}^n = R(E)\dot{+}N(E)$，

E^2=E证维数加和=n

$\forall \boldsymbol{x} \in R(E), \boldsymbol{y} \in N(E)$

$$(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{y}^H \boldsymbol{x} = \boldsymbol{y}^H E \boldsymbol{z} = (E\boldsymbol{y})^H \boldsymbol{z} = 0$$

式中，$\boldsymbol{x} = E\boldsymbol{z}$.

因此，

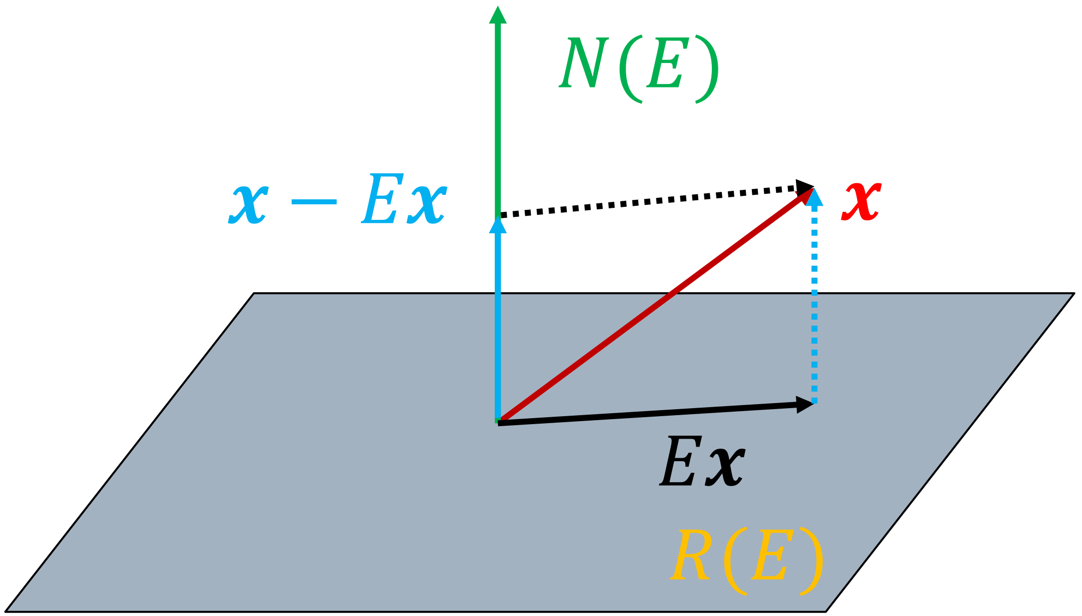
$$R(E) \perp N(E)$$

$$\mathbb{C}^n = R(E) \oplus N(E)$$

（a，a）=0，长度为0，因此a为零向量。

思考:为什么Hermite幂等阵称为正交投影阵?

分析: $\mathbb{C}^n = R(E) \oplus N(E)$



例3.7.3 求向量$\boldsymbol{b}\in\mathbb{C}^n$在$V_m=\mathrm{span}(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$上的正交投影，其中向量组$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$是$\mathbb{C}^n$空间的$m$个线性无关向量，$m\le n$.

分析：以$V_m=V_1$为例，

$$\begin{aligned}\mathrm{Proj}_{V_1}\boldsymbol{b}&=\frac{(\boldsymbol{b},\boldsymbol{a}_1)}{(\boldsymbol{a}_1,\boldsymbol{a}_1)}\boldsymbol{a}_1\\&=(\boldsymbol{a}_1^H\boldsymbol{b})(\boldsymbol{a}_1^H\boldsymbol{a}_1)^{-1}\boldsymbol{a}_1\\&=\boldsymbol{a}_1(\boldsymbol{a}_1^H\boldsymbol{a}_1)^{-1}\boldsymbol{a}_1^H\boldsymbol{b}\\&\triangleq E_1\boldsymbol{b}\end{aligned}$$

例 3.7.3(续)

分析: $\boldsymbol{b}$在V_1的正交投影:

$$\mathrm{Proj}_{V_1}\boldsymbol{b} = \boldsymbol{a}_1(\boldsymbol{a}_1^H\boldsymbol{a}_1)^{-1}\boldsymbol{a}_1^H\boldsymbol{b} \triangleq E_1\boldsymbol{b}$$

联想$\boldsymbol{b}$在V_m的正交投影:

定义$A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m] \in \mathbb{C}^{n\times m}$, 仿照$E_1$构造$V_m$中的正交投影矩阵

$$\boldsymbol{E_m} = A(A^HA)^{-1}A^H$$

思考: E_m有定义吗, 即A^HA可逆吗?

回答: 由于A列满秩, 故复方阵A^HA可逆.

