1.考虑矩阵的Hardmard积$\mathbf{A}\odot\mathbf{B}$，即矩阵对应元素的乘积，证明：

1）如果$\mathbf{A},\mathbf{B},\mathbf{D}$都是$m\times m$矩阵，并且$\mathbf{D}$是对角阵，则$(\mathbf{DA})\odot(\mathbf{BD})$

$=\mathbf{D}(\mathbf{A}\odot\mathbf{B})\mathbf{D}$

2）$\mathrm{rank}(\mathbf{A}\odot\mathbf{B})\leq\mathrm{rank}(\mathbf{A})\mathrm{rank}(\mathbf{B})$

矩阵Hardmard积定义 $[\mathbf{A} \odot \mathbf{B}]_{ij} = a_{ij} b_{ij}$.

1）令 $\mathbf{D} = \mathrm{diag}\{d_1, d_2, \dots, d_m\}$，则

$$(\mathbf{DA}) \odot (\mathbf{BD}) = \begin{bmatrix} d_1 a_{11} & \cdots & d_1 a_{1m} \\ \vdots & \ddots & \vdots \\ d_m a_{m1} & \cdots & d_m a_{mm} \end{bmatrix} \odot \begin{bmatrix} d_1 b_{11} & \cdots & d_m b_{1m} \\ \vdots & \ddots & \vdots \\ d_1 b_{m1} & \cdots & d_m b_{mm} \end{bmatrix}$$

$$= \begin{bmatrix} d_1 a_{11} b_{11} d_1 & \cdots & d_1 a_{1m} b_{1m} d_m \\ \vdots & & \vdots \\ d_m a_{m1} b_{m1} d_1 & \cdots & d_m a_{mm} b_{mm} d_m \end{bmatrix} = \mathbf{D}(\mathbf{A} \odot \mathbf{B})\mathbf{D}$$

2）很容易证明 $\mathbf{A} \odot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \odot \mathbf{B} + \mathbf{A} \odot \mathbf{C}$ （*）

令$\mathbf{A}$的秩为r，B的秩为s，根据SVD分解有

$$\mathbf{A} = \sum_{k=1}^{r} \lambda_k \mathbf{x}_k \mathbf{y}_k^T, \mathbf{B} = \sum_{l=1}^{s} \gamma_l \mathbf{v}_l \mathbf{w}_l^T,$$

$$\mathbf{A} \odot \mathbf{B} = \left(\sum_{k=1}^{r} \lambda_k \mathbf{x}_k \mathbf{y}_k^T \right) \odot \left(\sum_{l=1}^{s} \gamma_l \mathbf{v}_l \mathbf{w}_l^T \right)$$

根据（*）将求和号和奇异值提出

$$\Rightarrow \text{上式} = \sum_k \sum_l \lambda_k \gamma_l (\mathbf{x}_k \mathbf{y}_k^T) \odot (\mathbf{v}_l \mathbf{w}_l^T)$$

$$= \sum_k \sum_l \lambda_k \gamma_l (\mathbf{x}_k \odot \mathbf{v}_l)(\mathbf{y}_k^T \odot \mathbf{w}_l^T)$$

$$\left((\mathbf{x}_k \mathbf{y}_k^T)\ \mathrm{e}\ (\mathbf{v}_l \mathbf{w}_l^T)\right)_{ij} = x_i y_j \cdot v_i w_j$$

$$= x_i v_i \cdot y_j w_j = \left((\mathbf{x}_k\ \mathrm{e}\ \mathbf{v}_l)(\mathbf{y}_k^T\ \mathrm{e}\ \mathbf{w}_l^T)\right)_{ij}$$

注意$\mathbf{x}_k, \mathbf{v}_l$是列向量，$\mathbf{y}_k^T, \mathbf{w}_l^T$是行向量

由$\operatorname{rank}\left((\mathbf{x}_k \odot \mathbf{v}_l) \odot \left(\mathbf{y}_k^T \odot \mathbf{w}_l^T\right)\right) \le 1$及$\operatorname{rank}(\mathbf{A}+\mathbf{B}) \le \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B})$

得证

$$\operatorname{rank}(\mathbf{A} \odot \mathbf{B}) \le \sum_k \sum_l \operatorname{rank}\left((\mathbf{x}_k \odot \mathbf{v}_l) \odot \left(\mathbf{y}_k^T \odot \mathbf{w}_l^T\right)\right) \le rs = \operatorname{rank}(\mathbf{A}) \operatorname{rank}(\mathbf{B})$$

2. 考虑矩阵$\mathbf{A}=[a_{mn}]_{M\times N}$,有p诱导范数定义$\|\mathbf{A}\|_p=\max_{\mathbf{x}}\frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p}$

1） 证明$\|\mathbf{A}\|_1=\max_{1\le n\le N}\sum_{m=1}^{M}|a_{mn}|$，$\|\mathbf{A}\|_\infty=\max_{1\le m\le M}\sum_{n=1}^{N}|a_{mn}|$

2） 给出$\|\mathbf{A}\|_0$背后的意义

答案:

求矩阵的p诱导范数，利用夹逼去证。

零诱导范数为A中列向量中最多非零元素的个数

2. 考虑矩阵$\mathbf{A}=\left[a_{mn}\right]_{M\times N}$，有p诱导范数定义 $\|\mathbf{A}\|_p=\max_{\mathbf{x}}\dfrac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}$

1） 证明$\|\mathbf{A}\|_1=\max_{1\le n\le N}\sum_{m=1}^{M}|a_{mn}|$， $\|\mathbf{A}\|_\infty=\max_{1\le m\le M}\sum_{n=1}^{N}|a_{mn}|$

2） 给出$\|\mathbf{A}\|_0$背后的意义

2)$\|\mathbf{A}\|_0=\max_{\mathbf{x}}\dfrac{\|\mathbf{A}\mathbf{x}\|_0}{\|\mathbf{x}\|_0}$, $\|\mathbf{x}\|_0$表示$\mathbf{x}$中非零元素的个数, $\|\mathbf{A}\mathbf{x}\|_0$表示$\sum_{n=1}^{N}a_{mn}x_n$中的非零元素个数，即$\mathbf{a}_m{}^T\mathbf{x}, 1\le m\le M$ 的非零元素个数。$\|\mathbf{A}\|_0$表示的是A的列向量中最多非零元素的个数，即$\|\mathbf{A}\|_0=\max_{n}\|\mathbf{a}_n\|_0$。

1) $\mathbf{A}=[\mathbf{a}_1,\mathbf{a}_2,\mathrm{L}\ \ ,\mathbf{a}_N],\mathbf{a}_n\in C^M,\mathbf{x}=[x_1,x_2,\mathrm{L}\ \ ,x_n]^T.$

①$$\|\mathbf{A}\|_1=\max_x\frac{\|\mathbf{A}\mathbf{x}\|_1}{\|\mathbf{x}\|_1}=\max_x\frac{\left\|\sum_{n=1}^{N}x_n\mathbf{a}_n\right\|_1}{\|\mathbf{x}\|_1}\le\max_x\frac{\sum_{n=1}^{N}\|x_n\mathbf{a}_n\|_1}{\|\mathbf{x}\|_1}=\max_x\frac{\sum_{n=1}^{N}|x_n|\|\mathbf{a}_n\|_1}{\|\mathbf{x}\|_1}\le\max_x\frac{\sum_{n=1}^{N}|x_n|\cdot\max_n\|\mathbf{a}_n\|_1}{\sum_{n=1}^{N}|x_n|}$$

$$=\max_x\max_n\|\mathbf{a}_n\|_1=\max_n\sum_{m=1}^{M}|a_{mn}|$$

②记 $k=\arg\max_n\sum_{m=1}^{M}|a_{mn}|$，存在 $\boldsymbol{\xi}=\mathbf{e}_k=(0,0,\mathrm{K}\ ,1,0,\mathrm{K}\ ,0)^T$，使得 $\frac{\|\mathbf{A}\boldsymbol{\xi}\|_1}{\|\boldsymbol{\xi}\|_1}=\max_n\sum_{m=1}^{M}|a_{mn}|$

因此，$\|\mathbf{A}\|_1=\max_{\mathbf{x}}\frac{\|\mathbf{A}\mathbf{x}\|_1}{\|\mathbf{x}\|_1}\ge\frac{\|\mathbf{A}\boldsymbol{\xi}\|_1}{\|\boldsymbol{\xi}\|_1}=\max_{\mathrm{n}}\sum_{m=1}^{M}|a_{mn}|$

由①和②，得 $\|\mathbf{A}\|_1=\max_n\sum_{m=1}^{M}|a_{mn}|$

3.证明矩阵迹的如下几个性质

1）证明$\mathrm{tr}\left(\mathbf{A}^2\right) \le \mathrm{tr}\left(\mathbf{A}^T\mathbf{A}\right)$,当且仅当$\mathbf{A}$是对称时取等

2）对于对称矩阵$\mathbf{A}$和$\mathbf{B}$,有$\mathrm{tr}\left(\mathbf{AB}\right) \le 0.5\,\mathrm{tr}\left(\mathbf{A}^2 + \mathbf{B}^2\right)$

3）$\mathrm{tr}\left(\mathbf{ABCD}\right) = \mathrm{vec}\left(\mathbf{D}\right)^T\left(\mathbf{A} \otimes \mathbf{C}^T\right)\mathrm{vec}\left(\mathbf{B}^T\right)$

1）$\mathrm{tr}\left(\mathbf{A}^2\right)=\sum_i^N\sum_j^N a_{ij}a_{ji},\mathrm{tr}\left(\mathbf{A}^T\mathbf{A}\right)=\sum_i^N\sum_j^N a_{ij}{}^2$

$\mathrm{tr}\left(\mathbf{A}^T\mathbf{A}\right)-\mathrm{tr}\left(\mathbf{A}^2\right)=\frac{1}{2}\sum_i^N\sum_j^N\left(a_{ij}-a_{ji}\right)^2\geq 0$

$\Rightarrow \mathrm{tr}\left(\mathbf{A}^T\mathbf{A}\right)\geq \mathrm{tr}\left(\mathbf{A}^2\right)$,当且仅当$a_{ij}=a_{ji}$时等号成立

