目录

第一章 线性代数基础

1. 线性空间与子空间

定义 1：数域：
对四则运算封闭的数集
定义 2：基底：
线性空间V中n个无关的向量。任意$n+1$个向量都一定相关。n称为维数，$\dim V = n$
定义 3：子空间：
线性空间的子集，且也是线性空间

2. 空间分解与维数定理

定义 1：线性空间的和：
所有能由两个空间里的向量线性表出的向量构成的集合。$V_1 + V_2$
定理 1（维数定理）：
设V_1和V_2是 V的子空间，则
$\dim V_1 + \dim V_2 = \dim (V_1 + V_2) + \dim (V_1 \cap V_2)$
定义 2：直和：
$V_1 + V_2$中的向量只有唯一一种用$\alpha_1 \in V_1$和$\alpha_2 \in V_2$线性表示法。记为$V_1 \oplus V_2$
（注意，直和不代表V_1、V_2正交，只是说没有重叠）
定理 2（直和的等价命题）：
零向量表示唯一。即 $0 = \alpha_1 + \alpha_2$，必有$\alpha_1 = 0$，$\alpha_2 = 0$
$V_1 \cap V_2 = \{0\}$
推论 1：
$W = V_1 \oplus V_2 \Leftrightarrow \dim \mathrm{W} = \dim V_1 + \dim V_2$
定义 3：多个子空间的直和：
$V_1 \oplus V_2 \oplus \ldots \oplus V_s$
定理 3（多空间直和的等价命题）：
1. 零向量表示唯一
2. $V_i \cap \sum V_j = \{0\}$，$i \neq j$
3. $\dim W = \sum \dim V_i$

4. 线性流形与凸包

定义 1：线性流形：
子空间平移一个向量。直线、平面、超平面
定理 1（向量组合）：
α_i，$i = 0 \sim s$是R^n的任意$s+1$个向量，x是其凸组合

所有x构成的集合组成一个线性流形P，P的维数等于向量组$\alpha_i-\alpha_0$，$i=1\sim s$的秩

定理 2（线性流形相等）：

两个线性流形相等的充要条件是原子空间相等，且平移向量之差属于其子空间

定理 3：

两条直线包含在一个三维线性流形中

定理 4：

两条直线$x=\alpha_0+\alpha_1 t$、$x=\beta_0+\beta_1 t$位于一个平面，充要条件是$\alpha_0-\beta_0$，α_1，β_1线性相关

推论 1：

两条直线$x=\alpha_0+\alpha_1 t$、$x=\beta_0+\beta_1 t$相交而不重合，充要条件是α_1，β_1线性无关，$\alpha_0-\beta_0$可用α_1，β_1线性表出

定理 5：

维数分别为k和h的两个流形包含在一个维数不大于$k+h+1$的流形中

定理 6：

维数分别为k和h的两个流形P、Q有公共向量，则$\dim P\cap Q\geq k+h-n$

定义 2：凸集：

集合内两点为线段中的点都属于该集合

定理 7：

凸集的交集是凸集

5. 特征值与特征向量

一、特征值和特征向量的概念

定义 1：特征值与特征向量：

A是$n\times n$复矩阵，若$Ax=\lambda x$，则λ是特征值，x是从属于λ的特征向量

特征多项式：$f(\lambda)=|\lambda E-A|=0$

若$f(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\dots(\lambda-\lambda_r)^{n_r}$，称$n_i$是$\lambda_i$的代数重数

若$\operatorname{rank}(\lambda_i E-A)=n-m_i$，则称$m_i$是$\lambda_i$的几何重数

代数重数大于等于几何重数

定理 1（Jordan 标准形）：

$P^{-1}AP=J=\begin{pmatrix}J_1(\lambda_1)&\cdots&0\\ \vdots&\ddots&\vdots\\ 0&\cdots&J_r(\lambda_r)\end{pmatrix}$，$J$为 Jordan 标准形，Jordan 块$J_i(\lambda_i)=\begin{pmatrix}\lambda_i&1&\cdots&0&0\\ 0&\lambda_i&\ddots&0&0\\ \vdots&\vdots&\ddots&\ddots&\vdots\\ 0&0&\cdots&\lambda_i&1\\ 0&0&\cdots&0&\lambda_i\end{pmatrix}$

求 Jordan 标准型：

1. 求特征值$\lambda_1,\dots,\lambda_s$。$|\lambda E-A|=0$
2. 求J。设λ_i的代数重数为r，若$\operatorname{rank}(\lambda_i E-A)=n-r$，即$\lambda_i$几何重数为$r$，则不用求 Jordan 标准型，可相似对角化。

若$\operatorname{rank}(\lambda_i E-A)=n-r_c>n-r$，即$\lambda_i$的几何重数为$r_c$（$r_c<r$），则$\lambda_i$对应$r_c$个 Jordan 块。

如有必要，算λ_i的 1 阶 Jordan 块个数=$\operatorname{rank}(\lambda_i E-A)^2+\mathrm{n}-2\operatorname{rank}(\lambda_i E-A)$，从而推知$\lambda_i$的 Jordan 块阶数组成。

3. 求$P^{-1}AP=J$的P。$P=(X_1,X_2,\dots,X_n)$，列方程组$AP=PJ$，求出P。

例：

求$A=\begin{pmatrix}1&2&3&4\\ &1&2&3\\ &&1&2\\ &&&1\end{pmatrix}$的 Jordan 标准型$J$，并求$P$使得$P^{-1}AP=J$。

$|\lambda E-A|=(\lambda-1)^4=0$，$\lambda=1$的代数重数为 4。

$\operatorname{rank}(E-A)=\operatorname{rank}\begin{pmatrix}0&-2&-3&-4\\ &0&-2&-3\\ &&0&-2\\ &&&0\end{pmatrix}=3>4-4=0$，几何重数为$4-3=1$，有 1 个 Jordan 标准形，阶数为 4。

故$J=\begin{pmatrix}1&1&0&0\\ 0&1&1&0\\ 0&0&1&1\\ 0&0&0&1\end{pmatrix}$。

$AP=PJ$，即$A(X_1,X_2,X_3,X_4)=(X_1,X_2,X_3,X_4)J$，即$(AX_1,AX_2,AX_3,AX_4)=(X_1,X_1+X_2,X_2+X_3,X_3+X_4)$

即$(A-E)X_1=0$，$(A-E)X_2=X_1$，$(A-E)X_3=X_2$，$(A-E)X_4=X_3$，即$(A-E)^4X_4=0$，$(A-E)^3X_4=X_3\neq 0$。

$(A-E)^2=\begin{pmatrix}0&0&4&12\\0&0&0&4\\0&0&0&0\\0&0&0&0\end{pmatrix}$，$(A-E)^3=\begin{pmatrix}0&0&0&8\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$，$(A-E)^4=O$，$(A-E)^3X_4\neq 0$，可取$X_4=(1,1,1,1)^T$。

则$X_3=(A-E)X_4=(9,5,2,0)^T$，$X_2=(A-E)X_3=(16,4,0,0)^T$，$X_1=(A-E)X_2=(8,0,0,0)^T$

则$P=(X_1,X_2,X_3,X_4)=\begin{pmatrix}8&16&9&1\\0&4&5&1\\0&0&2&1\\0&0&0&1\end{pmatrix}$

定义 2：可对角化矩阵：

$$P^{-1}AP=\Lambda=\begin{pmatrix}\lambda_1&\cdots&0\\\vdots&\ddots&\vdots\\0&\cdots&\lambda_n\end{pmatrix}$$

定理 2（可对角化矩阵的等价命题）：

1. A 的特征向量线性无关，构成C^n的一组基底
2. Jordan 块都是一阶的
3. 所有特征值的代数重数等于几何重数

A 可逆等价于 A 不存在零特征值

A 的特征值之和等于 A 的迹

二、特征值和特征向量的几何性质

定义 1'：特征值与特征向量的几何定义：

T 是$V_n(C)$的一个线性变换，若$T\xi=\lambda\xi$，则λ是特征值，ξ是从属于λ的特征向量

定理 3：

不同特征值对应的特征向量线性无关

证明：令$\sum c_i\xi_i=0$，$i=1\sim s$，给每个式子分别左乘T，T^2，...，T^{s-1}，并用$T\xi=\lambda\xi$代换

写成矩阵有$(c_1\xi_1,c_2\xi_2,...,c_n\xi_n)\begin{pmatrix}1&\lambda_1&\cdots&\lambda_1^{s-1}\\1&\lambda_2&&\lambda_2^{s-1}\\\vdots&&\ddots&\vdots\\1&\lambda_n&\cdots&\lambda_n^{s-1}\end{pmatrix}=0$，矩阵为范德蒙德行列式不等于 0，$\xi_i\neq 0$，故只有$c_i=0$

定理 4：

几何重数（λ_i对应的特征向量所组成的子空间V_{λ_i}的维度）不大于其代数重数

定理 5（可对角化的充要条件）：

$C^n=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\ldots\oplus V_{\lambda_n}$

6. 初等矩阵及酉变换

一、初等矩阵的一般形式

定义 1：初等矩阵：

$E(u,v;\sigma)=E_n-\sigma uv^H$（单位阵–秩一阵）

性质 1：

若垂直，则$E(u,v;\sigma)$仅有$n-1$ 个线性无关的特征向量（$\lambda=1$ 代数重数为 n，几何重数为$n-1$，不可相似对角化）

若 u，v 不垂直，则$E(u,v;\sigma)$有 n 个线性无关的特征向量（$\lambda=1$ 重数为$n-1$，$\lambda=1-\sigma v^Hu$重数为 1，可相似对角化）

在$v^{\perp}$中的一组基就是$E(u,v;\sigma)$属于$\lambda=1$ 的特征向量

性质 2：

$E(u,v;\sigma)$的特征谱为 $1-\sigma v^Hu,1,1,\ldots,1$

性质 3：

$|E(u,v;\sigma)|=1-\sigma v^Hu$

性质 4：

当且仅当$\sigma v^Hu\neq 1$ 时可逆，且$E(u,v;\sigma)^{-1}=E(u,v;\frac{\sigma}{\sigma v^Hu-1})$

性质 5：

要使$E(u,v;\sigma)a=b$，只需满足$v^Ha\neq 0$，$\sigma u=\frac{a-b}{v^Ha}$

所有初等变换矩阵都能表示成$E(u,v;\sigma)$

1. 交换阵$E_{ij}=E(\varepsilon_i-\varepsilon_j,\varepsilon_i-\varepsilon_j;1)$
2. 倍加阵$E_{ij}(k)=E(\varepsilon_j,\varepsilon_i;-k)$
3. 倍乘阵$E_i(k)=E(\varepsilon_i,\varepsilon_i;1-k)$

二、初等下三角阵

定义 2：初等下三角阵：

对角元为 1，只有第 i 列第 i+1~n 行不为 0，其余均为 0（左乘 A，用于消去 A 的第 i 列）

任一可逆下三角阵可以表示成一个可逆对角阵和若干个初等下三角阵的乘积

三、初等酉阵

定义 3：初等酉阵（Householder 变换）：

$H(u)=E(u,u;2)=E_n-2uu^H$，其中$u^Hu=1$（长度为 1）

性质 1：

$H(u)=H(u)^H=H(u)^{-1}$

性质 2：

$|H(u)|=-1$，$H(u)$与$\begin{pmatrix}-1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{pmatrix}$酉相似

性质 3：

$H(u)$是镜象变换。$H(u)(a+ru)=a-ru$

性质 4：

$H(u)$是酉变换，保内积、长度、夹角、“形状”不变

四、酉变换与酉矩阵

定义 4：酉变换：

n维复空间的变换T满足$(T(x),T(y))=(x,y)$（保内积不变）

性质 1：

酉变换是线性变换

定理 1（酉变换的等价命题）：

$\|T(x)\|=\|x\|$（保长度不变）

一组标准正交基经过变换T后仍是标准正交基（保形状不变）

T在标准正交基下的矩阵A是酉矩阵

定理 2（酉矩阵的性质）：

保内积不变，保长度不变

A^H也是酉矩阵

酉矩阵相乘也是酉矩阵

酉矩阵的特征值的模为 1

7. 欧式空间上的度量

定义 1：内积：

对向量x，y的运算(x,y)满足

1. 非负性：$(x,x)\geq 0$，仅当$x=0$时有$(x,x)=0$
2. 对称性：$(x,y)=\overline{(y,x)}$
3. 线性性：$(\lambda x,y)=\bar{\lambda}(x,y)$，$(x+y,z)=(x,z)+(y,z)$

定义 2：长度（范数）：

$||x||=\sqrt{(x,x)}$

定理 1（内积的性质）：

齐次性：$||\lambda x|| = |\lambda|||x||$

平行四边形法则：$||x-y||^2 + ||x+y||^2 = 2||x||^2 + 2||y||^2$

柯西不等式：$|(x,y)| \leq \|x\|\|y\|$

三角不等式：$\|x+y\| \leq \|x\| + \|y\|$

定义 3：距离：

$d(x,y) = \|x-y\|$

性质 123：

对称性，非负性，三角不等式

定义 4：正交（垂直）：

$(\alpha,\beta) = 0$

性质 1：

勾股定理

定理 2（垂线最短）：

一个固定向量和一个子空间中各向量间的距离中，垂线最短

定理 3（垂线长度）：

向量α给出的点到线性流形$P = \alpha_0 + V_1$的距离等于向量$\alpha - \alpha_0$关于子空间V_1的正交分量β的长度

距离的求解：

①设线性流形$P = \alpha_0 + V_1$，其中V_1为导出组的解空间，α_0是非齐次方程组的一个特解。根据给出的线性方程组，求出其特解和解空间，特解为α_0，解空间显示V_1的基向量。

②求出${V_1}^{\perp}$的基向量（也即是把V_1当做齐次方程组，求解空间），相当于把空间进行直和分解为$V_1 \oplus {V_1}^{\perp}$。

③将$\alpha - \alpha_0$用V_1和${V_1}^{\perp}$的基向量进行线性表出，β即为仅用${V_1}^{\perp}$的基向量进行表出的分量

④求解β的长度即可

定义 5：Gram 行列式：

$$G(\alpha_1,\alpha_2,\dots,\alpha_k) = \begin{vmatrix} (\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_k) \\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_k) \\ \vdots & & \ddots & \vdots \\ (\alpha_k,\alpha_1) & (\alpha_k,\alpha_2) & \cdots & (\alpha_k,\alpha_k) \end{vmatrix}$$（可用于计算体积）

施密特正交化：

$\beta_1 = \alpha_1$

$\beta_2 = \alpha_2 - \frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1$

$\beta_3 = \alpha_3 - \frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1 - \frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2$

定理 4：

$\alpha_1,\alpha_2,\dots,\alpha_k$线性相关的充要条件是$G(\alpha_1,\alpha_2,\dots,\alpha_k) = 0$

定理 5：

正交化（不能单位化）不改变 Gram 行列式。

$G(\alpha_1,\alpha_2,\dots,\alpha_k) = G(\beta_1,\beta_2,\dots,\beta_k) = ||\beta_1||^2||\beta_2||^2 \dots ||\beta_k||^2$ （等于体积的平方）

推论 1：

$G(\alpha_1,\alpha_2)$表示平行四边形面积平方，$G(\alpha_1,\alpha_2,\alpha_3)$表示平行六面体体积平方

定理 6：

$\alpha_1,\alpha_2,\dots,\alpha_k$是线性子空间$V_1$的基底，则$\alpha$到线性流形$P = \alpha_0 + V_1$的距离$d$有$d^2 = \frac{G(\alpha_1,\alpha_2,\dots,\alpha_k,\alpha-\alpha_0)}{G(\alpha_1,\alpha_2,\dots,\alpha_k)}$

定义 6：平行多面体的体积：

$V(\alpha_1) = ||\alpha_1||$

$V(\alpha_1,\alpha_2,\dots,\alpha_n) = V(\alpha_1,\alpha_2,\dots,\alpha_{n-1})h_n$

其中h_n是向量α_n关于向量$\alpha_1,\alpha_2,\dots,\alpha_{n-1}$张成的子空间的正交分量长度

定理 7：

$V(\alpha_1,\alpha_2,...,\alpha_n)=\sqrt{G(\alpha_1,\alpha_2,...,\alpha_k)}=|D|$

其中D是$\alpha_1,\alpha_2,...,\alpha_n$在标准正交基下的坐标构成的行列式

定理 8（正交多面体最大）：

$G(\alpha_1,\alpha_2,...,\alpha_k)\le||\alpha_1||^2||\alpha_2||^2...||\alpha_k||^2$，当且仅当$\alpha_1,\alpha_2,...,\alpha_k$两两正交或有至少一个 0 向量时等号成立