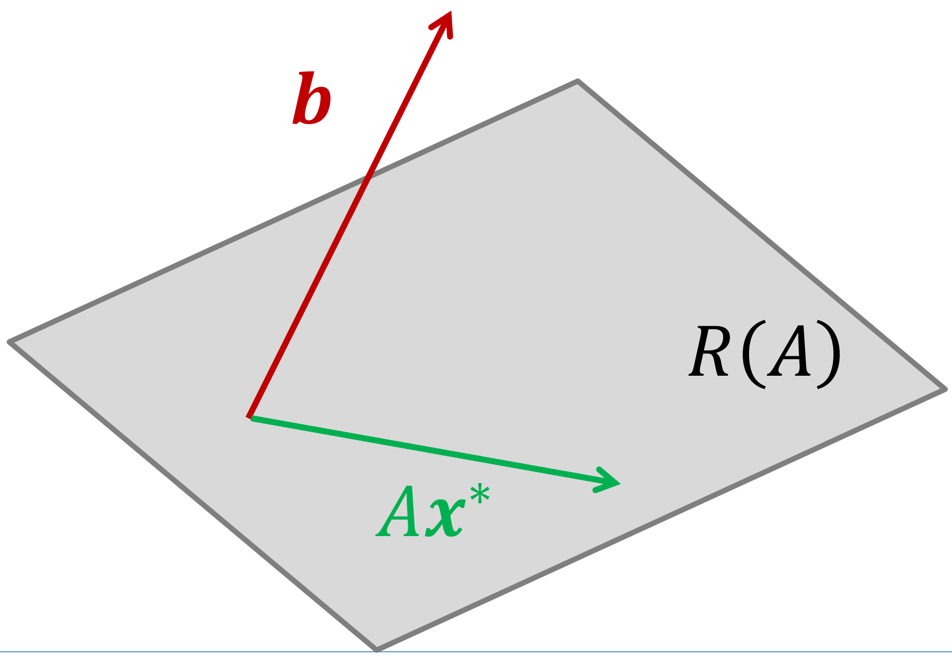
例3.7.3 (续)

再思考:E_m是正交投影矩阵吗?

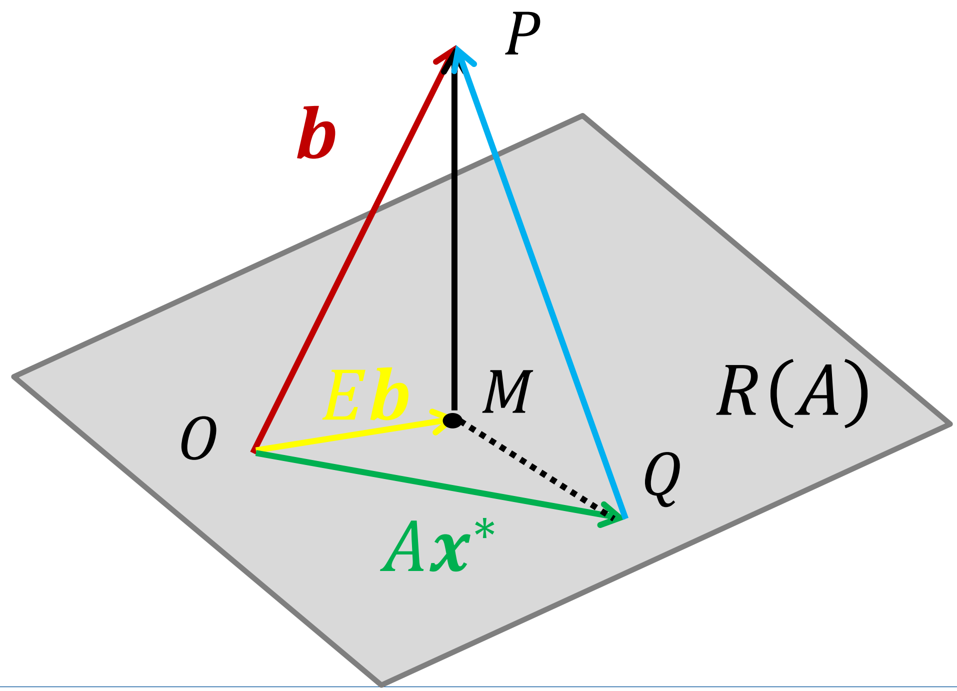
$$E_m^H=(A(A^HA)^{-1}A^H)^H$$
$$=A\left((A^HA)^{-1}\right)^HA^H$$
$$=A\left((A^HA)^H\right)^{-1}A^H$$
$$=A(A^HA)^{-1}A^H=E_m$$
$$E_m^2=A(A^HA)^{-1}A^HA(A^HA)^{-1}A^H$$
$$=A(A^HA)^{-1}A^H=E_m$$

因此,$E_m\boldsymbol{b}$一定是向量$\boldsymbol{b}$在V_m上的正交投影.

例3.7.4(最小二乘问题) 考查线性方程组$A\boldsymbol{x}=\boldsymbol{b}$，其中，$A\in\mathbb{C}_n^{m\times n}$和$\boldsymbol{b}\in\mathbb{C}^m$给定，$\boldsymbol{x}\in\mathbb{C}^n$待定. 求向量$\boldsymbol{x}$使得$\|A\boldsymbol{x}-\boldsymbol{b}\|$最小.



例3.7.4(**最小二乘问题**)考查线性方程组$A\boldsymbol{x}=\boldsymbol{b}$，其中，$A\in\mathbb{C}_n^{m\times n}$和$\boldsymbol{b}\in\mathbb{C}^m$给定，$\boldsymbol{x}\in\mathbb{C}^n$待定. 求向量$\boldsymbol{x}$使得$\|A\boldsymbol{x}-\boldsymbol{b}\|$最小.



例3.7.4(最小二乘问题)

分析: 求$\|A\boldsymbol{x}-\boldsymbol{b}\|$最小可等价转化为求解线性方程组

$$A\boldsymbol{x}=E\boldsymbol{b}=A(A^HA)^{-1}A^H\boldsymbol{b}$$

观察得$\boldsymbol{x}=(A^HA)^{-1}A^H\boldsymbol{b}$是方程的一个解.

若方程两端同时乘以A^H, 得

$$A^HA\,\boldsymbol{x}=A^H\boldsymbol{b}$$

由于A列满秩, 有唯一解

$$\boldsymbol{x}=(A^HA)^{-1}A^H\boldsymbol{b}$$

小结:

基本概念:单纯矩阵与正规矩阵谱分解、幂等矩阵、正交投影矩阵

重要结论:谱阵性质、幂等矩阵性质

重要计算:单纯矩阵与正规矩阵谱分解式、正交投影矩阵构造