2）由于$\mathbf{A}$、$\mathbf{B}$ 均为对称矩阵 $\Rightarrow a_{ij}=a_{ji},b_{ij}=b_{ji}$

$\mathrm{tr}\left(\mathbf{AB}\right)=\sum_i^N\sum_j^N a_{ij}b_{ji}=\sum_i^N\sum_j^N a_{ij}b_{ij},$

$\mathrm{tr}\left(\mathbf{A}^2+\mathbf{B}^2\right)=\mathrm{tr}\left(\mathbf{A}^T\mathbf{A}\right)+\mathrm{tr}\left(\mathbf{B}^T\mathbf{B}\right)=\sum_i^N\sum_j^N a_{ij}{}^2+\sum_i^N\sum_j^N b_{ij}{}^2,$

$\mathrm{tr}\left(\mathbf{A}^2+\mathbf{B}^2\right)-2\mathrm{tr}\left(\mathbf{AB}\right)=\sum_i^N\sum_j^N\left(a_{ij}-b_{ij}\right)^2\geq 0,$

$\Rightarrow \frac{1}{2}\mathrm{tr}\left(\mathbf{A}^2+\mathbf{B}^2\right)\geq \mathrm{tr}\left(\mathbf{AB}\right)$,当且仅当$a_{ij}=b_{ij}$时等号成立

3) $\operatorname{tr}\left(\mathbf{A}^T\mathbf{B}\right)=\sum\limits_{i,j} a_{ij}b_{ij}=\operatorname{vec}(\mathbf{A})^T\cdot\operatorname{vec}(\mathbf{B})$

令$\mathbf{B}=\left(\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_N\right),\mathbf{I}=\left(\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_N\right),\mathbf{B}=\mathbf{B}\mathbf{I}^T=\sum\limits_{j=1}^{N}\mathbf{b}_j\mathbf{e}_j^{\ T};$

$\operatorname{vec}(\mathbf{ABC})=\operatorname{vec}(\sum\limits_{j=1}^{N}\mathbf{A}\mathbf{b}_j\mathbf{e}_j^{\ T}\mathbf{C})=\sum\limits_{j=1}^{N}\operatorname{vec}((\mathbf{A}\mathbf{b}_j)\cdot(\mathbf{C}^T\mathbf{e}_j)^T)$ （1）

根据$\operatorname{vec}(\mathbf{a}\cdot\mathbf{b}^T)$=$\mathbf{b}\otimes\mathbf{a}$以及$\left(\mathbf{A}\otimes\mathbf{C}\right)\left(\mathbf{B}\otimes\mathbf{D}\right)=\left(\mathbf{AB}\right)\otimes\left(\mathbf{CD}\right)$

（1）式 $=\sum\limits_{j=1}^{N}(\mathbf{C}^T\mathbf{e}_j)\otimes(\mathbf{A}\mathbf{b}_j)=\left(\mathbf{C}^T\otimes\mathbf{A}\right)\cdot(\sum\limits_{j=1}^{N}\mathbf{e}_j\otimes\mathbf{b}_j)$

$$=\left(\mathbf{C}^T\otimes\mathbf{A}\right)\operatorname{vec}(\sum_{j=1}^{N}\mathbf{b}_j\mathbf{e}_j^{\ T})=\left(\mathbf{C}^T\otimes\mathbf{A}\right)\operatorname{vec}\left(\mathbf{B}\right)$$

根据$\operatorname{tr}\left(\mathbf{A}\right)=\operatorname{tr}\left(\mathbf{A}^T\right)$

$\operatorname{tr}\left(\mathbf{ABCD}\right)=\operatorname{tr}\left(\left(\mathbf{ABCD}\right)^T\right)=\operatorname{tr}\left(\mathbf{D}^T\cdot\mathbf{C}^T\mathbf{B}^T\mathbf{A}^T\right)=\operatorname{vec}(\mathbf{D})^T\cdot\operatorname{vec}(\mathbf{C}^T\mathbf{B}^T\mathbf{A}^T)$

$$=\operatorname{vec}\left(\mathbf{D}\right)^T\left(\mathbf{A}\otimes\mathbf{C}^T\right)\operatorname{vec}\left(\mathbf{B}^T\right)$$

4. 对于正定矩阵

1）对于$N \times N$的正定矩阵$\mathbf{A}$，证明 $\det(\mathbf{A}) \le \prod_{i=1}^{N} a_{ii}$

2）构造一个凸函数，使其Hessian矩阵为正定矩阵 $\mathbf{A}$

3）如果矩阵$\mathbf{A}$满足$\mathbf{A} = \mathbf{A}\mathbf{A}$且$\mathbf{A}$是Hermitian矩阵，则$\mathbf{A}$是半正定的

1）令$\mathbf{D}=diag\{a_{11}^{\frac{1}{2}},a_{22}^{\frac{1}{2}},\ldots,a_{nn}^{\frac{1}{2}}\}$，并有$\mathbf{C}=\mathbf{D}^{-1}\mathbf{A}\mathbf{D}^{-1}$，则$\mathbf{C}$正定且对角元素均为1

$$\det(\mathbf{C})=\prod\lambda_i\le\left(\frac{1}{n}\sum\lambda_i\right)^n=\left(\frac{1}{n}\mathrm{tr}(\mathbf{C})\right)^n=1,$$

$$\Rightarrow\det(\mathbf{A})=\det(\mathbf{DCD})=\det(\mathbf{C})\left(\prod_i a_{ii}\right)\le\prod_i a_{ii}$$