定理 9：

在α与V_1中所有向量的夹角中，以α在V_1上的正交投影 r 之间的夹角最小

定理 10：

线性流形$P_1=\alpha_1+V_1$和$P_2=\alpha_2+V_2$之间的距离等于向量$\alpha_1-\alpha_2$关于线性子空间V_1+V_2的正交分量的长度

8. 酉空间的分解与投影

二、正交补子空间

定义 2：正交补子空间

两个子空间$V_1\perp V_2$，且$V_1+V_2=V_n(C)$。

定理 3：

酉空间每个子空间都有唯一的正交补

三、投影与幂等矩阵

定义 3：投影（投影算子、幂等算子）：

线性变换T满足$T^2=T$

定理 4：

投影的值域和核互为直和补。$V_n(C)=R(T)\oplus N(T)$

（值域$R(T)$：像β的全体。$R(T)=\{\beta|\beta=T\alpha,\alpha\in V_n(C)\}$，$\alpha$称为原像，$T$称为投影）

（核$N(T)$：零子空间。$N(T)=\{\alpha|T\alpha=0,\alpha\in V_n(C)\}$）

定理 5：

设$V_n(C)=V_1\oplus V_2$，则存在投影T，使$R(T)=V_1$，$N(T)=V_2$。

定理 6（幂等矩阵的性质）：

1. A^H和$E-A$也是幂等矩阵
2. A的特征值非 0 即 1，且可对角化
3. $\mathrm{rank}(A)=\mathrm{tr}(A)$
4. $A(E-A)=(E-A)A=0$
5. $A\alpha=\alpha\Leftrightarrow\alpha\in R(T)$（值域内的向量投影等于自身）
6. $N(A)=R(E-A)$，$R(A)=N(E-A)$

四、正交投影

定义 4：正交投影

设 T 为$V_n(C)$上的投影，$V_n(C)=R(T)\oplus N(T)$。如果$R(T)^\perp=N(T)$，则称T是正交投影

（值核正交的投影）

定理 7（正交投影的充要条件）：

$A=A^H$（Hermite 矩阵）

$C^n=R(A)\oplus N(A)$，$R(A)^\perp=N(A)$

第二章 向量与矩阵的范数

1. 向量的范数

定义 1：范数：

正定性：$\|x\| \geq 0$，当且仅当$x = 0$时有$\|x\| = 0$

齐次性：$\|\lambda x\| = |\lambda|\|x\|$

三角不等式：$\|x + y\| \leq \|x\| + \|y\|$

范数不是长度，属于广义的“长度”

p范数（Holder 范数）：$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$（$p \geq 1$）

1 范数（模的和）：$\|x\|_1 = \sum_{i=1}^{n} |x_i|$

2 范数（长度）：$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}$

∞范数（最大模）：$\|x\|_\infty = \max\limits_{1\leq i\leq n} |x_i|$

（$0 < p < 1$时，p范数不是范数）

引理 1：

u、v是非负实数，p、q是正实数，有$p > 1$和$\frac{1}{p} + \frac{1}{q} = 1$，则$uv \leq \frac{1}{p}u^p + \frac{1}{q}v^q$

定理 1（Holder 不等式）：

若$p, q > 1$且$\frac{1}{p} + \frac{1}{q} = 1$，则$\sum_{i=1}^{n} |x_i||y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}}$

证明：令$u = \frac{|x_i|}{\|x\|_p}$，$v = \frac{|y_i|}{\|y\|_p}$，根据引理 1，有$\frac{|x_i||y_i|}{\|x\|_p\|y\|_p} \leq \frac{|x_i|^p}{p\|x\|_p^p} + \frac{|y_i|^q}{q\|y\|_q^q}$

则$\sum_{i=1}^{n} \frac{|x_i||y_i|}{\|x\|_p\|y\|_p} \leq \sum_{i=1}^{n} \left(\frac{|x_i|^p}{p\|x\|_p^p} + \frac{|y_i|^q}{q\|y\|_q^q}\right) = \frac{\sum_{i=1}^{n} |x_i|^p}{p\|x\|_p^p} + \frac{\sum_{i=1}^{n} |y_i|^q}{q\|y\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1$（$\sum_{i=1}^{n} |x_i|^p = \|x\|_p^p$）

故$\sum_{i=1}^{n} |x_i||y_i| \leq \|x\|_p\|y\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}}$

特例：$p = q = 2$，Holder 不等式特化为柯西不等式

定理 2：

如果$\|\cdot\|$是范数（n维），那么$\|A\cdot\|$也是范数（$A_{m\times n}$，m维）

A-范数：$\|x\| = \sqrt{x^H A x}$，其中A为 Hermite 正定矩阵

定义 2：范数等价：

对于任意x，都有$\begin{cases} \|x\|_a \leq C_1\|x\|_b \\ \|x\|_b \leq C_2\|x\|_a \end{cases}$，则$\|\cdot\|_a$和$\|\cdot\|_b$范数等价。

定理 3：

$V_n(P)$（有限维空间）上的任意两个向量范数均等价

定理 4（向量序列的极限的两种定义等价）：

$\lim\limits_{k\to\infty} x^{(k)} = a \Leftrightarrow \lim\limits_{k\to\infty} \left\|x^{(k)} - a\right\| = 0$（任意范数。因为范数都等价）

逐点收敛：$\lim\limits_{k\to\infty} x_i^{(k)} = a_i$（$i = 1,2,\ldots,n$）或$\lim\limits_{k\to\infty} x^{(k)} = a$

依范数收敛：$\lim\limits_{k\to\infty} \left\|x^{(k)} - a\right\| = 0$

2. 矩阵的范数

定义 1：矩阵范数：

正定性：$\|A\| \geq 0$，当且仅当$A = 0$时有$\|A\| = 0$

齐次性：$\|\lambda A\| = |\lambda|\|A\|$

三角不等式：$\|A+B\| \leq \|A\| + \|B\|$

矩阵 p 范数：$\|A\|_{mp} = \left(\sum_{j=1}^{n}\sum_{i=1}^{m}\left|a_{ij}\right|^{p}\right)^{\frac{1}{p}}$（$p \geq 1$）

矩阵 1 范数（模的和）：$\|A\|_{m1} = \sum_{j=1}^{n}\sum_{i=1}^{m}\left|a_{ij}\right|$

矩阵 2 范数（“长度”）：$\|A\|_{m2} = \left(\sum_{j=1}^{n}\sum_{i=1}^{m}\left|a_{ij}\right|^{2}\right)^{\frac{1}{2}}$

矩阵∞范数（最大模）：$\|A\|_{m\infty} = \max\limits_{i,j}\left\{\left|a_{ij}\right|\right\}$

定理 1：
$P^{m\times n}$上的任意两个矩阵范数等价
定义 2：范数相容：
相容：$\|AB\|_c \leq \|A\|_a\|B\|_b$。（自）相容：$\|AB\| \leq \|A\|\|B\|$。

$\|\cdot\|_{m1}$、$\|\cdot\|_{m2}$相容，$\|\cdot\|_{m\infty}$不相容；矩阵a范数$\|A\|_a = n\max\limits_{i,j}\left\{\left|a_{ij}\right|\right\}$是相容的，$A$是$n$阶方阵

定理 2（$\|\cdot\|_{m2}$，Frobenius 范数$\|\cdot\|_F$的性质）：
（A是n阶方阵）
按列分块$A = (\alpha_1, \alpha_2, \ldots, \alpha_n)$，$\|A\|_F^2 = \|A\|_{m2}^2 = \sum_{i=1}^{n}\|\alpha_i\|_2^2$
$\|A\|_{m2}^2 = \operatorname{tr}(A^HA) = \sum_{i=1}^{n}\lambda_i(A^HA)$ （A^HA的迹、A^HA的特征值的和）

$\|A\|_{m2} = \left\|U^HAV\right\|_{m2} = \left\|UAV^H\right\|_{m2}$（酉不变）

推论 1：
$\|UAV\|_{m2} = \|A\|_{m2}$

3. 算子范数

一、算子范数

定义 1：与向量范数相容的矩阵范数：
若任意均有$\|Ax\|_a \leq \|A\|_m\|x\|_a$，则$\|\cdot\|_m$是与向量范数$\|\cdot\|_a$相容的矩阵范数
定理 1（从属于向量范数的算子范数）：

$\|A\|_a = \max\limits_{x\neq\theta}\frac{\|Ax\|_a}{\|x\|_a}$（$= \max\limits_{\|u\|_a=1}\|Au\|_a$）是与向量范数$\|x\|_a$相容的矩阵范数

证：

相容：$\|A\|_a = \max\limits_{x\neq\theta}\frac{\|Ax\|_a}{\|x\|_a}$，则$\|A\|_a \geq \frac{\|Ax\|_a}{\|x\|_a}$

非负性：设$A \neq 0$，则存在$x_0 \neq \theta$使$Ax_0 \neq 0$，那么$\|A\|_a \geq \frac{\|Ax_0\|_a}{\|x_0\|_a} > 0$

齐次性：$\|\lambda A\|_a = \max\limits_{x\neq\theta}\frac{\|\lambda Ax\|_a}{\|x\|_a} = \max\limits_{x\neq\theta}\frac{|\lambda|\|Ax\|_a}{\|x\|_a} = |\lambda|\max\limits_{x\neq\theta}\frac{\|Ax\|_a}{\|x\|_a} = |\lambda|\|A\|_a$

三角不等式：$\|A+B\|_a = \max\limits_{x\neq\theta}\frac{\|(A+B)x\|_a}{\|x\|_a} = \max\limits_{x\neq\theta}\frac{\|Ax+Bx\|_a}{\|x\|_a} \leq \max\limits_{x\neq\theta}\frac{\|Ax\|_a+\|Bx\|_a}{\|x\|_a}$

（都取最大值再加起来≥加起来再取总的最大值）$\leq \max\limits_{x\neq\theta}\frac{\|Ax\|_a}{\|x\|_a} + \max\limits_{x\neq\theta}\frac{\|Bx\|_a}{\|x\|_a} = \|A\|_a + \|B\|_a$

推论 1：
算子范数自相容
定理 2：

相容的矩阵范数$\|\cdot\|_m$都存在向量范数使其相容（取$\|x\| = \left\|xa^H\right\|_m$，$a$为任意非 0 列向量）

定理 3：
相容的矩阵范数$\|A\|_m \geq |\lambda_i|$，其中λ_i是A的特征值

二、算子范数的计算

算子 1 范数$\|A\|_1 \overset{\text{def}}{=} \max\limits_{x\neq\theta}\frac{\|Ax\|_1}{\|x\|_1} = \max\limits_{j}\sum_{i=1}^{n}|a_{ij}|$**（列和范数）**

证：

（1. 证明$\|A\|_1 \le \max\limits_{j}\sum_{i=1}^{n}|a_{ij}|$）

$\|Ax\|_1 = \sum_{i=1}^{n}\left|\sum_{j=1}^{n}a_{ij}x_j\right| \le \sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}||x_j|$（和的模≤模的和，积的模=模的积）

$= \sum_{j=1}^{n}\left(\sum_{i=1}^{n}|a_{ij}|\right)|x_j|$（交换求和次序、提公因式）$\le \left(\max\limits_{j}\sum_{i=1}^{n}|a_{ij}|\right)\sum_{j=1}^{n}|x_j|$（$\sum_{i=1}^{n}|a_{ij}|$放大为$\max\limits_{j}\sum_{i=1}^{n}|a_{ij}|$）

$= \left(\max\limits_{j}\sum_{i=1}^{n}|a_{ij}|\right)\|x\|_1$

即$\frac{\|Ax\|_1}{\|x\|_1} \le \max\limits_{j}\sum_{i=1}^{n}|a_{ij}|$

（2. 证明$\|A\|_1 \ge \max\limits_{j}\sum_{i=1}^{n}|a_{ij}|$）

设第s列有最大列和（$1 \le s \le n$），令$x_0 = \varepsilon_s = (0,\ldots,0,1,0,\ldots,0)^T$，有$\|x_0\|_1 = 1$，

按列分块$A = (\alpha_1,\ldots,\alpha_s,\ldots,\alpha_n)$，则$Ax_0 = A\varepsilon_s = \alpha_s$，则$\|Ax_0\|_1 = \|\alpha_s\|_1 = \sum_{i=1}^{n}|a_{is}| = \max\limits_{j}\sum_{i=1}^{n}|a_{ij}|$

故$\max\limits_{x\neq\theta}\frac{\|Ax\|_1}{\|x\|_1} \ge \frac{\|Ax_0\|_1}{\|x_0\|_1} = \max\limits_{j}\sum_{i=1}^{n}|a_{ij}|$

综合（1）（2），有$\|A\|_1 = \max\limits_{j}\sum_{i=1}^{n}|a_{ij}|$，证毕。

算子∞范数$\|A\|_\infty \overset{\text{def}}{=} \max\limits_{x\neq\theta}\frac{\|Ax\|_\infty}{\|x\|_\infty} = \max\limits_{i}\sum_{j=1}^{n}|a_{ij}|$**（行和范数）**

证：

（1. 证明$\|A\|_\infty \le \max\limits_{i}\sum_{j=1}^{n}|a_{ij}|$）

$\|Ax\|_\infty = \max\limits_{i}\left\{\left|\sum_{j=1}^{n}a_{ij}x_j\right|\right\} \le \max\limits_{i}\left\{\sum_{j=1}^{n}|a_{ij}||x_j|\right\}$（和的模≤模的和，积的模=模的积）

$\le \max\limits_{i}\left\{\sum_{j=1}^{n}|a_{ij}|\|x\|_\infty\right\}$（$|x_j|$放大为$\|x\|_\infty = \max\limits_{1\le j\le n}|x_j|$）$= \max\limits_{i}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}\cdot\|x\|_\infty$（提公因式）

即$\frac{\|Ax\|_\infty}{\|x\|_\infty} \le \max\limits_{i}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}$

（2. 证明$\|A\|_\infty \ge \max\limits_{i}\sum_{j=1}^{n}|a_{ij}|$）

设第s行有最大行和（$1 \le s \le n$），$\alpha_s = (a_{sj}) = \left(|a_{sj}|e^{i\theta_j}\right)$（用指数形式表示复数）

令$x_0 = \left(e^{-i\theta_1}, e^{-i\theta_2},\ldots,e^{-i\theta_n}\right)^T$，有$\|x_0\|_\infty = 1$，

按列分块$A = (\alpha_1,\alpha_2,\ldots,\alpha_n)$，则$Ax_0 = \alpha_1e^{-i\theta_1} + \alpha_2e^{-i\theta_2} + \ldots + \alpha_ne^{-i\theta_n}$

则$\|Ax_0\|_\infty = \max\limits_{1\le i\le n}\left|\alpha_{i1}e^{-i\theta_1} + \alpha_{i2}e^{-i\theta_2} + \ldots + \alpha_{in}e^{-i\theta_n}\right| \ge \left|\alpha_{s1}e^{-i\theta_1} + \alpha_{s2}e^{-i\theta_2} + \ldots + \alpha_{sn}e^{-i\theta_n}\right|$（第$s$个分量）

$= |a_{s1}| + |a_{s2}| + \ldots + |a_{sn}| = \max\limits_{i}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}$（第$s$行有最大行和）

故$\max\limits_{x\neq\theta}\frac{\|Ax\|_\infty}{\|x\|_{\infty 1}} \ge \frac{\|Ax_0\|_\infty}{\|x_0\|_\infty} = \max\limits_{i}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}$

综合（1）（2），有$\|A\|_\infty = \max\limits_{i}\sum_{j=1}^{n}|a_{ij}|$，证毕。

定义 2：谱半径：

$r(A)=\max_i |\lambda_i|$（包含所有λ_i的最小圆半径）

算子 2 范数$\|A\|_2 \stackrel{\text{def}}{=} \max_{x\neq\theta}\frac{\|Ax\|_2}{\|x\|_2}=\sqrt{r(A^HA)}$**（谱范数）**

证：

（1. 证明$\|A\|_2^2 \le r(A^HA)$）

$\|Ax\|_2^2=(Ax,Ax)=(Ax)^HAx=x^HA^HAx\ge 0$（长度≥ 0），故$A^HA$是半正定 Hermite 矩阵

又A^HA是正规阵（$A^HA(A^HA)^H=(A^HA)^HA^HA$）（第三章第 2 节定义 3）

故$A^HA=U\,\mathrm{diag}\,(\lambda_1,\lambda_2,\dots,\lambda_n)\,U^H=U\Lambda U^H$（第三章第 2 节定理 5），其中$U$是酉矩阵

不妨设$\lambda_1\ge\lambda_2\ge\dots\ge\lambda_n\ge 0$（半正定）

$x_1,x_2,\dots,x_n$是对应于$\lambda_1,\lambda_2,\dots,\lambda_n$的单位正交的特征向量（$U=(x_1,x_2,\dots,x_n)$）

故$\|Ax\|_2^2=x^HA^HAx=x^HU\Lambda U^Hx=y^H\Lambda y$（用$y=U^Hx$化二次型为标准型）

$=\lambda_1|y_1|^2+\lambda_2|y_2|^2+\dots+\lambda_n|y_n|^2\le\lambda_1(|y_1|^2+|y_2|^2+\dots+|y_n|^2)=\lambda_1\|y\|_2^2=\lambda_1\|x\|_2^2$（酉不变）

故$\frac{\|Ax\|_2^2}{\|x\|_2^2}\le\lambda_1=r(A^HA)$

（2. 证明$\|A\|_2^2\ge r(A^HA)$）

令$x_0=x_1$（对应于λ_1的单位特征向量），则$A^HAx_1=\lambda_1x_1$

故$\|Ax_1\|_2^2=x_1^HA^HAx_1=\lambda_1x_1^Hx_1=\lambda_1=r(A^HA)$

故$\max_{x\neq\theta}\frac{\|Ax\|_2^2}{\|x\|_2^2}\ge\frac{\|Ax_1\|_2^2}{\|x_1\|_2^2}=r(A^HA)$

综合（1）（2），有$\|A\|_2^2=r(A^HA)$，即$\|A\|_2=\sqrt{r(A^HA)}$，证毕。

三、谱范数

定理 4：

$\|A\|_2=\|A^H\|_2=\|A^T\|_2=\|\bar{A}\|_2$（证明$A^HA$和$AA^H$特征值相同，进而证明$\|A\|_2^2=r(A^HA)=r(AA^H)=\|A^H\|_2^2$）

$\|A^HA\|_2=\|AA^H\|_2=\|A\|_2^2$（$A\alpha=\lambda\alpha\Rightarrow A^2\alpha=\lambda^2\alpha$，则$r(A^2)=r(A)^2$，故$\|A^HA\|_2^2=r\big((A^HA)^2\big)=r(A^HA)^2=\big(\|A\|_2^2\big)^2$）

酉不变：$\|UAV\|_2=\|A\|_2$（用$\|A\|_2=\|A^H\|_2$证）

定理 5：

1. $\|A\|_2=\max_{\|x\|_2=\|y\|_2=1}|y^HAx|$

证：

（先证$\|A\|_2\ge\max_{\|x\|_2=\|y\|_2=1}|y^HAx|$）

$|y^HAx|\le\|y\|_2\|Ax\|_2$（柯西）$\le\|A\|_2\|x\|_2\|y\|_2$（相容）$=\|A\|_2$

（再证$\|A\|_2\le\max_{\|x\|_2=\|y\|_2=1}|y^HAx|$）

取x_0使得$\|x_0\|_2=1$，取$y_0=\frac{Ax_0}{\|Ax_0\|_2}$（自然有$\|y_0\|_2=1$），有$\max_{\|x\|_2=\|y\|_2=1}|y^HAx|\ge|y_0^HAx_0|=\|Ax_0\|_2=\|A\|_2$

故$\|A\|_2=\max_{\|x\|_2=\|y\|_2=1}|y^HAx|$

2. $\|A\|_2^2\le\|A\|_1\|A\|_\infty$

证：

$\|A\|_2^2=r(A^HA)\le\|A^HA\|_1$（相容的矩阵范数≥特征值）$\le\|A^H\|_1\|A\|_1=\|A\|_\infty\|A\|_1$（$\|A^H\|_1=\|A\|_\infty$）

6. 范数的应用

一、矩阵逆的摄动

定义 1：条件数：

$K_p(A)=\|A\|_p\left\|A^{-1}\right\|_p$是矩阵$A$相对矩阵范数$\|\cdot\|_p$的条件数（$K_p(A)\geq 1$）

定理 1：

当算子a范数$\|A\|_a<1$时，$E-A$可逆，且$\left\|(E-A)^{-1}\right\|_a\leq\frac{1}{1-\|A\|_a}$

证明可逆：$\|A\|_a\geq\lambda_i$，$1-\lambda_i\geq 1-\|A\|_a>0$，即$E-A$的的特征值$>0$
证明不等式：$E=(E-A)(E-A)^{-1}=(E-A)^{-1}-A(E-A)^{-1}$，则$(E-A)^{-1}=E+A(E-A)^{-1}$

则$\left\|(E-A)^{-1}\right\|_a=\left\|E+A(E-A)^{-1}\right\|_a\leq\|E\|_a+\left\|A(E-A)^{-1}\right\|_a$（三角不等式）$\leq\|E\|_a+\|A\|_a\left\|(E-A)^{-1}\right\|_a$（自相容）

即$\left\|(E-A)^{-1}\right\|_a\leq 1+\|A\|_a\left\|(E-A)^{-1}\right\|_a$，故$\left\|(E-A)^{-1}\right\|_a\leq\frac{1}{1-\|A\|_a}$

定理 2：

A可逆，δA为摄动矩阵（即计算机数值不稳定，准确矩阵为$A+\delta A$），且$\left\|A^{-1}\delta A\right\|_a<1$，则

1. $A+\delta A$可逆
2. $(A+\delta A)^{-1}=(E+F)A^{-1}$，其中$\|F\|_a\leq\frac{\left\|A^{-1}\delta A\right\|_a}{1-\left\|A^{-1}\delta A\right\|_a}$
3. $\frac{\left\|A^{-1}-(A+\delta A)^{-1}\right\|_a}{\left\|A^{-1}\right\|_a}\leq\frac{\left\|A^{-1}\delta A\right\|_a}{1-\left\|A^{-1}\delta A\right\|_a}$

推论 1：

$\|F\|_a\leq\frac{K(A)\frac{\|\delta A\|_a}{\|A\|_a}}{1-K(A)\frac{\|\delta A\|_a}{\|A\|_a}}$，$\frac{\left\|A^{-1}-(A+\delta A)^{-1}\right\|_a}{\left\|A^{-1}\right\|_a}\leq\frac{K(A)\frac{\|\delta A\|_a}{\|A\|_a}}{1-K(A)\frac{\|\delta A\|_a}{\|A\|_a}}$

（若$K(A)=\|A\|\left\|A^{-1}\right\|$的值越大，则$(A+\delta A)^{-1}$和$A^{-1}$的相对误差$\frac{\left\|A^{-1}-(A+\delta A)^{-1}\right\|_a}{\left\|A^{-1}\right\|_a}$就越大。即使$\delta A$相对于$A$可能很小，也可能引起$A^{-1}$很大的偏差。$K(A)$较大的方阵称为病态的，较小称为良态的）

第三章 矩阵的分解

1. 矩阵的三角分解

一、n阶方阵的三角分解

定义 1：正线上三角矩阵：
上三角矩阵的对角元均为正实数。如果对角元均为 1，则称为单位上三角矩阵。
定义 2：正线下三角矩阵：
下三角矩阵的对角元均为正实数。如果对角元均为 1，则称为单位下三角矩阵。
性质：
1. 上三角矩阵R的逆R^{-1}也是上三角矩阵，且对角元是R对角元的倒数
2. 两个上三角矩阵R_1和R_2的乘积R_1R_2也是上三角矩阵，且对角元是R_1和R_2对角元之积
3. 酉矩阵U的逆U^{-1}也是酉矩阵
4. 两个酉矩阵U_1和U_2的乘积U_1U_2也是酉矩阵

定理 1（复矩阵UR分解）：
n阶复方阵A可唯一分解为$A = U_1R$，或唯一分解为$A = LU_2$
其中U_1、U_2是酉矩阵，R是正线上三角复矩阵，L是正线下三角复矩阵
推论 1（实矩阵QR分解）：
n阶实矩阵A可唯一分解为$A = Q_1R$，或唯一分解为$A = LQ_2$
其中Q_1、Q_2是正交矩阵，R是正线上三角实矩阵，L是正线下三角实矩阵
求QR分解（以三阶实矩阵为例）：
1. 把A按列分块$A = (\alpha_1, \alpha_2, \alpha_3)$
2. 施密特正交化：$\begin{cases} \beta_1 = \alpha_1 \\ \beta_2 = \alpha_2 - \frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1 \\ \beta_3 = \alpha_3 - \frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1 - \frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2 \end{cases}$，单位化$\begin{cases} \gamma_1 = \frac{\beta_1}{\|\beta_1\|} \\ \gamma_2 = \frac{\beta_2}{\|\beta_2\|} \\ \gamma_3 = \frac{\beta_3}{\|\beta_3\|} \end{cases}$
3. 把α_i移到等式左边，$\begin{cases} \alpha_1 = \beta_1 \\ \alpha_2 = \frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1 + \beta_2 \\ \alpha_3 = \frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1 + \frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2 + \beta_3 \end{cases}$，即$A = (\alpha_1, \alpha_2, \alpha_3) = (\beta_1, \beta_2, \beta_3)\begin{pmatrix} 1 & \frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)} & \frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)} \\ 0 & 1 & \frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)} \\ 0 & 0 & 1 \end{pmatrix}$

又$(\beta_1, \beta_2, \beta_3) = (\gamma_1, \gamma_2, \gamma_3)\begin{pmatrix} \|\beta_1\| & & \\ & \|\beta_2\| & \\ & & \|\beta_3\| \end{pmatrix}$，则

$$A = (\alpha_1, \alpha_2, \alpha_3) = (\gamma_1, \gamma_2, \gamma_3)\begin{pmatrix} \|\beta_1\| & & \\ & \|\beta_2\| & \\ & & \|\beta_3\| \end{pmatrix}\begin{pmatrix} 1 & \frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)} & \frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)} \\ 0 & 1 & \frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)} \\ 0 & 0 & 1 \end{pmatrix} = (\gamma_1, \gamma_2, \gamma_3)\begin{pmatrix} \|\beta_1\| & (\alpha_2, \beta_1) & (\alpha_3, \beta_1) \\ 0 & \|\beta_2\| & (\alpha_3, \beta_2) \\ 0 & 0 & \|\beta_3\| \end{pmatrix}$$

其中$Q = (\gamma_1, \gamma_2, \gamma_3)$是正交阵，$R = \begin{pmatrix} \|\beta_1\| & (\alpha_2, \beta_1) & (\alpha_3, \beta_1) \\ 0 & \|\beta_2\| & (\alpha_3, \beta_2) \\ 0 & 0 & \|\beta_3\| \end{pmatrix}$是上三角矩阵，有$A = QR$

推论 2（实对称正定矩阵分解）：
唯一分解$A = R^TR$，R是正线上三角实矩阵

$$A = Q^{-1}\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}Q\text{（}Q\text{为正交矩阵）}$$

$$= Q^T\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \sqrt{\lambda_2} & \\ & & \sqrt{\lambda_3} \end{pmatrix}\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \sqrt{\lambda_2} & \\ & & \sqrt{\lambda_3} \end{pmatrix}Q = P^TP\text{（设}P = Q^T\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \sqrt{\lambda_2} & \\ & & \sqrt{\lambda_3} \end{pmatrix}Q\text{）}$$

$= (QR)^T(QR) = R^TQ^TQR = R^TR$（可逆阵$P$做$QR$分解）

推论 3（正定 Hermite 矩阵分解）：

把推论 2 推广到复矩阵

定理 2（复矩阵LU分解）：

A各阶顺序主子式不为 0，则A可唯一分解为$A = LR^* = L^*R = L^*DR^*$

其中R、L是上三角、下三角阵，R^*、L^*是单位上三角、下三角阵，D是对角阵

二、任意方阵的三角分解

定义 3：满秩矩阵：

秩等于列数称列满秩，秩等于行数称行满秩

定理 3（满秩矩阵的三角分解）：

1. A行满秩，则存在$A = (L \quad O)U$，其中L为m阶正线下三角阵，U为n阶酉矩阵
2. A列满秩，则存在$A = U\begin{pmatrix} R \\ O \end{pmatrix}$，其中$R$为$n$阶正线上三角阵，$U$为$m$阶酉矩阵

定理 4（满秩矩阵的三角分解 2）：

1. A行满秩，则唯一$A = LU$，其中L为m阶正线下三角阵，U为$m \times n$类酉型矩阵（酉矩阵摘出来几行或者几列）
2. A列满秩，则唯一$A = UR$，其中R为n阶正线上三角阵，U为$m \times n$类酉型矩阵

定理 5（三角分解）：

$m \times n$矩阵A的秩为r，则存在$A = U\begin{pmatrix} L & O \\ O & O \end{pmatrix}V$，其中$L$为正线下三角阵，$U$、$V$为$m$、$n$阶酉矩阵。

证：

$AP = (\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_n) = (A_1 \quad A_2)$，其中$P$是置换矩阵（$P^{-1}$仍为置换阵），$A_1 = (\alpha_1, \alpha_2, \ldots, \alpha_r)$是极大无关组。

故可写作$AP = A_1(E_r \quad C)$，分解$m \times r$列满秩矩阵$A_1 = U\begin{pmatrix} R \\ O \end{pmatrix}$，即$AP = U\begin{pmatrix} R \\ O \end{pmatrix}(E_r \quad C) = U\begin{pmatrix} R & RC \\ O & O \end{pmatrix}$

$(R \quad RC)$是$r \times n$行满秩矩阵，分解为$(L \quad O)V_1$，故$AP = U\begin{pmatrix} R & RC \\ O & O \end{pmatrix} = U\begin{pmatrix} L & O \\ O & O \end{pmatrix}V_1$

故$A = U\begin{pmatrix} L & O \\ O & O \end{pmatrix}V_1P^{-1} = U\begin{pmatrix} L & O \\ O & O \end{pmatrix}V$

推论 4：

存在$A = U\begin{pmatrix} R & O \\ O & O \end{pmatrix}V$。同上。

2. 矩阵的谱分解

一、单纯矩阵的谱分解

定义 1：代数重复度、几何重复度（代数重数、几何重数）：

代数重数：特征值的重数，特征多项式根的重数

几何重数：特征子空间（$Ax = \lambda_i x$的解空间）的维数，线性无关的特征向量的个数

定理 1：

几何重数不大于代数重数

定义 2：单纯矩阵：

代数重数和几何重数相等

定理 2（单纯矩阵的充要条件）：

可相似对角化

定理 3（谱分解）：

单纯矩阵A可分解为一系列幂等矩阵A_i的加权和，即$A = \sum_{i=1}^n \lambda_i A_i$，$\lambda_i$是$A$的特征值

单纯矩阵$A = P\Lambda P^{-1}$，P是可逆矩阵，分块$P = (v_1, v_2, \ldots, v_n)$，$P^{-1} = \left(w_1^T, w_2^T, \ldots, w_n^T\right)^T$

则$A = P\Lambda P^{-1} = (v_1, v_2, \ldots, v_n)\Lambda\left(w_1^T, w_2^T, \ldots, w_n^T\right)^T = \sum_{i=1}^n \lambda_i v_i w_i^T = \sum_{i=1}^n \lambda_i A_i$，其中$A_i = v_i w_i^T$是幂等矩阵

A_i的性质：$\begin{cases} A_iA_j = \begin{cases} A_i，i=j（幂等性）\\ O，i\neq j（分离性）\end{cases} \\ \sum_{i=1}^{n} A_i = E（可加性）\end{cases}$

A的多项式$f(A)=\sum_{i=1}^{n} f(\lambda_i)A_i$

Hamilton-Caylay 定理：A的特征多项式为$f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\ldots+a_{n-1}\lambda+a_n$

则$f(A)=A^n+a_1A^{n-1}+\ldots+a_{n-1}A+a_nE=0$

则$A^n=-(a_1A^{n-1}+\ldots+a_{n-1}A+a_nE)$，即$A^n$是$A^{n-1},A^{n-2},\ldots,A^2,A,E$的线性组合

即任意A^m（$m\geq n$）都是$A^{n-1},A^{n-2},\ldots,A^2,A,E$的线性组合

也有$E=-\frac{1}{a_n}(A^n+a_1A^{n-1}+\ldots+a_{n-1}A)=-\frac{1}{a_n}(A^{n-1}+a_1A^{n-2}+\ldots+a_{n-1}E)A$

即$A^{-1}=-\frac{1}{a_n}(A^{n-1}+a_1A^{n-2}+\ldots+a_{n-1}E)$，故$A^{-1}=\sum_{i=1}^{n} f(\lambda_i)A_i=-\frac{1}{a_n}\sum_{i=1}^{n}\left(\lambda_i^{n-1}+a_1\lambda_i^{n-2}+\ldots+a_{n-1}\right)A_i$