2）构造凸函数$f(x)=\frac{1}{2}x^H\mathbf{A}x$，则$H\left(f(x)\right)=\frac{1}{2}\left(\mathbf{A}+\mathbf{A}^H\right)=\mathbf{A}$

3）由$\mathbf{A}=\mathbf{AA}$，$\mathbf{A}^H=\mathbf{A}$

$$\Rightarrow x^H\mathbf{A}x=x^H\mathbf{AA}x=x^H\mathbf{A}^H\mathbf{A}x=(\mathbf{A}x)^H\cdot\mathbf{A}x\ge 0$$

$\Rightarrow\mathbf{A}$半正定

5.考虑上三角矩阵$\mathbf{A}$

1）如果$\mathbf{A}$ 是一个酉矩阵，则$\mathbf{A}$ 是一个对角矩阵，且对角元素的范数都是1

2）如果$\mathbf{A}$ 是一个上三角（块）酉矩阵，$\mathbf{A}=\begin{pmatrix}\mathbf{P} & \mathbf{B}\\ \mathbf{0} & \mathbf{Q}\end{pmatrix}$，其中$\mathbf{P}$ 为$m\times m$ 矩阵，$\mathbf{Q}$是 $n\times n$ 矩阵，那么$\mathbf{P}$ 和 $\mathbf{Q}$ 都是酉矩阵，且$\mathbf{B}=\mathbf{0}$

1）A为上三角矩阵，且为酉矩阵 $\Rightarrow \mathbf{A}^H = \mathbf{A}^{-1}, \mathbf{A}^H$为下三角矩阵，由于上三角矩阵的逆依旧是上三角矩阵 $\Rightarrow \mathbf{A}^H$也是上三角矩阵 $\Rightarrow \mathbf{A}^H$必须是对角阵。

令$\mathbf{A} = diag\{\lambda_1, \lambda_2, \ldots, \lambda_n\}, \mathbf{A}^H\mathbf{A} = \mathbf{I} = diag\left\{|\lambda_1|^2, |\lambda_2|^2, \ldots, |\lambda_n|^2\right\} \Rightarrow |\lambda_i| = 1$

2)A为上三角分块酉矩阵，由1）知 $\mathbf{B} = \mathbf{0}$,

且由$\mathbf{A}^H\mathbf{A} = \mathbf{I} \Rightarrow \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix}\begin{pmatrix} \mathbf{P}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^H \end{pmatrix} = \begin{pmatrix} \mathbf{PP}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{QQ}^H \end{pmatrix} = \mathbf{I}$

$\Rightarrow \mathbf{PP}^H = \mathbf{I}, \mathbf{QQ}^H = \mathbf{I}$

即P和Q均为酉矩阵

6. 设V_1与V_2分别为齐次线性方程组$x_1 + x_2 + \ldots + x_n = 0$和$x_1 = x_2 = \ldots = x_n$的解空间，证明$\mathbb{R}^n = V_1 \oplus V_2$

解：

方程组$x_1+x_2+\dots+x_n=0$的解空间是$n-1$维的，

$a_1=(-1,1,0,\ \dots\ ,0)^T$， $a_2=(-1,0,1,\ \dots\ ,0)^T,\dots,$

$a_{n-1}=(-1,0,0,\ \dots\ ,1)^T$,即$V_1=\mathrm{span}(a_1,a_2,\ \dots\ ,a_{n-1})$

方程组$x_1=x_2=\dots=x_n$的解空间是1维的，

$a=(1,1,1,\ \dots\ ,1)^T$, 即$V_2=\mathrm{span}(a)$

由于$a_1,a_2,\ \dots\ ,a_{n-1},a$线性无关， $V_1+V_2=\mathrm{span}(a_1,a_2,\ \dots\ ,a_{n-1},a)=\mathbb{R}^n$

又因$\dim V_1+\dim V_2=n-1+1=n$,根据维数定理：

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

因此$\dim(V_1\cap V_2)=\{\mathbf{0}\}$,命题成立

7. 设n阶方阵$\mathbf{A}=(a_{ij})_{n\times n}$，且$\sum_{j=1}^{n}\left|a_{ij}\right|<1, i=1,2,\ \ldots\ ,n$，

证明$\mathbf{A}$的每一个特征值的绝对值$|\lambda|<1$

解：

设$\mathbf{Ax} = \lambda \mathbf{x}$， $\mathbf{x}=(x_1, x_2, \ldots, x_n)^T$，且设$|x_k| = \max(|x_1|, |x_2|, \ldots, |x_n|)$

取$\mathbf{Ax} = \lambda \mathbf{x}$的第k个方程：

$$\lambda x_k = \sum_{j=1}^{n} a_{kj} x_j$$

于是

$$|\lambda||x_k| = \left|\sum_{j=1}^{n} a_{kj} x_j\right| \le \sum_{j=1}^{n} |a_{kj}||x_j|$$