定理 4（谱分解定理）：

A有k个不同特征值，则A是单纯矩阵的充要条件是存在k个矩阵A_i（$1\leq i\leq k$），使$\begin{cases} A_iA_j = \begin{cases} A_i，i=j（幂等性）\\ O，i\neq j（分离性）\end{cases} \\ \sum_{i=1}^{k} A_i = E（可加性）\\ A=\sum_{i=1}^{k}\lambda_iA_i（分解）\end{cases}$

（谱分解和单纯矩阵是等价的）

证：（必要性）A是单纯矩阵，则由定理 3 得$A=\sum_{i=1}^{n} l_iB_i$。如果特征值l_i互不相同，结论自然成立。

如果有相同l_i，则将相同特征值进行合并，得到$A=\sum_{i=1}^{k}\lambda_iA_i$，其中$A_i=\sum_{j=1}^{r_i} B_{ij}$

（B_{ij}指特征值λ_i对应的矩阵，r_i为λ_i的重数）

则$A_iA_j=\begin{cases} \sum_{j=1}^{r_i} B_{ij}\sum_{j=1}^{r_i} B_{ij}=\sum_{j=1}^{r_i} B_{ij}^2=\sum_{j=1}^{r_i} B_{ij}=A_i，i=j（幂等性）\\ \sum_{j_1=1}^{r_i} B_{ij_1}\sum_{j_2=1}^{r_j} B_{jj_2}=O，i\neq j（分离性）\end{cases}$，且$\sum_{i=1}^{k} A_i=\sum_{i=1}^{n} B_i=E$（可加性）

故必要性得证。

（充分性）设$\operatorname{rank} A_i=\alpha_i$（$1\leq i\leq k$），有$\sum_{i=1}^{k}\alpha_i=\sum_{i=1}^{k} rank\, A_i=\sum_{i=1}^{k} tr\,(A_i)=tr\left(\sum_{i=1}^{k} A_i\right)=tr\,(E)=n$

（A_i是幂等矩阵$A_i^2=A_i$，也就有$\lambda^2=\lambda$，特征值非 0 即 1，故秩=迹）

则对A_i进行三角分解（第三章第 1 节定理 5），有$A_i=U_i\begin{pmatrix} L_i & O \\ O & O\end{pmatrix}V_i$

令$X_i=U_i\begin{pmatrix} L_i \\ 0\end{pmatrix}$（$X_i$为$n\times\alpha_i$矩阵），$V_i=\begin{pmatrix} Y_{i1} \\ Y_{i2}\end{pmatrix}$（$Y_{i1}$为$\alpha_i\times n$矩阵），有$A_i=X_iY_{i1}$

令$X=(X_1,\ldots,X_k)$，$Y=\begin{pmatrix} Y_{11} \\ \vdots \\ Y_{k1}\end{pmatrix}$（由$\sum_{i=1}^{k}\alpha_i=n$得$X$、$Y$均为$n$阶方阵），$XY=\sum_{i=1}^{k} X_iY_{i1}=\sum_{i=1}^{k} A_i=E$（用可加性）

于是X、Y互为逆矩阵，则$YX=\begin{pmatrix} Y_{11} \\ \vdots \\ Y_{k1}\end{pmatrix}(X_1,\ldots,X_k)=\begin{pmatrix} Y_{11}X_1 & \cdots & Y_{11}X_k \\ \vdots & \ddots & \vdots \\ Y_{k1}X_1 & \ldots & Y_{k1}X_k\end{pmatrix}=E$

对应块相等，有$Y_{i1}X_j=\begin{cases} E，i=j \\ O，i\neq j\end{cases}$，从而$A_iX_j=X_iY_{i1}X_j=\begin{cases} X_i，i=j \\ O，i\neq j\end{cases}$，

于是$AX=\sum_{i=1}^{k}\lambda_iA_i\,(X_1,\ldots,X_k)=\sum_{i=1}^{k}\lambda_i(A_iX_1,\ldots,A_iX_k)=\sum_{i=1}^{k}\lambda_i(0,\ldots,0,X_i,0,\ldots,0)$

$=(\lambda_1X_1,\lambda_2X_2,\ldots,\lambda_kX_k)=(X_1,\ldots,X_k)\begin{pmatrix} \lambda_1E_{\alpha_1} & & \\ & \ddots & \\ & & \lambda_kE_{\alpha_k}\end{pmatrix}=X\Lambda$（$\Lambda$是对角矩阵）

$AX=X\Lambda$即$A=X\Lambda X^{-1}$，所以A是单纯矩阵，充分性得证。

谱分解：

1. 求特征值λ_i以及对应的特征向量α_i（注意这里不同特征值对应的特征向量不能正交化）
2. 令$P=(\alpha_1,\alpha_2,\dots,\alpha_n)$，求出$P^{-1}$，令$P^{-1}=\begin{pmatrix}\beta_1^T\\ \vdots\\ \beta_n^T\end{pmatrix}$，得$A_i=\alpha_i\beta_i^T$，则$A=\sum_{i=1}^n\lambda_iA_i$

二、正规矩阵及其分解

定义 3：正规矩阵：

如果$A^HA=AA^H$，则称A为正规矩阵。

如果实矩阵A有$A^TA=AA^T$，则称A为实正规矩阵。

对角矩阵、酉矩阵、Hermite 矩阵、反 Hermite 矩阵都是正规矩阵

正交矩阵、实对称矩阵、实反对称矩阵都是实正规矩阵

引理 1：

A是正规矩阵，A和B酉相似（$U^{-1}BU=U^HBU=A$，U是酉矩阵），则B也是正规矩阵。

引理 2（Schur 分解）：

任意矩阵A和上三角矩阵R酉相似（$URU^{-1}=A$），R的主对角线元素为A的特征值。

引理 3：

三角矩阵A是正规矩阵，当且仅当A是对角矩阵。

定理 5：

A是正规矩阵的充要条件是A和对角矩阵酉相似。

（充分性）$U\Lambda U^{-1}=A$，Λ是正规矩阵，所以A是正规矩阵。

（必要性）A是正规矩阵，$URU^{-1}=A$，R是上三角矩阵，也是酉矩阵，则R是对角矩阵。

定理 6（正规矩阵的谱分解定理）：

A有k个不同特征值，则（A是正规矩阵，当且仅当谱分解的A_i都是 Hermite 矩阵）

A是正规矩阵的充要条件是存在k个矩阵A_i（$1\le i\le k$），使$\begin{cases}A_iA_j=\begin{cases}A_i，i=j\text{（幂等性）}\\ O，i\ne j\text{（分离性）}\end{cases}\\ \sum_{i=1}^k A_i=E\text{（可加性）}\\ A=\sum_{i=1}^k\lambda_iA_i\text{（分解）}\\ A_i^H=A_i\text{（Hermite）}\end{cases}$

定理 7（正规矩阵的性质）：

1. 存在U^HAU和U^HA^HU均为对角矩阵（A和A^H可酉相似对角化）
2. A是单纯矩阵（正规矩阵是单纯矩阵的特例）
3. 若$Ax=\lambda_ix$（$x\ne 0$），则$A^Hx=\overline{\lambda}_ix$（$A$和$A^H$特征值共轭，且对应特征向量相同）
4. 属于A的不同特征值的特征向量正交（和实正规矩阵结论相同）

三、与 Jordan 标准形矩阵相似的矩阵的分解

3. Hermite 矩阵及其分解

定义 1：Hermite 矩阵：

如果$A^H=A$则称A为 Hermite 矩阵；如果$A^H=-A$则称A为反 Hermite 矩阵

定理 1（Hermite 矩阵的性质）：

1. $(A\alpha,\beta)=(\alpha,A\beta)$（$A$和内积运算可以交换顺序）
2. A的特征值均为实数
3. 属于A的不同特征值的特征向量正交

（Hermite 矩阵是正规矩阵的特例）

推论 1：

Hermite 矩阵A有$\operatorname{rank}A=r$，则A与矩阵$\begin{pmatrix}E_p & & \\ & -E_{r-p} & \\ & & 0\end{pmatrix}$合同，$p$为正惯性系数，$q=r-p$为负惯性系数

定理 2：

反 Hermite 矩阵A的特征值均为纯虚数。

定义 2：正定二次型、正定矩阵：

$f(X) = X^H AX > 0$ 称为正定二次型，A为正定矩阵

定理 3（Cholesky 分解）：

$A = U\Lambda U^H = U\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}U^H = U\Lambda^{\frac{1}{2}}VV^H\Lambda^{\frac{1}{2}}U^H = \left(U\Lambda^{\frac{1}{2}}V\right)\left(U\Lambda^{\frac{1}{2}}V\right)^H = LL^H$，$L$是下三角阵

定理 4（Hermite 矩阵A正定的等价条件）：

1. 特征值全为正实数
2. A与E合同
3. A的顺序主子式全为正

定理 5：（正定 Hermite 矩阵A的性质）

1. A的主对角线元素均大于零
2. 存在正定 Hermite 矩阵，使得$A = B^2$
3. A的任意k行和对应的k列组成的主子阵$\begin{pmatrix} a_{i_1i_1} & \cdots & a_{i_1i_k} \\ \vdots & \ddots & \vdots \\ a_{i_ki_1} & \cdots & a_{i_ki_k} \end{pmatrix}$（$1 \le i_1 < \ldots < i_k \le n$）是正定的
4. $|A| \le \prod_{i=1}^n a_{ii}$，当且仅当$A$为对角矩阵时等号成立

定理 6：（Hermite 矩阵A半正定的等价条件）：

1. 特征值全非负
2. A与$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$合同
3. A的顺序主子式非负

定理 7：（半正定 Hermite 矩阵A的性质）

1. A的主对角线元素均非负
2. 存在半正定 Hermite 矩阵，使得$A = B^2$

定理 8：（同时对角化）

A是正定 Hermite 矩阵，B是 Hermite 矩阵，则存在可逆矩阵T，使得$T^H AT = E$，$T^H BT = \Lambda$。

（Λ是对角阵，其对角元为$A^{-1}B$的特征值）

推论 2：

若A、B均为正定 Hermite 矩阵，$A - B$半正定，则A^{-1}、B^{-1}均为正定 Hermite 矩阵，$B^{-1} - A^{-1}$半正定

证：

$T^H AT = E$，$T^H BT = \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$（$\lambda_i > 0$）。$T^H(A - B)T = E - \Lambda = \begin{pmatrix} 1 - \lambda_1 & & \\ & \ddots & \\ & & 1 - \lambda_n \end{pmatrix}$（$1 - \lambda_i \ge 0$）

则$(T^H AT)^{-1} = T^{-1}A^{-1}(T^H)^{-1} = T^{-1}A^{-1}(T^{-1})^H = E$，所以$A^{-1}$是正定 Hermite 矩阵。

同理，$T^{-1}B^{-1}(T^{-1})^H = \Lambda^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & & \\ & \ddots & \\ & & \frac{1}{\lambda_n} \end{pmatrix}$，$B^{-1}$是正定 Hermite 矩阵。

$T^{-1}(B^{-1} - A^{-1})(T^{-1})^H = \Lambda^{-1} - E = \begin{pmatrix} \frac{1}{\lambda_1} - 1 & & \\ & \ddots & \\ & & \frac{1}{\lambda_n} - 1 \end{pmatrix} = \begin{pmatrix} \frac{1-\lambda_1}{\lambda_1} & & \\ & \ddots & \\ & & \frac{1-\lambda_n}{\lambda_n} \end{pmatrix}$（$\frac{1-\lambda_1}{\lambda_1} \ge 0$），所以$B^{-1} - A^{-1}$半正定

4. 矩阵的最大秩分解

定理 1（最大秩分解、满秩分解）：

$m \times n$矩阵A的秩为r，则存在$m \times r$列满秩矩阵B和$r \times n$行满秩矩阵D，使得$A = BD$

证：

$AP = (\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_n) = (A_1 \quad A_1C)$，其中$P$是置换矩阵，$A_1 = (\alpha_1, \alpha_2, \ldots, \alpha_r)$是极大无关组。

故可写作$AP = A_1(E_r \quad C)$，则$A = A_1(E_r \quad C)P^{-1}$
则$B = A_1$（$m \times r$列满秩），$D = (E_r \quad C)P^{-1}$（$r \times n$行满秩）
满秩分解步骤：
将A化为行最简形$\widetilde{A}$，$\widetilde{A}$每行第 1 个 1 所在的列对应在A列元素组成的$m \times r$列满秩矩阵就是B，$\widetilde{A}$中前r行就是D。
（比如说$A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 2 & 4 \end{pmatrix}$，$\widetilde{A} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$，则$B$为$A$的第 2 列$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$，$D$为$\widetilde{A}$的第 1 行$(0 \quad 1 \quad 2)$）
（也可以按列变换）
定理 2（两个最大秩分解的关系）：
1. $B_1 = B_2Q$，$D_1 = Q^{-1}D_2$（Q为可逆阵）
2. $D_1^H\left(D_1D_1^H\right)^{-1}\left(B_1^HB_1\right)^{-1}B_1^H = D_2^H\left(D_2D_2^H\right)^{-1}\left(B_2^HB_2\right)^{-1}B_2^H$（M－P 广义逆矩阵唯一）

5. 矩阵的奇异值分解

定理 1：
1. $\operatorname{rank} A = \operatorname{rank} A^HA = \operatorname{rank} AA^H$
2. A^HA、AA^H的特征值均为非负实数
3. A^HA、AA^H的非零特征值相同

证：
1. 如果$A^HAx = 0$，必有$x^HA^HAx = \|Ax\|^2 = 0$，即$Ax = 0$，也就是说$A^HAx = 0$ 的解空间包含了$Ax = 0$ 的解空间
则$\operatorname{rank} A^HA \geq \operatorname{rank} A$。又$\operatorname{rank} A \geq \operatorname{rank} A^HA$，所以$\operatorname{rank} A = \operatorname{rank} A^HA$。又有$\operatorname{rank} A^H = \operatorname{rank} AA^H$，$\operatorname{rank} A = \operatorname{rank} A^H$，得证。
2. 若$A^HA\alpha = \lambda\alpha$，则$(A\alpha, A\alpha) = \alpha^HA^HA\alpha = \lambda\alpha^H\alpha$，其中$\alpha^H\alpha > 0$，$(A\alpha, A\alpha) \geq 0$，故$\lambda \geq 0$。$AA^H$的特征值同理。
3. 若$A^HA\alpha = \lambda\alpha$，则$AA^HA\alpha = \lambda A\alpha$，即$\lambda$也是$AA^H$的特征值，特征向量为$A\alpha$。同理，可证$A^HA$、$AA^H$的非零特征值互相存在。
设λ是A^HA的p重非零特征值，A^HA对应于λ的特征子空间的一组基是$\alpha_1, \alpha_2, \ldots, \alpha_p$。
令$k_1A\alpha_1 + k_2A\alpha_2 + \ldots + k_pA\alpha_p = 0$，即$A\left(k_1\alpha_1 + k_2\alpha_2 + \ldots + k_p\alpha_p\right) = 0$，则$A^HA\left(k_1\alpha_1 + k_2\alpha_2 + \ldots + k_p\alpha_p\right) = 0$。
因为$\alpha_1, \alpha_2, \ldots, \alpha_p$是$A^HA$对应于$\lambda$的特征向量，则$A^HA\left(k_1\alpha_1 + k_2\alpha_2 + \ldots + k_p\alpha_p\right) = \lambda\left(k_1\alpha_1 + k_2\alpha_2 + \ldots + k_p\alpha_p\right) = 0$。
因为$\lambda \neq 0$，所以只能$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_p\alpha_p = 0$。因为$\alpha_1, \alpha_2, \ldots, \alpha_p$线性无关，只能$k_1 = k_2 = \ldots = k_p = 0$。
因此$A\alpha_1, A\alpha_2, \ldots, A\alpha_p$线性无关。
说明λ是AA^H的至少p重非零特征值。反之，AA^H的p重非零特征值也是A^HA的至少p重非零特征值。
也即说明A^HA、AA^H的非零特征值重数相同。
定义 1：奇异值：
A^HA的特征值有$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$，则称$\sigma_i = \sqrt{\lambda_i}$为矩阵$A$的正奇异值。
定义 2：酉等价：
$A = UBV$，U、V为酉矩阵，则A、B酉等价。
定理 2（酉等价的性质）：
A、B酉等价，则A、B具有相同的正奇异值。
证：$A^HA = V^HB^HU^HUBV = V^HB^HBV$，则$A^HA$和$B^HB$酉相似，则$A^HA$和$B^HB$特征值相同，则$A$和$B$正奇异值相同。
定理 3（奇异值分解，SVD 分解）：

$m \times n$矩阵A的秩为r，则存在$A = U\begin{pmatrix} D & O \\ O & O \end{pmatrix}V$，其中$D = \begin{pmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_r \end{pmatrix}$，其中$|\delta_i| = \sigma_i$，$U$、$V$为$m$、$n$阶酉矩阵。

奇异值分解步骤：

①（求V、D）A^HA酉相似对角化，即$VA^HAV^H = \begin{pmatrix} D^HD & O \\ O & O \end{pmatrix}$，其中$D^HD = \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2 \end{pmatrix}$，则$D = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}$

V分块为$V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$，$V_1$是$r \times n$矩阵，$V_2$是$(n-r) \times n$矩阵。则$VA^HAV^H = \begin{pmatrix} V_1A^HAV_1^H & V_1A^HAV_2^H \\ V_2A^HAV_1^H & V_2A^HAV_2^H \end{pmatrix} = \begin{pmatrix} D^HD & O \\ O & O \end{pmatrix}$

对应相等，有$V_1A^HAV_1^H = D^HD$，$V_2A^HAV_2^H = O = \left(AV_2^H\right)^HAV_2^H$，所以$AV_2^H = O$

②（求U）令$U_1^H = (D^H)^{-1}V_1A^H$是$r \times m$类酉型矩阵，将U_1^H扩充成m阶矩阵酉矩阵$U^H = \begin{pmatrix} U_1^H \\ U_2^H \end{pmatrix}$（$U_1^Hx = 0$ 的基础解系正交单位

化可得到U_2），即求出U。

③（验证$A=U\begin{pmatrix}D & O\\ O & O\end{pmatrix}V$）

第四章 特征值的估计与摄动

1.特征值界的估计

定理 1（Schur）：

A的特征值为$\lambda_1,\lambda_2,...,\lambda_n$，则$\sum_{i=1}^{n}|\lambda_i|^2\le\|A\|_F^2=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2$（$F$范数即矩阵 2 范数$\|\cdot\|_{m2}$）

证：由第三章第 2 节引理 2（Schur）有$URU^H=A$，R是上三角矩阵，主对角线元素为A的特征值。
矩阵 2 范数酉不变，故$\|A\|_F^2=\|R\|_F^2\ge\sum_{i=1}^{n}|\lambda_i|^2$

定理 2（Hirsch）：

$A=B+C$，其中$\begin{cases}B=(b_{ij})=\frac{(A+A^H)}{2}\text{（Hermite 部分）}\\ C=(c_{ij})=\frac{(A-A^H)}{2}\text{（反 Hermite 部分）}\end{cases}$，则$\begin{cases}|\lambda_i|\le n\max\limits_{i,j}|a_{ij}|\\ |\mathrm{Re}\,(\lambda_i)|\le n\max\limits_{i,j}|b_{ij}|\\ |\mathrm{Im}\,(\lambda_i)|\le n\max\limits_{i,j}|c_{ij}|\end{cases}$（$n\max\limits_{i,j}|a_{ij}|=\|\cdot\|_a$范数）

证：$U^HAU=R$，$U^HA^HU=R^H$，R是上三角矩阵，则$\begin{cases}U^HBU=\frac{(R+R^H)}{2}\\ U^HCU=\frac{(R-R^H)}{2}\end{cases}$

则$\begin{cases}|\mathrm{Re}\,(\lambda_i)|^2\le\sum_{i=1}^{n}|\mathrm{Re}\,(\lambda_i)|^2=\sum_{i=1}^{n}\left|\frac{\lambda_i+\bar{\lambda}_i}{2}\right|^2\le\sum_{i=1}^{n}\sum_{j=1}^{n}|b_{ij}|^2\le n^2\max\limits_{i,j}|b_{ij}|^2\\ |\mathrm{Im}\,(\lambda_i)|^2\le\sum_{i=1}^{n}|\mathrm{Im}\,(\lambda_i)|^2=\sum_{i=1}^{n}\left|\frac{\lambda_i-\bar{\lambda}_i}{2}\right|^2\le\sum_{i=1}^{n}\sum_{j=1}^{n}|c_{ij}|^2\le n^2\max\limits_{i,j}|c_{ij}|^2\end{cases}$（定理 1 得第二个≤），故$\begin{cases}|\mathrm{Re}\,(\lambda_i)|\le n\max\limits_{i,j}|b_{ij}|\\ |\mathrm{Im}\,(\lambda_i)|\le n\max\limits_{i,j}|c_{ij}|\end{cases}$

又由定理 1 得$|\lambda_i|^2\le\sum_{i=1}^{n}|\lambda_i|^2\le\|A\|_{m2}^2\le n^2\max\limits_{i,j}|a_{ij}|^2$，故$|\lambda_i|\le n\max\limits_{i,j}|a_{ij}|$，故得证。

定理 3（Bendixson）：

如果限制A为实矩阵：$|\mathrm{Im}\,(\lambda_i)|\le\sqrt{\frac{n(n-1)}{2}}\max\limits_{i,j}|c_{ij}|$

证：A为实矩阵，则$c_{ii}=0$。
第三章第 3 节定理 2 知，反 Hermite 矩阵C的特征值为纯虚数，并由于C是实矩阵，其特征值成共轭对出现。

不妨设共有s对共轭特征值，则$2\sum_{i=1}^{s}|\mathrm{Im}\,(\lambda_i)|^2=\sum_{i=1}^{n}|\mathrm{Im}\,(\lambda_i)|^2\le\sum_{i=1}^{n}\sum_{j=1}^{n(j\ne i)}|c_{ij}|^2\le n(n-1)\max\limits_{i,j}|c_{ij}|^2$

因此$|\mathrm{Im}\,(\lambda_i)|^2\le\frac{n(n-1)}{2}\max\limits_{i,j}|c_{ij}|^2$，故得证

推论：
Hermite 矩阵的特征值为实数，反 Hermite 矩阵的特征值为纯虚数。

2.Gerschgorin 圆盘定理

定义 1：Gerschgorin 圆（盖尔圆）：

称$S_i=\{z\in C:|z-a_{ii}|\le R_i\}$为矩阵$A$在复平面上的第$i$个（行）盖尔圆，其中（$R_i=\sum_{j\ne i}|a_{ij}|$，$i=1,2,...,n$）称为$S_i$的半径

定理 1（Gerschgorin 圆盘定理 1）：
$\lambda_i\in S=\bigcup_{j=1}^{n}S_j$（$i=1,2,...,n$），即矩阵$A$的任一特征值都属于所有盖尔圆的并集。

证：设λ为$A=(a_{ij})$的任一特征值，对应的特征向量为$x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$，即$Ax=\lambda x$，展开为$\sum_{j=1}^{n}a_{ij}x_j=\lambda x_i$，（$i=1,2,...,n$）

假设$|x_k|=\max\limits_{1\le i\le n}(|x_i|)$（最大绝对值），由$x\ne 0$得$|x_k|>0$，由上式代入$i=k$有$\sum_{j=1}^{n}a_{kj}x_j=\lambda x_k$。

把左式中的$a_{kk}x_k$移项到右式，得$\sum_{j=1,j\neq k}^{n}a_{kj}x_j=\lambda x_k-a_{kk}x_k=x_k(\lambda-a_{kk})$

两边取绝对值有$|x_k||\lambda-a_{kk}|=\left|\sum_{j=1,j\neq k}^{n}a_{kj}x_j\right|\le\sum_{j=1,j\neq k}^{n}|a_{kj}||x_j|$（$|x_j|$放大为$|x_k|$）$\le|x_k|\sum_{j=1,j\neq k}^{n}|a_{kj}|=|x_k|R_k$

因此$|\lambda-a_{kk}|\le R_k$，从而$\lambda\in S_k\subset S$

定理 2（Gerschgorin 圆盘定理 2）：

A的k个并集形成一个联通区域（相交或相切），且它与余下的$n-k$个圆盘都不相交，则在这个区域中恰好有A的k个特征值。

证：$A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&\ddots&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix}=\begin{pmatrix}a_{11}&\cdots&0\\ \vdots&\ddots&\vdots\\ 0&\cdots&a_{nn}\end{pmatrix}+\begin{pmatrix}0&\cdots&a_{1n}\\ \vdots&\ddots&\vdots\\ a_{n1}&\cdots&0\end{pmatrix}=D+B$（$D$为对角矩阵，$B$为对角之外的矩阵）

令$A_\varepsilon=D+\varepsilon B$（$0\le\varepsilon\le1$），$R_i(A_\varepsilon)=R_i(\varepsilon B)=\varepsilon R_i(A)$。

当$\varepsilon=0$时，$A_\varepsilon=D$，所有的盖尔圆都退化为一个点，每个点里都有一个特征值（如果有几个相同a_{ii}，则有几个特征值）。

当$\varepsilon=1$时，$A_\varepsilon=A$。因为特征值是A的元素的连续函数，所以有多少盖尔圆相交，也即在该区域内原本有多少点，随着ε从 0 变到 1，这些点只能在联通区域内运动，而无法离开其边界。而外边不属于联通区域的特征值也无法进入这一区域，所以这一区域只能有这些特征值。

（联通部分中不一定每个盖尔圆里都有且仅有一个特征值）

推论 1：

$\lambda\in G=\bigcup_{j=1}^{n}G_j$，即矩阵$A$的任一特征值都属于所有列盖尔圆$G_j=\left\{z\in C:|z-a_{jj}|\le\sum_{i=1,i\neq j}^{n}|a_{ij}|\right\}$的并集。

推论 2：

$\lambda\in T=\left(\bigcup_{i=1}^{n}S_i\right)\cap\left(\bigcup_{j=1}^{n}G_j\right)$。$A$的任一特征值都属于行盖尔圆和列盖尔圆的交集。

推论 3：

设矩阵A的n个圆盘两两互不相交，则A是单纯矩阵（相似于对角阵）。

证：n个圆盘两两互不相交，每个圆盘里有且仅有一个特征值，故n个特征值互不相同，因此一定可以相似对角化

推论 4：

设n阶实矩阵A的n个圆盘两两互不相交，则A的特征值全为实数。

证：A是实矩阵，a_{ii}是实数，盖尔圆一定在实轴上。如果A存在复特征值，则一定成共轭对出现，即关于实轴对称。这样的话一定会出现某个盖尔圆内有至少两个特征值，由于n个圆盘互不相交，每个圆内一定有且仅有一个特征值，出现矛盾。因此A不存在复特征值。

定理 3（放缩盖尔圆）：

$Q_i=\left\{z\in C:|z-a_{ii}|\le\frac{1}{p_i}\sum_{j=1,j\neq i}^{n}p_j|a_{ij}|\right\}$，$P_j=\left\{z\in C:|z-a_{ii}|\le p_j\sum_{i=1,i\neq j}^{n}\frac{1}{p_i}|a_{ij}|\right\}$

则$\lambda\in T=\left(\bigcup_{i=1}^{n}Q_i\right)\cap\left(\bigcup_{j=1}^{n}P_j\right)$

证：$D^{-1}AD=B$，A和B相似则具有相同的特征值。取$D=\begin{pmatrix}p_1&\cdots&0\\ \vdots&\ddots&\vdots\\ 0&\cdots&p_n\end{pmatrix}$，$D^{-1}=\begin{pmatrix}\frac{1}{p_1}&\cdots&0\\ \vdots&\ddots&\vdots\\ 0&\cdots&\frac{1}{p_n}\end{pmatrix}$，则$B=\begin{pmatrix}a_{11}&\cdots&\frac{p_n}{p_1}a_{1n}\\ \vdots&\ddots&\vdots\\ \frac{p_1}{p_n}a_{n1}&\cdots&a_{nn}\end{pmatrix}$

所以行盖尔圆条件变为$|z-a_{ii}|\le\frac{1}{p_i}\sum_{j=1,j\neq i}^{n}p_j|a_{ij}|$，列也同理。由推论 2 得证。

定义 2：对角占优：

若$|a_{ii}|\ge R_i(A)$，则称A为行对角占优矩阵。如果在列上成立，则称列对角占优。如果不等式严格成立，称严格对角占优。

（对角元的绝对值大于（等于）其余元素绝对值的和）

定理 4（严格对角占优的性质）：

1. A为可逆矩阵，且$\lambda_i\in S=\bigcup_{i=1}^{n}S_i$，$S_i=\{z\in C:|z-a_{ii}|<a_{ii}\}$（$i=1,2,\dots,n$）
2. 若A的所有主对角元均为正数，则A的所有特征值都有正实部
3. 若A为 Hermite 阵，且A的所有主对角元均为正数，则A的所有特征值都是正数

证：

1. 盖尔圆圆心大于半径，0 不在任何一个盖尔圆内，故可逆。$z\in C:|z-a_{ii}|\le R_i<a_{ii}$
2. A的所有行盖尔圆圆心都在实轴，所有盖尔圆都在虚轴右侧，所以特征值的实部均为正
3. A为 Hermite 阵，则A的特征值全为实数。由 2 知全为正数。

定理 5（利用盖尔圆估计谱范数）：

1. $r(A) \le \min\left\{\max\limits_i \sum_{j=1}^n |a_{ij}|, \max\limits_j \sum_{i=1}^n |a_{ij}|\right\}$

2. 放缩盖尔圆，当取得适当的$\{p_i\}$时，能使得盖尔圆的范围最小。

补充：谱半径不是范数。

4. Hermite 矩阵特征值的变分特征

（用解析函数的方法求特征值）

定义：Rayleigh 商（瑞利商）：

对 Hermite 矩阵A，有$R(x) = \frac{x^H Ax}{x^H x}$，$x \ne 0$（$= y^H Ay$，$\|y\| = 1$）

定理 1（Rayleigh-Ritz）：

将 Hermite 矩阵A的特征值降序排列（$\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_n$），有

1. $\lambda_n x^H x \le x^H Ax \le \lambda_1 x^H x$（对于任意$x$）

2. $\begin{cases} \lambda_1 = \max\limits_{x \ne 0} R(x) = \max\limits_{\|x\|=1} x^H Ax \\ \lambda_n = \min\limits_{x \ne 0} R(x) = \min\limits_{\|x\|=1} x^H Ax \end{cases}$

证：

1. $x^H Ax = x^H U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^H x$（令$y = U^H x$）$= y^H \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} y = \lambda_1 |y_1|^2 + \ldots + \lambda_n |y_n|^2$

放大有$x^H Ax \le \lambda_1 y^H y = \lambda_1 x^H x$，缩小有$x^H Ax \ge \lambda_n y^H y = \lambda_n x^H x$。

2. 由 1 知$R(x) \le \lambda_1$，当取$x = x_1$为λ_1对应的特征向量时，即$Ax_1 = \lambda_1 x_1$，此时有$x_1^H Ax_1 = \lambda_1 x_1^H x_1$，即$R(x_1) = \lambda_1$

所以$\lambda_1 = R(x_1) = \max\limits_{x \ne 0} R(x) = \max\limits_{\|x\|=1} x^H Ax$（取$\|x_1\| = 1$）。$\lambda_n$同理。

推论 1：

若$a = R(x) = \frac{x^H Ax}{x^H x}$，则至少有一个特征值$\lambda \le a$，也至少有一个特征值$\lambda \ge a$。

推论 2：

设p_i为对应于λ_i的单位正交特征向量， 则p_1、p_n分别为$R(x)$的极大点和极小点，即$R(p_1) = \lambda_1$，$R(p_n) = \lambda_n$

定理 2（Courant-Fischer）：

将 Hermite 矩阵A的特征值升序排列（$\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$），给定 $1 \le k \le n$，有

1. $\min\limits_{\omega_1,\omega_2,\ldots,\omega_{n-k} \in C^n} \max\limits_{\substack{x \ne 0, x \in C^n \\ x \perp \omega_1,\omega_2,\ldots,\omega_{n-k}}} R(x) = \lambda_k$（有$n-k$个约束，求最大值，能得到第$k$小特征值）

2. $\max\limits_{\omega_1,\omega_2,\ldots,\omega_{k-1} \in C^n} \min\limits_{\substack{x \ne 0, x \in C^n \\ x \perp \omega_1,\omega_2,\ldots,\omega_{k-1}}} R(x) = \lambda_k$（有$k-1$ 个约束，求最小值，能得到第k小特征值）

证：

1. （先证明λ_k是$\max\limits_{x \perp \omega_1,\omega_2,\ldots,\omega_{n-k}} \frac{x^H Ax}{x^H x}$的下界）$x \perp \omega_1, \omega_2, \ldots, \omega_{n-k}$，即$x \perp L(\omega_1, \omega_2, \ldots, \omega_{n-k})$（生成子空间），即$x \in L^\perp$