即有

$$|\lambda| \le \sum_{j=1}^{n} |a_{kj}| \frac{|x_j|}{|x_k|} \le \sum_{j=1}^{n} |a_{kj}| < 1$$

8. 证明正规矩阵**A**若是一个上三角矩阵，

则必是对角矩阵

解：

对$\mathbf{A}$分块处理：$\mathbf{A}=\begin{bmatrix} a_{11} & \boldsymbol{\alpha}^T \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$，$\mathbf{B}$是n-1阶上三角矩阵；

$$\mathbf{A}^H\mathbf{A}=\begin{bmatrix} a_{11}^* & \mathbf{0}^T \\ \boldsymbol{\alpha}^* & \mathbf{B}^H \end{bmatrix}\begin{bmatrix} a_{11} & \boldsymbol{\alpha}^T \\ \mathbf{0} & \mathbf{B} \end{bmatrix}=\begin{bmatrix} |a_{11}|^2 & a_{11}^*\boldsymbol{\alpha}^T \\ a_{11}\boldsymbol{\alpha}^* & \boldsymbol{\alpha}^*\boldsymbol{\alpha}^T+\mathbf{B}^H\mathbf{B} \end{bmatrix}$$

$$\mathbf{A}\mathbf{A}^H=\begin{bmatrix} a_{11} & \boldsymbol{\alpha}^T \\ \mathbf{0} & \mathbf{B} \end{bmatrix}\begin{bmatrix} a_{11}^* & \mathbf{0}^T \\ \boldsymbol{\alpha}^* & \mathbf{B}^H \end{bmatrix}=\begin{bmatrix} |a_{11}|^2+\boldsymbol{\alpha}^T\boldsymbol{\alpha}^* & \boldsymbol{\alpha}^T\mathbf{B}^H \\ \mathbf{B}\boldsymbol{\alpha}^* & \mathbf{B}\mathbf{B}^H \end{bmatrix}$$

由于正规矩阵：$\mathbf{A}^H\mathbf{A}=\mathbf{A}\mathbf{A}^H$

则有：$\boldsymbol{\alpha}^T\boldsymbol{\alpha}^*=0$，因此$\boldsymbol{\alpha}=\mathbf{0}$，

同时，$\mathbf{B}^H\mathbf{B}=\mathbf{B}\mathbf{B}^H$，因此$\mathbf{B}$为$n-1$阶上三角正规矩阵；

上述过程证明了，$\mathbf{A}$为分块对角矩阵；

根据上述证明过程递推，可得$\mathbf{A}$为对角矩阵。

9. 求实值函数$f(\mathbf{X})=\mathbf{a}^T\mathbf{X}\mathbf{X}^T\mathbf{b}$的Jacobian矩阵，

其中$\mathbf{X}\in\mathbb{R}^{\mathrm{m}\times n},\mathbf{a},\mathbf{b}\in\mathbb{R}^{\mathrm{m}}$

解：

$$f(\mathbf{X}) = \sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{p=1}^{n} a_k x_{kp} x_{lp} b_l$$

$$\left[\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}^T}\right]_{ij} = \frac{\partial f(\mathbf{X})}{\partial x_{ji}} = \sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{p=1}^{n} \frac{\partial (a_k x_{kp} x_{lp} b_l)}{\partial x_{ji}} = \sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{p=1}^{n}\left[a_k x_{lp} b_l \frac{\partial x_{kp}}{\partial x_{ji}} + a_k x_{kp} b_l \frac{\partial x_{lp}}{\partial x_{ji}}\right]$$

$$= \sum_{l=1}^{m} a_j x_{li} b_l + \sum_{k=1}^{m} a_k x_{ki} b_j = \left[\mathbf{X}^T \mathbf{b}\right]_i a_j + \left[\mathbf{X}^T \mathbf{a}\right]_i b_j$$

所以，Jacobian矩阵和梯度矩阵分别为

$$D_{\mathbf{X}} f(\mathbf{X}) = \mathbf{X}^T(\mathbf{b}\mathbf{a}^T + \mathbf{a}\mathbf{b}^T),\quad \nabla_{\mathbf{X}} f(\mathbf{X}) = (\mathbf{a}\mathbf{b}^T + \mathbf{b}\mathbf{a}^T)\mathbf{X}$$

10. 证明：函数$f(\mathbf{x})=8x_1+12x_2+x_1^2-2x_2^2$仅有一个驻点，且此点既不是最大值点，也不是最小值点，而是一个鞍点

解：

$$\frac{\partial f}{\partial x_1}=8+2x_1 \qquad \frac{\partial f}{\partial x_2}=12-4x_2$$

$$\nabla f(\mathbf{x})=\begin{bmatrix} 8+2x_1 \\ 12-4x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$，仅有一个驻点 $\mathbf{x}^*=\begin{bmatrix} -4 \\ 3 \end{bmatrix}$

$\nabla^2 f(\mathbf{x})=\begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$是非正定的，（$\begin{vmatrix} 2 & 0 \\ 0 & -4 \end{vmatrix}=-8<0$）