A是 Hermite 矩阵，则$A = U\Lambda U^H$，则$\max\limits_{x \perp \omega_1,\omega_2,\ldots,\omega_{n-k}} \frac{x^H Ax}{x^H x} = \max\limits_{x \perp \omega_1,\omega_2,\ldots,\omega_{n-k}} \frac{x^H U\Lambda U^H x}{x^H UU^H x}$（酉变换$y = U^H x$）$= \max\limits_{y \perp U^H\omega_1, U^H\omega_2,\ldots,U^H\omega_{n-k}} \frac{y^H \Lambda y}{y^H y}$

（$x \perp \omega_1 \Leftrightarrow 0 = (x, \omega_1) = x^H \omega_1$（$y = U^H x$，即$x = Uy$）$= (Uy)^H \omega_1 = y^H U^H \omega_1 = (y, U^H \omega_1) = 0 \Leftrightarrow y \perp U^H \omega_1$）

（单位化$z = \frac{y}{\|y\|_2}$）$= \max\limits_{\substack{\|z\|_2 = 1 \\ z \perp U^H\omega_1, U^H\omega_2,\ldots,U^H\omega_{n-k}}} z^H \Lambda z$（标准型）$= \max\limits_{\substack{\|z\|_2 = 1 \\ z \perp U^H\omega_1, U^H\omega_2,\ldots,U^H\omega_{n-k}}} \lambda_1 |z_1|^2 + \ldots + \lambda_n |z_n|^2$

（加条件：$z_1 = \ldots = z_{k-1} = 0$）$\ge \max\limits_{\substack{\|z\|_2 = 1 \\ z \perp U^H\omega_1, U^H\omega_2,\ldots,U^H\omega_{n-k} \\ z_1 = \ldots = z_{k-1} = 0}} \lambda_k |z_k|^2 + \ldots + \lambda_n |z_n|^2$（$\lambda_i$都缩小为$\lambda_k$）$\ge \max\limits_{\substack{\|z\|_2 = 1 \\ z \perp U^H\omega_1, U^H\omega_2,\ldots,U^H\omega_{n-k}}} \lambda_k \|z\|_2^2 = \lambda_k$

即$\max\limits_{x \perp \omega_1,\omega_2,\ldots,\omega_{n-k}} \frac{x^H Ax}{x^H x} \ge \lambda_k$。

（2 再证明存在一系列$\omega_1,\omega_2,\dots,\omega_{n-k}$使$\max\limits_{x\perp\omega_1,\omega_2,\dots,\omega_{n-k}}\frac{x^HAx}{x^Hx}=\lambda_k$）令$U=(u_1,\dots,u_n)$

令$\omega_i=u_{n-i+1}$，即$\omega_1=u_n,\omega_2=u_{n-1},\dots,\omega_{n-k}=u_{k+1}$。（取特殊值）

则$U^H\omega_1=U^Hu_n=\left(u_1^H,\dots,u_n^H\right)^Hu_n=(0,\dots,0,1)^H=\varepsilon_n$。且$z\perp\varepsilon_n$相当于$z_n=0$

则$\max\limits_{x\perp u_n,u_{n-1},\dots,u_{k+1}}R(x)=\max\limits_{\substack{\|z\|_2=1\\ z\perp U^Hu_n,U^Hu_{n-1},\dots,U^Hu_{k+1}}}\lambda_k\|z\|_2^2=\max\limits_{\substack{\|z\|_2=1\\ z_n=z_{n-1}=\dots=z_{k+1}=0}}\lambda_k\|z\|_2^2\le\lambda_k$

取$x=u_k$为λ_k对应的单位特征向量。（因为$Au_i=\lambda_iu_i$）

则$\max\limits_{\substack{\|z\|_2=1\\ x\perp u_n,u_{n-1},\dots,u_{k+1}}}R(x)\ge R(u_k)=\frac{u_k^HAu_k}{u_k^Hu_k}=\lambda_k$（$u_k\perp u_n,u_{n-1},\dots,u_{k+1}$，确实满足条件）

因为$\max\limits_{x\perp u_n,u_{n-1},\dots,u_{k+1}}R(x)\le\lambda_k$，又$\max\limits_{x\perp u_n,u_{n-1},\dots,u_{k+1}}R(x)\ge\lambda_k$，所以$\max\limits_{x\perp u_n,u_{n-1},\dots,u_{k+1}}R(x)=\lambda_k$

λ_k是$\max\limits_{x\perp\omega_1,\omega_2,\dots,\omega_{n-k}}\frac{x^HAx}{x^Hx}$的下界，下界具有可达性，因此$\min\limits_{\omega_1,\omega_2,\dots,\omega_{n-k}\in C^n}\max\limits_{\substack{x\neq0,x\in C^n\\ x\perp\omega_1,\omega_2,\dots,\omega_{n-k}}}R(x)=\lambda_k$。2.的证明同理。

用定理 2 算特征值：

1. 用$R(x)$算λ_n，u_n
2. 把$x\perp u_n$当做新增条件，算λ_{n-1}，u_{n-1}
3. 以此类推，直至全部算出

例：计算$A=\begin{pmatrix}5&0&0\\0&1&2\\0&2&1\end{pmatrix}$的特征值

0. $\begin{cases}\lambda_1=-1\\ \lambda_2=3\\ \lambda_3=5\end{cases}$

1. $\lambda_3=\max\limits_{x_1^2+x_2^2+x_3^2=1}5x_1^2+x_2^2+x_3^2+4x_2x_3=5$，$5E-A=\begin{pmatrix}0&0&0\\0&4&-2\\0&-2&4\end{pmatrix}$，$u_3=\begin{pmatrix}1\\0\\0\end{pmatrix}$

2. $\lambda_2=\max\limits_{\substack{x_1^2+x_2^2+x_3^2=1\\ x\perp u_3}}5x_1^2+x_2^2+x_3^2+4x_2x_3=\max\limits_{x_2^2+x_3^2=1}x_2^2+x_3^2+4x_2x_3=3$，$3E-A=\begin{pmatrix}-2&0&0\\0&2&-2\\0&-2&2\end{pmatrix}$，$u_2=\begin{pmatrix}0\\1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}$

3. $\lambda_1=\max\limits_{\substack{x_2^2+x_3^2=1\\ x\perp u_2}}x_2^2+x_3^2+4x_2x_3=\max\limits_{2x_3^2=1}-2x_3^2=-1$，$-E-A=\begin{pmatrix}-6&0&0\\0&-2&-2\\0&-2&-2\end{pmatrix}$，$u_2=\begin{pmatrix}0\\1/\sqrt{2}\\-1/\sqrt{2}\end{pmatrix}$

第五章 矩阵分析

1. 矩阵序列与矩阵级数

定义 1：矩阵序列收敛：

矩阵序列$\{A^{(k)}\}$，$A^{(k)}=\left(a_{ij}^{(k)}\right)_{m\times n}$，若$\lim\limits_{k\to\infty}a_{ij}^{(k)}=a_{ij}$，$i=1,2,\ldots,m$，$j=1,2,\ldots,n$，则称矩阵序列$\{A^{(k)}\}$收敛于$A=(a_{ij})$

称矩阵A为矩阵序列$\{A^{(k)}\}$的极限，记作$\lim\limits_{k\to+\infty}A^{(k)}=A$或$A^{(k)}\to A$（$k\to+\infty$）

定理 1（收敛序列的运算）：

1. $\lim\limits_{k\to+\infty}\left(aA^{(k)}+bB^{(k)}\right)=a\lim\limits_{k\to+\infty}A^{(k)}+b\lim\limits_{k\to+\infty}B^{(k)}$（线性）

2. $\lim\limits_{k\to+\infty}\left(A^{(k)}B^{(k)}\right)=\lim\limits_{k\to+\infty}A^{(k)}\lim\limits_{k\to+\infty}B^{(k)}$（乘的极限等于极限的乘）

3. 当$A^{(k)}$和$\lim\limits_{k\to+\infty}A^{(k)}$均可逆时，有$\lim\limits_{k\to+\infty}\left(A^{(k)}\right)^{-1}=\left(\lim\limits_{k\to+\infty}A^{(k)}\right)^{-1}$（逆的极限等于极限的逆）

定理 2（依范数收敛等价于逐点收敛）：

矩阵序列$\{A^{(k)}\}$收敛于矩阵A的充要条件是$\lim\limits_{k\to\infty}\left\|A^{(k)}-A\right\|=0$，$\|\cdot\|$是任意矩阵范数（依范数收敛）

定义 2：收敛矩阵：

A为n阶方阵，若$\lim\limits_{k\to\infty}A^k=O$，则称$A$为收敛矩阵

定理 3（收敛矩阵的充要条件）：

A为收敛矩阵的充要条件是谱半径$r(A)<1$。

证：

（充分性）作 Jordan 分解，有$A=PJP^{-1}$，J为 Jordan 标准型。于是有$A^k=PJ^kP^{-1}$，则A收敛等价于J收敛。

J收敛又等价于每个 Jordan 块$J_{r_i}(\lambda_i)$收敛，而$J_{r_i}^k(\lambda_i)=\begin{pmatrix}f_k(\lambda_i) & \cdots & \frac{f_k^{(r_i-1)}(\lambda_i)}{(r_i-1)!}\\ & \ddots & \vdots\\ & & f_k(\lambda_i)\end{pmatrix}$，其中$f_k(\lambda_i)=\lambda_i^k$。

所以当$|\lambda_i|<1$时有$f_k^{(l)}(\lambda_i)\to 0$，$l=0,1,\ldots,r_i-1$，故$J_{r_i}^k(\lambda_i)\to O$，$J^k\to O$，$A^k\to O$

（必要性）若$A^k\to O$，则$J_{r_i}^k(\lambda_i)\to O$，则$f_k(\lambda_i)=\lambda_i^k\to 0$，故$|\lambda_i|<1$

推论：

A是n阶方阵，对于任意$\varepsilon>0$，都存在一个与A和ε有关的常数c，使得$\left|\left(A^k\right)_{ij}\right|\le c[r(A)+\varepsilon]^k$。

（说明A^k的变化速度主要由谱半径$r(A)$决定）

定义 3：矩阵级数：

$\{A^{(k)}\}$是$m\times n$的矩阵序列，称无穷和$A^{(1)}+A^{(2)}+\ldots+A^{(k)}+\ldots$为矩阵级数，记为$\sum_{k=1}^{\infty}A^{(k)}$。

称$S^{(N)}=\sum_{k=1}^{N}A^{(k)}$为矩阵级数的部分和，如果$\{S^{(N)}\}$收敛于$S$，则称矩阵级数$\sum_{k=1}^{\infty}A^{(k)}$收敛于$S$，记为$\sum_{k=1}^{\infty}A^{(k)}=S$

（矩阵级数$\sum_{k=1}^{\infty}A^{(k)}$收敛的充要条件是$mn$个数项级数$\sum_{k=1}^{\infty}a_{ij}^{(k)}$都收敛）

定义 4：矩阵级数绝对收敛：

如果mn个数项级数$\sum_{k=1}^{\infty}a_{ij}^{(k)}$都绝对收敛，则称矩阵级数$\sum_{k=1}^{\infty}A^{(k)}$绝对收敛。

定理 4（矩阵级数绝对收敛等价于范数和级数收敛）：

矩阵级数$\sum_{k=1}^{\infty}A^{(k)}$绝对收敛的充要条件是正项级数$\sum_{k=1}^{\infty}\left\|A^{(k)}\right\|$收敛，$\|\cdot\|$是任意矩阵范数。

证：

（必要性）若$\sum_{k=1}^{\infty}A^{(k)}$绝对收敛，则$\sum_{k=1}^{\infty}a_{ij}^{(k)}$都绝对收敛，则对于任意$N$、$i$、$j$，都存在正数$M$，使得$\sum_{k=1}^{N}\left|a_{ij}^{(k)}\right|<M$

（即M是mn个数项级数$\sum_{k=1}^{\infty}\left|a_{ij}^{(k)}\right|$的最大上界）从而$\sum_{k=1}^{N}\left\|A^{(k)}\right\|_1=\sum_{k=1}^{N}\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}^{(k)}\right|\right)<mnM$，故$\sum_{k=1}^{\infty}\left\|A^{(k)}\right\|_1$收敛。

又由于任何矩阵范数都等价（第二章第 2 节定理 1），所以$\sum_{k=1}^{\infty}\left\|A^{(k)}\right\|\le C_1\sum_{k=1}^{\infty}\left\|A^{(k)}\right\|_1$，因此任意矩阵范数均收敛。

（充分性）若正项级数$\sum_{k=1}^{\infty}\left\|A^{(k)}\right\|$收敛，则$\sum_{k=1}^{\infty}\left\|A^{(k)}\right\|_1$收敛，则$\sum_{k=1}^{\infty}\left|a_{ij}^{(k)}\right|\le\sum_{k=1}^{\infty}\left\|A^{(k)}\right\|_1$也收敛，则$\sum_{k=1}^{\infty}A^{(k)}$绝对收敛。

定理 5（矩阵幂级数）：
A是n阶方阵，A的幂级数（Neumann 级数）$\sum_{k=0}^{\infty}A^k=E+A+A^2+\ldots+A^k+\ldots$收敛的充要条件是$A$为收敛矩阵。
此时幂级数的和$\sum_{k=0}^{\infty}A^k=S=(E-A)^{-1}$
证：

（必要性）$\sum_{k=0}^{\infty}A^k$收敛，当且仅当mn个数项级数$\sum_{k=0}^{\infty}\left(A^k\right)_{ij}$收敛，此时必有$\lim\limits_{k\to\infty}\left(A^k\right)_{ij}=0$，故$\lim\limits_{k\to\infty}A^k=O$，即$A$为收敛矩阵。

（充分性）A为收敛矩阵当且仅当$r(A)<1$，故$E-A$可逆。
又$\left(E+A+A^2+\ldots+A^k\right)(E-A)=E-A^{k+1}$，则$E+A+A^2+\ldots+A^k=\left(E-A^{k+1}\right)(E-A)^{-1}$
由于$k\to\infty$时，有$A^{k+1}\to O$，则有$E+A+A^2+\ldots+A^k=(E-A)^{-1}$。即$\sum_{k=0}^{\infty}A^k$收敛于$(E-A)^{-1}$
定理 6（矩阵幂级数“收敛半径”）：
幂级数$f(z)=\sum_{k=0}^{\infty}c_kz^k$的收敛半径为$r$，如果方阵$A$满足$r(A)<r$，则矩阵幂级数$\sum_{k=0}^{\infty}c_kA^k$绝对收敛；如果$r(A)>r$，则发散。
推论：
如果幂级数$f(z)=\sum_{k=0}^{\infty}c_kz^k$在整个复平面收敛，则$\sum_{k=0}^{\infty}c_kA^k$总是绝对收敛。

2. 矩阵函数

一、矩阵函数的定义

定义：矩阵函数：
设幂级数$\sum_{k=0}^{\infty}a_kz^k$的收敛半径为$r$，且当$|z|<r$时，幂级数收敛于函数$f(z)$，即$f(z)=\sum_{k=0}^{\infty}a_kz^k$，$|z|<r$。
如果n阶方阵A满足$r(A)<r$，则称收敛的矩阵幂级数$\sum_{k=0}^{\infty}a_kA^k$的和为矩阵函数，记为$f(A)$，即$f(A)=\sum_{k=0}^{\infty}a_kA^k$。
（矩阵函数是用对应的泰勒级数定义的）
常用公式：

$e^A=E+A+\frac{A^2}{2!}+\ldots+\frac{1}{k!}A^k+\ldots=\sum_{k=0}^{\infty}\frac{1}{k!}A^k$，$A$是任意方阵

$\sin A=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\frac{1}{7!}A^7+\ldots+\frac{(-1)^k}{(2k+1)!}A^{2k+1}+\ldots=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}A^{2k+1}$，$A$是任意方阵

$\cos A=E-\frac{1}{2!}A^2+\frac{1}{4!}A^4-\frac{1}{6!}A^6+\ldots+\frac{(-1)^k}{(2k)!}A^{2k}+\ldots=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}A^{2k}$，$A$是任意方阵

$(E-A)^{-1}=E+A+A^2+\ldots+A^k+\ldots=\sum_{k=0}^{\infty}A^k$，$r(A)<1$

$\ln(E+A)=A-\frac{1}{2}A^2+\frac{1}{3}A^3-\frac{1}{4}A^4+\ldots+\frac{(-1)^k}{k+1}A^{k+1}=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}A^{2k}$，$r(A)<1$

$\arctan A=A-\frac{1}{3}A^3+\frac{1}{5}A^5-\frac{1}{7}A^7+\ldots+\frac{(-1)^k}{2k+1}A^{2k+1}+\ldots=\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}A^{2k+1}$，$r(A)<1$

引入参数t，把变量A换成At，有$f(At)=\sum_{k=0}^{\infty}a_k(At)^k$。

二、矩阵函数值的计算

1. 利用相似对角化

若A相似于对角矩阵D，即$P^{-1}AP=D=\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}$，则$f(A)=\sum_{k=0}^{\infty}a_kA^k=\sum_{k=0}^{\infty}a_k(PDP^{-1})^k=P\left(\sum_{k=0}^{\infty}a_kD^k\right)P^{-1}$

$=P\begin{pmatrix}\sum_{k=0}^{\infty}a_k\lambda_1^k&&\\&\ddots&\\&&\sum_{k=0}^{\infty}a_k\lambda_n^k\end{pmatrix}P^{-1}=P\begin{pmatrix}f(\lambda_1)&&\\&\ddots&\\&&f(\lambda_n)\end{pmatrix}P^{-1}$，同理$f(At)=P\begin{pmatrix}f(\lambda_1t)&&\\&\ddots&\\&&f(\lambda_nt)\end{pmatrix}P^{-1}$

求特征根的时候可以用这个定理：

$f(x)=a_nx^n+\ldots+a_1x+a_0$是整数系数多项式，如果$\frac{s}{r}$是$f(x)=0$ 的有理根，则r能整除a_n，s能整除a_0

2. Jordan 标准形法

化 Jordan 标准型，$P^{-1}AP=J=\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix}$，$J_i=\begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{pmatrix}_{r_i\times r_i}$，

$$f(J_i)=\sum_{k=0}^{\infty}a_kJ_i^k=\sum_{k=0}^{\infty}a_k\begin{pmatrix} \lambda_i^k & C_k^1\lambda_i^{k-1} & \ldots & C_k^{r_i-1}\lambda_i^{k-r_i+1} \\ & \lambda_i^k & & \vdots \\ & & \ddots & C_k^1\lambda_i^{k-1} \\ & & & \lambda_i^k \end{pmatrix}=\begin{pmatrix} f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \ldots & \frac{1}{(r_i-1)!}f^{(r_i-1)}(\lambda_i) \\ & f(\lambda_i) & & \vdots \\ & & \ddots & \frac{1}{1!}f'(\lambda_i) \\ & & & f(\lambda_i) \end{pmatrix},$$

则$f(A)=\sum_{k=0}^{\infty}a_k(PJP^{-1})^k=P\left(\sum_{k=0}^{\infty}a_kJ^k\right)P^{-1}=P\begin{pmatrix} \sum_{k=0}^{\infty}a_kJ_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^{\infty}a_kJ_s^k \end{pmatrix}P^{-1}=P\begin{pmatrix} f(J_1) & & \\ & \ddots & \\ & & f(J_s) \end{pmatrix}P^{-1}$

3. 数项级数求和法

因 Hamilton-Caylay 定理（见第三章第 2 节定理 3 下方）

$f(\lambda)=|\lambda A-E|=\lambda^n+a_1\lambda^{n-1}+\ldots+a_{n-1}\lambda+a_n$，则$f(A)=A^n+a_1A^{n-1}+\ldots+a_{n-1}A+a_nE=0$

故总存在$\Psi(\lambda)=\lambda^m+b_1\lambda^{m-1}+\ldots+b_{m-1}\lambda+b_m$（$1\le m\le n$），使得$\Psi(A)=A^m+b_1A^{m-1}+\ldots+b_{m-1}A+b_mE=0$

或者写成$A^m=-(b_1A^{m-1}+\ldots+b_{m-1}A+b_mE)$，即$A^m$是$A^{m-1},A^{m-2},\ldots,A^2,A,E$的线性组合

即任意A^{m+l}都是$A^{m-1},A^{m-2},\ldots,A^2,A$的线性组合：$\begin{cases} A^{m+1}=k_0^{(1)}E+k_1^{(1)}A+\ldots+k_{m-1}^{(1)}A^{m-1} \\ \vdots \\ A^{m+l}=k_0^{(l)}E+k_1^{(l)}A+\ldots+k_{m-1}^{(l)}A^{m-1} \\ \vdots \end{cases}$

于是有$f(A)=\sum_{k=0}^{\infty}c_kA^k$

$=(c_0E+c_1A+\ldots+c_{m-1}A^{m-1})+c_m\left(k_0^{(1)}E+k_1^{(1)}A+\ldots+k_{m-1}^{(1)}A^{m-1}\right)+\ldots+c_{m+l}\left(k_0^{(l)}E+k_1^{(l)}A+\ldots+k_{m-1}^{(l)}A^{m-1}\right)+\ldots$

$=\left(c_0+\sum_{l=0}^{\infty}c_{m+l}k_0^{(l)}\right)E+\left(c_1+\sum_{l=0}^{\infty}c_{m+l}k_1^{(l)}\right)A+\ldots+\left(c_{m-1}+\sum_{l=0}^{\infty}c_{m+l}k_{m-1}^{(l)}\right)A^{m-1}$ （共m项）

可见，矩阵幂级数求和可转化为m（不超过n，n为阶数）个数项级数求和问题。

三、矩阵函数的一些性质

定理 1：

若$AB=BA$，则$e^Ae^B=e^Be^A=e^{A+B}$

证：

$e^Ae^B=\left(E+A+\frac{1}{2!}A^2+\ldots+\frac{1}{k!}A^k+\ldots\right)\left(E+B+\frac{1}{2!}B^2+\ldots+\frac{1}{k!}B^k+\ldots\right)$

$=\left(E+B+\frac{1}{2!}B^2+\ldots\right)+\left(A+AB+\frac{1}{2!}AB^2+\ldots\right)+\left(\frac{1}{2!}A^2+\frac{1}{2!}A^2B+\frac{1}{2!}\cdot\frac{1}{2!}A^2B^2+\ldots\right)+\ldots$

$=E+(A+B)+\left(\frac{1}{2!}A^2+AB+\frac{1}{2!}B^2\right)+\ldots=e^{A+B}$

推论 1：

$e^Ae^{-A}=e^{-A}e^A=E$，$(e^A)^{-1}=e^{-A}$

推论 2：

m为整数，则$(e^A)^m=e^{mA}$

定理 2：

若$AB=BA$，则$\begin{cases} \cos(A+B)=\cos A\cos B-\sin A\sin B \\ \sin(A+B)=\sin A\cos B+\cos A\sin B \end{cases}$

第六章 广义逆矩阵

1. 矩阵的单边逆

定义 1：左（右）逆：

A为$m\times n$矩阵，如果$GA=E_n$（$AG=E_m$），则称G是A的左（右）逆矩阵，记为$G=A_L^{-1}$（$G=A_R^{-1}$）。

如果A有左（右）逆，则称A是左（右）可逆的。

（可逆矩阵A的逆A^{-1}是单边逆的特殊情况，且$A^{-1}=A_L^{-1}=A_R^{-1}$）

定理 1（左（右）可逆的充要条件）：

左（右）可逆的充要条件是列（行）满秩。

证：

（充分性）A列满秩，则由第三章第 5 节定理 1 的 1. 有$\operatorname{rank}A^HA=\operatorname{rank}A=n$，即$A^HA$可逆。令$G=(A^HA)^{-1}A^H$，有$GA=E_n$。

（必要性）$GA=E_n$，则$\operatorname{rank}A\geq\operatorname{rank}E_n=n$。又$\operatorname{rank}A\leq n$，故$\operatorname{rank}A=n$，即$A$列满秩。

右可逆行满秩取$G=A^H(AA^H)^{-1}$，证法同理。

推论 1：

1. A左可逆的充要条件是$N(A)=\{0\}$
2. A右可逆的充要条件是$R(A)=C^m$

$N(A)=\{0\}$即$Ax=0$ 只有 0 解，其充要条件是$\operatorname{rank}A=n$，即A列满秩。

$R(A)=C^m$即$y=Ax$，y所在的线性子空间为整个C^m空间，其充要条件是$\operatorname{rank}A=m$，即A行满秩。

初等变换求左（右）逆：

A为$m\times n$列满秩矩阵，补充为$(A\quad E_m)$，行变换为$\begin{pmatrix}E_n & G\\ O & *\end{pmatrix}$，则$G$就是$A$的一个左逆矩阵。

A为$m\times n$行满秩矩阵，补充为$\begin{pmatrix}A\\ E_n\end{pmatrix}$，行变换为$\begin{pmatrix}E_m & O\\ G & *\end{pmatrix}$，则$G$就是$A$的一个右逆矩阵。

定理 2（左逆的通解）：

A左可逆，则$G=(A_1^{-1}-BA_2A_1^{-1}\quad B)P$是$A$的一个左逆矩阵，其中$n\times(m-n)$矩阵$B$为任意矩阵

P为置换阵，使得$PA=\begin{pmatrix}A_1\\ A_2\end{pmatrix}$，其中$A_1$是$n$阶可逆阵。

定理 3（右逆的通解）：

A右可逆，则$G=Q\begin{pmatrix}A_1^{-1}-A_1^{-1}A_2D\\ D\end{pmatrix}$是$A$的一个右逆矩阵，其中$(n-m)\times m$矩阵$D$为任意矩阵

Q为置换阵，使得$AQ=(A_1\quad A_2)$，其中A_1是m阶可逆阵。

定理 4（线性方程组有唯一解）：

A左可逆，A_L^{-1}是A的一个左逆，则方程组$Ax=b$有解的充要条件是$\left(E_m-AA_L^{-1}\right)b=0$，若成立则有唯一解$x=(A^HA)^{-1}A^Hb$

证：

（必要性）$Ax=b$有解，设x_0是$Ax=b$的解，则$Ax_0=b$，则$AA_L^{-1}Ax_0=AA_L^{-1}b=A\left(A_L^{-1}A\right)x_0=Ax_0=b$，则$\left(E_m-AA_L^{-1}\right)b=0$

（充分性）令$x_0=A_L^{-1}b$，因为$\left(E_m-AA_L^{-1}\right)b=0$ 即$AA_L^{-1}b=b$，则$Ax_0=AA_L^{-1}b=b$，则x_0是$Ax=b$的解。

（存在性）$(A^HA)^{-1}A^HA=E_n$，所以$(A^HA)^{-1}A^H$是A的一个左逆，所以$x=(A^HA)^{-1}A^Hb$的$Ax=b$的解。

（唯一性）设x_0、x_1都是$Ax=b$的解，则$Ax_0=b$，$Ax_1=b$，则$A(x_0-x_1)=0$，由推论 1.1 知$x_0-x_1=0$。

定理 4（线性方程组有无穷多解解）：

A右可逆，则$Ax=b$对任意b都有解。若$b\neq0$，则方程组的解可表示为$x=A_R^{-1}b$，其中A_R^{-1}是A的一个右逆矩阵。

证：

由推论 1.2 知$Ax=b$对任意b都有解。$Ax=b$左乘A_R^{-1}，有$A_R^{-1}Ax=A_R^{-1}b$，即$x=A_R^{-1}b$。

$A^H(AA^H)^{-1}$是A的一个右逆，故$x=A^H(AA^H)^{-1}b$是$Ax=b$的一个解。

2. 广义逆矩阵A^-

定义 1：广义逆矩阵：

A为$m\times n$矩阵，若存在$n\times m$矩阵G，使得对于$\forall b\in R(A)$，都有$AGb=b$，则称G是A的广义逆矩阵，记作$G=A^-$。

（任何矩阵都有广义逆）

定理 1（广义逆的充要条件，或另一种定义）：

矩阵G是A的广义逆的充要条件是$AGA = A$。

满足 Penrose 第一个矩阵方程的G的集合。$A\{1\} = \{G|AGA = A, \forall G \in C^{n\times m}\}$

证：

（必要性）G是A的广义逆，则$AGb = b$，$\forall b \in R(A)$。又$A = (\alpha_1, \alpha_2, \dots, \alpha_n)$，$\alpha_i \in R(A)$，则$AG\alpha_i = \alpha_i$，因此$AGA = A$。

（充分性）$\forall b \in R(A)$，取$x \in C^n$使$Ax = b$，则$AGb = AGAx$。又$AGA = A$，则$AGb = AGAx = Ax = b$，因此$AGb = b$。

推论 1：

$\operatorname{rank} A^- \geq \operatorname{rank} A$

证：$AA^-A = A$，因此$\operatorname{rank} A^- \geq \operatorname{rank} A$。

定理 3（A^-的性质）：

1. $(A^-)^H = (A^H)^-$，$(A^-)^T = (A^T)^-$（可交换）
2. AA^-、A^-A均为幂等矩阵，且$\operatorname{rank} A = \operatorname{rank} AA^- = \operatorname{rank} A^-A$
3. λ^-A^-是λA的广义逆，其中$\lambda^- = \begin{cases} 0 & ,\lambda = 0 \\ \lambda^{-1} & ,\lambda \neq 0 \end{cases}$（实际上$\lambda = 0$ 时λ^-可取任意值，即O^-是任意矩阵）
4. $B = SAT$，其中S、T分别为m、n阶可逆阵，则$B^- = T^{-1}A^-S^{-1}$（广义逆$(AB)^- \neq B^-A^-$）
5. $R(AA^-) = R(A)$，$N(A^-A) = N(A)$
6. 若$ABA = A$，$(AB)^H = AB$，则$AB = P_{R(A)}$（AB是正交投影矩阵，且其值域为$R(A)$）

证：

1. $AA^-A = A$，则$A^H(A^-)^HA^H = A^H$，所以$(A^H)^- = (A^-)^H$。$(A^-)^T$同理。
2. 幂等：$(AA^-)^2 = AA^-AA^- = (AA^-A)A^- = AA^-$。

显然有$\operatorname{rank} A \geq \operatorname{rank} AA^-$，又$\operatorname{rank} A = \operatorname{rank} AA^-A \leq \operatorname{rank} AA^-$，则$\operatorname{rank} A = \operatorname{rank} AA^-$。$A^-A$同理。

3. $\lambda = 0$ 时，$(\lambda A)(\lambda^-A^-)(\lambda A) = O = \lambda A$。$\lambda \neq 0$ 时，$(\lambda A)(\lambda^-A^-)(\lambda A) = \lambda\frac{1}{\lambda}\lambda AA^-A = \lambda A$。
4. $(SAT)(T^{-1}A^-S^{-1})(SAT) = SAA^-AT = SAT$，故$(SAT)^- = T^{-1}A^-S^{-1}$
5. $R(AA^-) = \{y|y = AA^-x, \forall x \in C^n\}$，　$R(A) = \{y|y = Ax, \forall x \in C^n\}$。

显然$y = AA^-x = A(A^-x)$，则$R(A) \supset R(AA^-)$。又由 2 知$\operatorname{rank} A = \operatorname{rank} AA^-$，故$R(A) = R(AA^-)$。

$N(A^-A) = \{x|A^-Ax = 0\}$，$N(A) = \{x|Ax = 0\}$。

显然若$Ax = 0$，必有$A^-Ax = 0$。则$N(A^-A) \supset N(A)$。又$\operatorname{rank} A = \operatorname{rank} A^-A$，故$N(A^-A) = N(A)$。

6. 由 2 知$(AB)^2 = AB$，又$(AB)^H = AB$，则AB是正交投影矩阵，$AB = P_{R(AB)}$。又由 5 知$R(AB) = R(A)$，则$AB = P_{R(A)}$。

3. 自反广义逆矩阵A_r^-

定义 1：自反广义逆：

A为$m \times n$矩阵，若存在$n \times m$矩阵G，使得$AGA = A$且$GAG = G$（$A\{1,2\}$），则G是A的自反广义逆矩阵，记为$G = A_r^-$。

显然自反广义逆A_r^-是特殊的广义逆A^-。

满足$(A_r^-)_r^- = A$（自反性），有$A = G_r^-$。

列酉型矩阵的共轭转置是自反广义逆，即满足$A^H = A_r^-$

行满秩矩阵的右逆也是自反广义逆。

定理 1（存在性）：

任何矩阵都有自反广义逆矩阵。

证：

任取A为$m \times n$矩阵。如果$A = O$，则$A_r^- = O$。

如果$A \neq O$，设$\operatorname{rank} A = r$，则存在$A = P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q$，其中$P$、$Q$分别为$m$、$n$阶可逆阵。