所以$\mathbf{x}^*$不是最小值点，同理，考虑 $\min\left(-f(\mathbf{x})\right)$

$\nabla^2\left[-f(\mathbf{x})\right]$也是非正定的，故$\mathbf{x}^*$不是最大值点，

所以$\mathbf{x}^*$是鞍点

11. 对于Household矩阵 $\mathbf{H}=\mathbf{I}-2\mathbf{u}\mathbf{u}^H$

（1）证明：$\mathbf{u}$是$\mathbf{H}$的一个特征向量，并求出相对应的特征值；

（2）证明：$\mathbf{u}$与其他特征向量均正交；

（3）对$\mathbf{H}$进行特征值分解。

(1) $\mathbf{Hu}=\mathbf{u}-\mathbf{u}\mathbf{u}^H\mathbf{u}=\left(1-2\|\mathbf{u}\|^2\right)\mathbf{u}$

故$\mathbf{u}$是$\mathbf{H}$的特征向量，且对应的特征值为$1-2\|\mathbf{u}\|^2$

(2) $\because \mathbf{H}^H=\mathbf{H}$（Hermitian矩阵）

$\therefore \mathbf{H}$可酉对角化　$\therefore \mathbf{H}$的特征向量相互正交

(3) $\because \text{rank}（\mathbf{u}\mathbf{u}^H）=\text{rank}（\mathbf{u}）=1$，$\mathbf{u}\neq 0$

$\therefore \lambda_{\min}(\mathbf{H})=1-2\|\mathbf{u}\|^2$，其余$n-1$个特征值均为1

$$\therefore \text{对}\mathbf{H}\text{进行EVD：}\ \mathbf{H}=\begin{pmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_{n-1} & \mathbf{u}\end{pmatrix}\begin{pmatrix}1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 1-2\|\mathbf{u}\|^2\end{pmatrix}\begin{pmatrix}\mathbf{u}_1^H \\ \vdots \\ \mathbf{u}_{n-1}^H \\ \mathbf{u}^H\end{pmatrix}$$

12.证明：

（1）$N \times N$实对称矩阵$\mathbf{A}$能够表达成形式：$\mathbf{A} = \sum_{n=1}^{N} \lambda_n \mathbf{Q}_n$，

$\{\lambda_n\}_{n=1}^{N}$是$\mathbf{A}$的特征值；

（2）$\mathbf{Q}_n$是非负定矩阵；

（3）$\mathbf{Q}_i \mathbf{Q}_i = \mathbf{Q}_i$，$\mathbf{Q}_i \mathbf{Q}_j = \mathbf{0}$.

$\because$ $\mathbf{A}$是实对称矩阵

$\therefore$ 对$\mathbf{A}$进行EVD：$\mathbf{A}=\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T=\sum_{i=1}^{n}\lambda_i\mathbf{u}_i\mathbf{u}_i^T=\sum_{i=1}^{n}\lambda_i\mathbf{Q}_i,\mathbf{Q}_i=\mathbf{u}_i\mathbf{u}_i^T$

$\therefore \mathbf{u}_i^T\mathbf{u}_i=1,\mathbf{u}_i^T\mathbf{u}_j=0,\forall i\neq j,\mathbf{u}_i=0$

$\therefore \mathbf{Q}_i\mathbf{Q}_i=\mathbf{u}_i\mathbf{u}_i^T\mathbf{u}_i\mathbf{u}_i^T=\mathbf{u}_i\mathbf{u}_i^T=\mathbf{Q}_i,\mathbf{Q}_i\mathbf{Q}_j=\mathbf{u}_i\mathbf{u}_i^T\mathbf{u}_j\mathbf{u}_j^T=\mathbf{0}$

$\mathbf{x}^T\mathbf{Q}_i\mathbf{x}=\mathbf{x}^T\mathbf{u}_i\mathbf{u}_i^T\mathbf{x}=\left\|\mathbf{x}^T\mathbf{u}_i\right\|^2\geq 0,\forall\mathbf{x}\neq 0,\mathbf{u}_i\neq 0$,当且仅当$\mathbf{x}\perp\mathbf{u}_i$时，等号成立

$\therefore \mathbf{Q}_i$是非负定矩阵

14.对于幂等矩阵$\mathbf{A}$: $\mathbf{A}^2 = \mathbf{A}$,证明：

（1）$\mathbf{A}$的特征值为0或1;

（2）所有的幂等矩阵（除了单位阵）都是奇异矩阵；

（3）$\operatorname{rank}(\mathbf{A}) = \operatorname{tr}(\mathbf{A})$;

（4）$\forall k \geq 0, \mathbf{A}^k$与$\mathbf{A}$有相同的特征值和特征向量。

(1)设$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$

$\therefore \mathbf{A}^2\mathbf{u} = \lambda\mathbf{A}\mathbf{u} = \lambda^2\mathbf{u}$

$\because \mathbf{A}^2 = \mathbf{A}$

$\therefore \lambda^2 = \lambda$

$\therefore \lambda = 0$或1

(2)对$\mathbf{A}$进行EVD：$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H$

若$\mathbf{A}$的特征值全为1，则$\mathbf{\Sigma} = \mathbf{I}$，$\mathbf{A} = \mathbf{U}\mathbf{U}^H = \mathbf{I}$,此时$\mathbf{A}$为单位阵