则可以验证$G = Q^{-1}\begin{pmatrix} E_r & X \\ Y & YX \end{pmatrix}P^{-1} = A_r^-$。

A_r^-求法：

先通过行变换，把$(A \quad E_m)$化为行最简R，即$P(A \quad E_m) = (PA \quad P) = R$。

再通过列变换，把$\begin{pmatrix} R \\ E_{m+n} \end{pmatrix}$化为$\begin{pmatrix} E_r & O \\ O & O \\ & Q \end{pmatrix}$，即$\begin{pmatrix} R \\ E_{m+n} \end{pmatrix}Q = \begin{pmatrix} RQ \\ Q \end{pmatrix} = \begin{pmatrix} E_r & O \\ O & O \\ & Q \end{pmatrix}$

因此$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n} = PAQ$（注意：此处的$P$、$Q$和上方证明中的$P$、$Q$不是一回事）。

因此$Q^{-1}A_r^-P^{-1} = (PAQ)_r^- = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$，即$A_r^- = Q\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}P$

定理 2：

X、$Y = A^-$，则$XAY = A_r^-$。

证：因为$AXA = A$，$AYA = A$，有$(XAY)A(XAY) = XAYAXAY = X(AYA)XAY = XAXAY = X(AXA)Y = XAY$。

且$A(XAY)A = AXAYA = (AXA)YA = AYA = A$，因此$XAY = A_r^-$。

（特别地，如果取$Z = A^-$，则$ZAZ = A_r^-$）。

定理 3（A^-是A_r^-的条件）：

设A^-是A的广义逆，则A^-是A的自反广义逆的充要条件是$\operatorname{rank} A = \operatorname{rank} A^-$。

证：

（必要性）A^-是A的自反广义逆，即A也是A^-的广义逆。由第 2 节定理 3.2，$\operatorname{rank} A = \operatorname{rank} A^-A = \operatorname{rank} A^-$。

（充分性）A^-是A的广义逆，即$AA^-A = A$。

显然有$\operatorname{rank} A^-A \le \operatorname{rank} A^-$。又$\operatorname{rank} A^- = \operatorname{rank} A = \operatorname{rank} AA^-A \le \operatorname{rank} A^-A$，故$\operatorname{rank} A^-A = \operatorname{rank} A^-$。

显然有$R(A^-) \supset R(A^-A)$，又$\operatorname{rank} A^-A = \operatorname{rank} A^-$，则$R(A^-) = R(A^-A)$。

因此存在$n \times m$矩阵X，使得$A^- = A^-AX$（A^-和A^-A值域相同，其列向量组等价，可以线性表出）

故$A = AA^-A = A(A^-AX)A = (AA^-A)XA = AXA$，故$X$也是$A$的一个广义逆。由定理 2，有$A^-AX = A^-$是$A$的自反广义逆。

5. M-P 广义逆矩阵A^+

定义 1：M-P 广义逆矩阵：

A为$m \times n$矩阵，若存在$n \times m$矩阵G，使得$\begin{cases} AGA = A \\ GAG = G \\ (GA)^H = GA \\ (AG)^H = AG \end{cases}$（$A\{1,2,3,4\} = \{A^+\}$），则称$G$是$A$的 M-P 广义逆矩阵，记为$G = A^+$。

定理 1：

A为$m \times n$矩阵，秩为r，最大秩分解为$A = BD$，则$G = D^H(DD^H)^{-1}(B^HB)^{-1}B^H$就是$A$的 M-P 广义逆矩阵$A^+$。

证：

若$r = 0$ 即$A = O$，则$G = O$，满足四个条件。

若$r \ne 0$，有B为$m \times r$列满秩矩阵，D为$r \times n$行满秩矩阵。

1. $AGA = (BD)D^H(DD^H)^{-1}(B^HB)^{-1}B^H(BD) = B(DD^H)(DD^H)^{-1}(B^HB)^{-1}(B^HB)D = BD = A$
2. $GAG = D^H(DD^H)^{-1}(B^HB)^{-1}B^H(BD)D^H(DD^H)^{-1}(B^HB)^{-1}B^H$
$= D^H(DD^H)^{-1}(B^HB)^{-1}(B^HB)(DD^H)(DD^H)^{-1}(B^HB)^{-1}B^H = D^H(DD^H)^{-1}(B^HB)^{-1}B^H = G$
3. $(AG)^H = \left((BD)D^H(DD^H)^{-1}(B^HB)^{-1}B^H\right)^H = \left(B(DD^H)(DD^H)^{-1}(B^HB)^{-1}B^H\right)^H = \left(B(B^HB)^{-1}B^H\right)^H = (B^H)^H\left((B^HB)^{-1}\right)^H B^H$
$= B\left((B^HB)^H\right)^{-1}B^H = B(B^HB)^{-1}B^H = AG$
4. $(GA)^H = \left(D^H(DD^H)^{-1}(B^HB)^{-1}B^H(BD)\right)^H = \left(D^H(DD^H)^{-1}(B^HB)^{-1}(B^HB)D\right)^H = \left(D^H(DD^H)^{-1}D\right)^H = D^H\left((DD^H)^{-1}\right)^H(D^H)^H =$
$D^H\left((DD^H)^H\right)^{-1}D = D^H(DD^H)^{-1}D = GA$

定理 2（唯一性）：

A^+是唯一的。

定理 3（A^+的性质）：

1. $(A^+)^+ = A$（自反性）
2. $(A^T)^+ = (A^+)^T$，　$(A^H)^+ = (A^+)^H$（可交换顺序）
3. $A^+ = (A^HA)^+A^H = A^H(AA^H)^+$
4. $R(A^+) = R(A^H)$
5. $AA^+ = P_{R(A)}$，$A^+A = P_{R(A^H)}$

6. $R(A)=R(A^H)$的充要条件是$AA^+=A^+A$

证：

1. 由定义知A^+是一个A_r^-。
2. 第 2 节定理 3.1，有$(A^T)^-=(A^-)^T$，$(A^H)^-=(A^-)^H$。A^+是一个A^-，故也可交换。
3. $A=BD$，则$A^HA=(BD)^HBD=D^HB^HBD$

因为$\operatorname{rank}A^HA=\operatorname{rank}A=r$，且$D$为$r\times n$行满秩矩阵，$D^H$为$n\times r$列满秩矩阵，$B^HB$为$r$阶可逆阵

因此$(D^HB^HB)D$是A^HA的一个满秩分解

故$(A^HA)^+A^H=D^H(DD^H)^{-1}\left((D^HB^HB)^H(D^HB^HB)\right)^{-1}(D^HB^HB)^H(BD)^H=D^H(DD^H)^{-1}(B^HBDD^HB^HB)^{-1}B^HBDD^HB^H$

$=D^H(DD^H)^{-1}\left((B^HB)(DD^H)(B^HB)\right)^{-1}(B^HB)(DD^H)B^H=D^H(DD^H)^{-1}(B^HB)^{-1}(DD^H)^{-1}(B^HB)^{-1}(B^HB)(DD^H)B^H$

$=D^H(DD^H)^{-1}(B^HB)^{-1}B^H=A^+$。$A^H(AA^H)^+$同理。

4. 由 3 知$R(A^+)=R\left(A^H(AA^H)^+\right)\subset R(A^H)$，又$\operatorname{rank}A^+=\operatorname{rank}A=\operatorname{rank}A^H$，故$R(A^+)=R(A^H)$
5. A^+是一个A_r^-，故$AA^+=P_{R(A)}$。$A^+A=P_{R(A^+)}=P_{R(A^H)}$
6. （必要性）若$R(A)=R(A^H)$，则$AA^+=P_{R(A)}=P_{R(A^H)}=A^+A$。

（充分性）由第 2 节定理 3.5，$R(A)=R(AA^+)$，$R(A^+)=R(A^+A)$。又$AA^+=A^+A$，则$R(A)=R(A^+)=R(A^H)$。

定理 5（$(AB)^+=B^+A^+$的条件）：

$(AB)^+=B^+A^+$的充要条件是$\begin{cases}R(A^HAB)\subset R(B)\\R(BB^HA^H)\subset R(A^H)\end{cases}$

6. A^+的计算方法

一、最大秩分解法

引理 1（行（列）满秩矩阵的A^+）：

1. 如果A是行满秩矩阵，则$A^+=A^H(AA^H)^{-1}$
2. 如果A是列满秩矩阵，则$A^+=(A^HA)^{-1}A^H$

定理 1（最大秩分解法）：

$A=BD$是A的最大秩分解，则$A^+=D^+B^+=D^H(DD^H)^{-1}(B^HB)^{-1}B^H$

二、奇异值分解法

定理 2（奇异值分解法）：

A奇异值分解为$A=U\begin{pmatrix}D_r&O\\O&O\end{pmatrix}V=UDV$，$U$、$V$为酉矩阵，$D_r=\begin{pmatrix}\sigma_1&&\\&\ddots&\\&&\sigma_r\end{pmatrix}$，$\sigma_i$是$A$的正奇异值，则有

1. $A^+=V^HD^+U^H$，$D^+=\begin{pmatrix}D_r^{-1}&O\\O&O\end{pmatrix}$
2. $\|A^+\|_F^2=\sum_{i=1}^r\frac{1}{\sigma_i^2}$
3. $\|A^+\|_2=\frac{1}{\min\limits_{1\le i\le r}\{\sigma_i\}}$

证：1. 因为$DD^+=\begin{pmatrix}E_r&O\\O&O\end{pmatrix}_{m\times m}$，$D^+D=\begin{pmatrix}E_r&O\\O&O\end{pmatrix}_{n\times n}$，故有

$$\begin{cases}AA^+A=(UDV)(V^HD^+U^H)(UDV)=UD(VV^H)D^+(U^HU)DV=UDD^+DV=UDV=A\\A^+AA^+=(V^HD^+U^H)(UDV)(V^HD^+U^H)=V^HD^+(U^HU)D(VV^H)D^+U^H=V^HD^+DD^+U^H=V^HD^+U^H=A^+\\(AA^+)^H=\left((UDV)(V^HD^+U^H)\right)^H=\left(UD(VV^H)D^+U^H\right)^H=(UDD^+U^H)^H=UDD^+U^H=AA^+\\(A^+A)^H=\left((V^HD^+U^H)(UDV)\right)^H=\left(V^HD^+(U^HU)DV\right)^H=(V^HD^+DV)^H=V^HD^+DV=A^+A\end{cases}$$

2. $\|A^+\|_F^2=\|D^+\|_F^2=\sum_{i=1}^r\frac{1}{\sigma_i^2}$（矩阵 2 范数酉不变）

3. $\|A^+\|_2^2=\|D^+\|_2^2=r[D^+(D^+)^H]=\max\limits_{1\le i\le r}\left\{\frac{1}{\sigma_i^2}\right\}=\frac{1}{\min\limits_{1\le i\le r}\sigma_i^2}$（算子 2 范数酉不变）

7. 广义逆矩阵的应用

一、矩阵方程的通解

定理 1（矩阵方程有解的条件和通解）：

矩阵方程$AXB=D$有解的充要条件是存在A^-和B^-，使得$AA^-DB^-B=D$成立。

如果有解，通解为$X=A^-DB^-+Y-A^-AYBB^-$，其中Y为任意矩阵。

证：

（充分性）若$AA^-DB^-B=D$，则$X=A^-DB^-$就是矩阵方程的一个解。

（必要性）设$AXB=D$。因为$AA^-A=A$，$BB^-B=B$，则$D=AXB=(AA^-A)X(BB^-B)=AA^-(AXB)B^-B=AA^-DB^-B$

（通解是解）$AXB=A(A^-DB^-+Y-A^-AYBB^-)B=AA^-DB^-B+AYB-(AA^-A)Y(BB^-B)=AA^-DB^-B=D$

（解是通解）设$AGB=D$，其中G是任意解。$G=A^-DB^-+G-A^-DB^-=A^-DB^-+G-A^-AGBB^-$

即$AXB=D$的任意解都可以写成$A^-DB^-+Y-A^-AYBB^-$的形式。因此$X=A^-DB^-+Y-A^-AYBB^-$是矩阵方程的通解。

推论 1：

$AX=D$有解的充要条件是存在A^-，使得$AA^-D=D$成立。此时通解为$X=A^-D+Y-A^-AY$

推论 2：

$XB=D$有解的充要条件是存在B^-，使得$DB^-B=D$成立。此时通解为$X=DB^-+Y-YBB^-$

（分别令A、$B=E$可得推论 1、2）

推论 3：

$Ax=b$有解的充要条件是存在A^-，使得$AA^-b=b$成立。此时通解为$x=A^-b+(E_n-A^-A)u$，$\forall u\in C^n$

（特解为$x^*=A^-b$，导出组$Ax=0$ 的通解为$x=(E_n-A^-A)u$）

在计算时可以取$A^-=A^+$。如果$AA^+b\ne b$，则无解。

考题中特解$x^*=A^+b$需要具体求出来，通解$x=(E_n-A^-A)u$不要求算

二、相容方程的最小范数解

相容方程组：有解的方程组

定义 1：最小范数解：

方程组有解，所有解中 2 范数最小的解

定理 4（求最小范数解）：

设$D\in A\{1,3\}$（即$\begin{cases}ADA=A\\(DA)^H=DA\end{cases}$），则$Db$是相容方程组$Ax=b$的最小范数解，并且最小范数解唯一。

证：

方程组$Ax=b$的通解为$x=A^-b+(E_n-A^-A)u$，取$A^-=D$，则$x=Db+(E_n-DA)u$。

计算内积$(Db,(E_n-DA)u)=(DAx_0)^H(E_n-DA)u=x_0^HA^HD^H(E_n-DA)u=x_0^HA^HD^Hu-x_0^HA^HD^HDAu$

$=x_0^HA^HD^Hu-x_0^HA^HD^H(DA)^Hu=x_0^H(DA)^Hu-x_0^H(DADA)^Hu=x_0^H(DA)^Hu-x_0^H(DA)^Hu=0$，即$Db\perp(E_n-DA)u$

因此Db是以x为斜边的直角三角形的直角边，有$\|Db\|\le\|x\|$，即Db是最小范数解。

垂线段唯一，因此最小范数解唯一。

取$D=A^+$，有最小范数解为A^+b。

定理 5（定理 4 的逆命题）：

若Db是相容方程组$Ax=b$的最小范数解，则$D\in A\{1,3\}$。

三、不相容方程组的解

定义 2：最小二乘解：

存在$\bar{x}$对任意x都有$\|A\bar{x}-b\|\le\|Ax-b\|$，则$\bar{x}$是$Ax=b$的最小二乘解。

定理 6（求最小二乘解）：

设$D\in A\{1,4\}$（即$\begin{cases}AGA=A\\(AG)^H=AG\end{cases}$），则$x=Gb$是不相容方程组$Ax=b$的最小二乘解。

证：

因为$(Ax-b)=(Ax-AGb)+(AGb-b)$

计算内积$(Ax-AGb,AGb-b)=(Ax-AGb)^H(AGb-b)=(x^HA^H-b^H(AG)^H)(AGb-b)$
$=x^HA^HAGb-b^H(AG)^HAGb-x^HA^Hb+b^H(AG)^Hb=x^HA^H(AG)^Hb-b^HAGAGb-x^HA^Hb+b^HAGb$
$=x^H(AGA)^Hb-x^HA^Hb=0$，故$Ax-AGb\perp AGb-b$
所以$AGb-b$是以$Ax-b$为斜边的直角三角形的直角边，有$\|AGb-b\|\le\|Ax-b\|$，即Gb是最小二乘解

引理 1：

x是不相容方程组$Ax=b$的最小二乘解的充要条件是$Ax=AGb$。其中$G\in A\{1,4\}$。

定理 7（最小二乘解通解）：

不相容方程组$Ax=b$的最小二乘解的通解为$x=Gb+(E_n-A^-A)u$，$G\in A\{1,4\}$
（$A(x-Gb)=0$，由推论 3 得$x-Gb=(E_n-A^-A)u$）

定义 3：最佳逼近解：

设x_0是不相容方程组$Ax=b$的最小二乘解。如果对任意最小二乘解$\bar{x}$，均有$\|x_0\|\le\|\bar{x}\|$，则称x_0为最佳逼近解

定理 8（求最佳逼近解）：

$x_0=A^+b$是不相容方程组$Ax=b$的最佳逼近解。

例题：$Ax=b$

1. 最大秩分解$A=BD$
2. $A^+=D^H(DD^H)^{-1}(B^HB)^{-1}B^H$
3. 算A^+b
4. 算AA^+b。$\begin{cases}AA^+b=b & A^+b\text{是最小范数解}\\ AA^+b\ne b & A^+b\text{是最佳逼近解}\end{cases}$
5. 算$\begin{cases}\text{矩阵方程组通解}\\ \text{最小二乘解通解}\end{cases}$，$x=A^+b+(E_n-A^+A)u$（$(E_n-A^+A)$不用算出结果，表示一下）