若$\mathbf{A}$的特征值不全为1，则$\det(\mathbf{A}) = \prod_{i=1}^{n}\lambda_i = 0$,此时$\mathbf{A}$为奇异矩阵

证毕

(3) $\because \mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$, $\mathrm{rank}(\mathbf{A})$等于非0特征值个数

又 $\because \lambda_i = 0$或1

$\therefore \mathrm{rank}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i = \mathrm{tr}(\mathbf{A})$

(4)令$\forall i = 1, 2, \ldots, n, \mathbf{A}\mathbf{u}_i = \lambda \mathbf{u}_i$

$\because \mathbf{A}^2 = \mathbf{A}$

$\therefore \mathbf{A}^k = \mathbf{A}$

$\therefore \mathbf{A}^k \mathbf{u}_i = \lambda \mathbf{u}_i, \forall i = 1, 2, \ldots, n$

$\therefore \mathbf{A}^k$的特征值和特征向量与$\mathbf{A}$相同

16、（1）若$\mathbf{A}^3+2\mathbf{A}^2+\mathbf{A}-\mathbf{I}=0$，证明$\mathbf{A}$可逆，并求$\mathbf{A}^{-1}$；

（2）若$\mathbf{A}^2-\mathbf{A}-4\mathbf{I}=0$，证明$\mathbf{A}+\mathbf{I}$可逆，并求$\left(\mathbf{A}+\mathbf{I}\right)^{-1}$.

证明：

（1）由$\mathbf{A}^3+2\mathbf{A}^2+\mathbf{A}-\mathbf{I}=0 \Rightarrow \mathbf{A}^3+2\mathbf{A}^2+\mathbf{A}=\mathbf{I} \Rightarrow \mathbf{A}\left(\mathbf{A}^2+2\mathbf{A}+\mathbf{I}\right)=\mathbf{I}$

即存在矩阵$\mathbf{B}=\mathbf{A}^2+2\mathbf{A}+\mathbf{I}$，使得$\mathbf{AB}=\mathbf{I}$

$\therefore$矩阵$\mathbf{A}$可逆，其逆矩阵为$\mathbf{A}^{-1}=\mathbf{B}=\mathbf{A}^2+2\mathbf{A}+\mathbf{I}.$

（2）由$\mathbf{A}^2-\mathbf{A}-4\mathbf{I}=0 \Rightarrow (\mathbf{A}+\mathbf{I})(\mathbf{A}-2\mathbf{I})=2\mathbf{I} \Rightarrow (\mathbf{A}+\mathbf{I})\left[\frac{1}{2}(\mathbf{A}-2\mathbf{I})\right]=\mathbf{I}$

即存在矩阵$\mathbf{B}=\frac{1}{2}(\mathbf{A}-2\mathbf{I})$，使得$(\mathbf{A}+\mathbf{I})\mathbf{B}=\mathbf{I}$

$\therefore$矩阵$\mathbf{A}+\mathbf{I}$可逆，其逆矩阵为$(\mathbf{A}+\mathbf{I})^{-1}=\mathbf{B}=\frac{1}{2}(\mathbf{A}-2\mathbf{I}).$

17、若向量β可由向量组$\alpha_1,\alpha_2,\ldots,\alpha_r$线性表示，证明：向量$\beta$的表示法是唯一的充分必要条件是$\alpha_1,\alpha_2,\ldots,\alpha_r$线性无关.

证明：

（1）**必要性**.（向量β的表示法是唯一的$\Longrightarrow \alpha_1,\alpha_2,\ldots,\alpha_r$ 线性无关）

由于向量β可由向量组$\alpha_1,\alpha_2,\ldots,\alpha_r$线性表示，且表示法唯一，则存在唯一一组数$k_1,k_2,\ldots,k_r$，使得

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r \quad \text{......(式1)}$$

假设$\alpha_1,\alpha_2,\ldots,\alpha_r$线性相关，那么存在一组不全为零的数$l_1,l_2,\ldots,l_r$使得

$l_1\alpha_1 + l_2\alpha_2 + \ldots + l_r\alpha_r = 0$

不妨设$l_1 \neq 0$，则有

$$\alpha_1 = -\frac{l_2}{l_1}\alpha_2 - \ldots - \frac{l_r}{l_1}\alpha_r = c_2\alpha_2 + \ldots + c_r\alpha_r \quad \text{......(式2)}$$

将式2代入式1可得β的新的线性表示式，这与β的表示法唯一矛盾，

$\therefore \alpha_1,\alpha_2,\ldots,\alpha_r$线性无关.

（2）充分性.（$\alpha_1, \alpha_2, ..., \alpha_r$ 线性无关⟹向量β的表示法是唯一的）

已知向量β可由向量组$\alpha_1, \alpha_2, ..., \alpha_r$线性表示，且$\alpha_1, \alpha_2, ..., \alpha_r$ 线性无关

假设向量β的线性表示式不唯一，即存在两组不同的数$k_1, k_2, ..., k_r$和$l_1, l_2, ..., l_r$使得

$$\begin{cases} \beta = k_1\alpha_1 + k_2\alpha_2 + \dots + k_r\alpha_r \\ \beta = l_1\alpha_1 + l_2\alpha_2 + \dots + l_r\alpha_r \end{cases}$$

两式相减，得

$$(k_1 - l_1)\alpha_1 + (k_2 - l_2)\alpha_2 + \dots + (k_r - l_r)\alpha_r = 0$$

此时由于系数$k_i - l_i, i = 1,2, \dots r$不全为零，因此得$\alpha_1, \alpha_2, ..., \alpha_r$线性相关，矛盾

∴向量β的线性表示式唯一.

18、若$\mathbf{S}$、$\mathbf{T}$分别为实对称、反实对称矩阵，证明

$\mathbf{A}=\left(\mathbf{I}+\mathbf{T}+i\mathbf{S}\right)\left(\mathbf{I}-\mathbf{T}-i\mathbf{S}\right)^{-1}$为酉矩阵.

证明：

$$\mathbf{A}\mathbf{A}^{H}=(\mathbf{I}+\mathbf{T}+i\mathbf{S})(\mathbf{I}-\mathbf{T}-i\mathbf{S})^{-1}(\mathbf{I}-\mathbf{T}-i\mathbf{S})^{-H}(\mathbf{I}+\mathbf{T}+i\mathbf{S})^{H}$$

$$=(\mathbf{I}+\mathbf{T}+i\mathbf{S})(\mathbf{I}-\mathbf{T}-i\mathbf{S})^{-1}(\mathbf{I}+\mathbf{T}+i\mathbf{S})^{-1}(\mathbf{I}-\mathbf{T}-i\mathbf{S})$$

$$=(\mathbf{I}+\mathbf{T}+i\mathbf{S})\left[(\mathbf{I}+\mathbf{T}+i\mathbf{S})(\mathbf{I}-\mathbf{T}-i\mathbf{S})\right]^{-1}(\mathbf{I}-\mathbf{T}-i\mathbf{S})$$

由于$\left[\mathbf{I}+(\mathbf{T}+i\mathbf{S})\right]\left[\mathbf{I}-(\mathbf{T}+i\mathbf{S})\right]=\mathbf{I}-(\mathbf{T}+i\mathbf{S})^{2}=\left[\mathbf{I}-(\mathbf{T}+i\mathbf{S})\right]\left[\mathbf{I}+(\mathbf{T}+i\mathbf{S})\right]$

$$\therefore \mathbf{A}\mathbf{A}^{H}=(\mathbf{I}+\mathbf{T}+i\mathbf{S})\left[(\mathbf{I}-\mathbf{T}-i\mathbf{S})(\mathbf{I}+\mathbf{T}+i\mathbf{S})\right]^{-1}(\mathbf{I}-\mathbf{T}-i\mathbf{S})$$

$$=(\mathbf{I}+\mathbf{T}+i\mathbf{S})(\mathbf{I}+\mathbf{T}+i\mathbf{S})^{-1}(\mathbf{I}-\mathbf{T}-i\mathbf{S})^{-1}(\mathbf{I}-\mathbf{T}-i\mathbf{S})$$

$$=\mathbf{I}$$

同理可证$\mathbf{A}^{H}\mathbf{A}=\mathbf{I}$

$\therefore$ $\mathbf{A}$为酉矩阵.

19、证明：如果S_1和S_2是$\mathbb{R}^{m\times n}$中的凸集，那么它们的部分和

$$S=\left\{\left(x,y_1+y_2\right)\middle|x\in\mathbb{R}^m,y_1,y_2\in\mathbb{R}^n,\left(x,y_1\right)\in S_1,\left(x,y_2\right)\in S_2\right\}$$

也是凸的.

证明：

在S中任取两点$\left(\bar{x},\bar{y}_1+\bar{y}_2\right),\left(\tilde{x},\tilde{y}_1+\tilde{y}_2\right)$

其中$\left(\bar{x},\bar{y}_1\right),\left(\tilde{x},\tilde{y}_1\right)\in S_1,\left(\bar{x},\bar{y}_2\right),\left(\tilde{x},\tilde{y}_2\right)\in S_2$

对于$\forall 0\le\theta\le 1$，$\theta\left(\bar{x},\bar{y}_1+\bar{y}_2\right)+(1-\theta)\left(\tilde{x},\tilde{y}_1+\tilde{y}_2\right)$

$=\left(\theta\bar{x}+(1-\theta)\tilde{x},\left(\theta\bar{y}_1+(1-\theta)\tilde{y}_1\right)+\left(\theta\bar{y}_2+(1-\theta)\tilde{y}_2\right)\right)$

由于S_1和S_2都是凸集

$\therefore\left(\theta\bar{x}+(1-\theta)\tilde{x},\theta\bar{y}_1+(1-\theta)\tilde{y}_1\right)\in S_1,\left(\theta\bar{x}+(1-\theta)\tilde{x},\theta\bar{y}_2+(1-\theta)\tilde{y}_2\right)\in S_2$

$\therefore\theta\left(\bar{x},\bar{y}_1+\bar{y}_2\right)+(1-\theta)\left(\tilde{x},\tilde{y}_1+\tilde{y}_2\right)\in S$

$\therefore S$也是凸的.