矩阵论笔记3

张天阳*
University of Chinese Academy of Science

December 11, 2017

*E-mail：zhangtianyang17@mails.ucas.edu.cn

Contents

Part I
矩阵分解

1 Gauss消去法:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{13} \\ a_{21} & a_{22} & \cdots & a_{23} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n3} \end{bmatrix}$$

首先行变换左乘，列变换右乘
每次高斯消去都是左乘第k次变换的矩阵L_k:

$$L_k = I - \boldsymbol{l}_k \boldsymbol{e}^T \tag{1}$$

$$\boldsymbol{l}_k = [0_1, 0_2, 0_3, ..., 0_k, c_{k+1}, c_{k+2}, c_{k+3}, ..., cn]^T, \quad 0_k\ means\ kth\ 0 \tag{2}$$

因为每次高斯消去法都是将第k 列消去，所以就是减去一个第k 行从a_{kk} 开始往下都是$\frac{a_{k(k+i)}}{a_{kk}}$乘以第k行再减去自己，于是相当于乘以:

$$\boldsymbol{l}_k \boldsymbol{e}^T = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \cdots & 0_{k,k} & \cdots & \vdots \\ 0 & \cdots & \frac{a_{k,(k+1)}}{a_{k,k}} & \cdots & 0 \\ 0 & \cdots & \vdots & \cdots & 0 \\ 0 & \cdots & \frac{a_{k,(n)}}{a_{k,k}} & \cdots & 0 \end{bmatrix} \tag{3}$$

$$L = L_1 L_2 ... L_n = I - \boldsymbol{l}_1 \boldsymbol{e}^T - \boldsymbol{l}_2 \boldsymbol{e}^T - ... - \boldsymbol{l}_{n-1} \boldsymbol{e}^T \tag{4}$$

则上三角矩阵LA可得。

2 三角分解:

$$A = LU \tag{5}$$

先求出Gauss消去法的L。
矩阵分解中L是Gauss中的L^{-1}

$$L = L_1 L_2 ... L_n = I + \boldsymbol{l}_1 \boldsymbol{e}^T + \boldsymbol{l}_2 \boldsymbol{e}^T + ... + \boldsymbol{l}_{n-1} \boldsymbol{e}^T \tag{6}$$

而U就是Gauss消去法所得到的上三角矩阵。

2.1 Doolittle分解

即将矩阵分解为LDU其中U为对角值为1的矩阵，而D则是对角矩阵，其实就是讲矩阵分解为LU之后将U逐行提取公因式，写作D。
而LU则称为Doolittle分解。

2.2 Cholesky分解

Choleshy分解是将是对称正定矩阵A分解为

$$A = G^T G \tag{7}$$

其是现将矩阵分解为：LDU形式:
可得

$$G = L\sqrt{D} \tag{8}$$

即D的每个对角线的元素（D为对角矩阵）都开跟号

3 分块矩阵求逆

3.1 行变换解析

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ -CA^{-1}A + C & -CA^{-1}B + D \end{bmatrix}, \quad Gauss\ Method \tag{9}$$

相当于左乘（行变换），注：分块矩阵乘法是原来在左边的乘完还在左边。

$$X = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ 0 & \cdots & x_{ij} & \cdots & 1 & \cdots & 0 \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}, only\ X_{ij} = x_{ij}$$

$$X = I + x_{ij} e_i e_j^T$$

这样的变换矩阵X，左乘A，XA相当于将A的第 i 行加上第 j 列乘以x_{ij}

解释上面为什么是第一行左乘CA^{-1}在加到第二行，因为高斯消去法需要让第二行第一个元素为O所以要将第一行乘以（左乘是因为行变换是左乘，CA^{-1}是因为$A^{-1}C$的话消不掉A会变为$A^{-1}CA$)。**该行都同时乘以A的话是左乘，而该列都同时乘以A的话是右乘。**

3.2 分块矩阵求逆

对下面分块矩阵求逆：应该先拓展，然后将其行变换为单位阵然后拓展阵自然变成它的逆。

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \Rightarrow \begin{bmatrix} A & B & I & O \\ C & D & O & I \end{bmatrix}$$

下面的式子第二行应该归一化，因为不然化三角时会让式子很长。

$$\begin{bmatrix} A & B & I & O \\ C & D & O & I \end{bmatrix} = \begin{bmatrix} A & B & I & O \\ -CA^{-1}A+C & -CA^{-1}B+D & O-CA^{-1}B & I-CA^{-1}B \end{bmatrix}$$

$$= \begin{bmatrix} A & B & I & O \\ O & I & -(-CA^{-1}B+D)^{-1}CA^{-1}B & (-CA^{-1}B+D)^{-1}(I-CA^{-1}B) \end{bmatrix}$$

$$= \begin{bmatrix} A & O & I+B^{-1}(-CA^{-1}B+D)^{-1}CA^{-1}B & -B^{-1}(-CA^{-1}B+D)^{-1} \\ O & I & -(-CA^{-1}B+D)^{-1}CA^{-1}B & (-CA^{-1}B+D)^{-1}(I-CA^{-1}B) \end{bmatrix}$$

$$= \begin{bmatrix} I & O & A^{-1}[I+B^{-1}(D-CA^{-1}B)^{-1}CA^{-1}B] & -A^{-1}[B^{-1}(D-CA^{-1}B)^{-1}] \\ O & I & -(D-CA^{-1}B)^{-1}CA^{-1}B & (D-CA^{-1}B)^{-1}(I-CA^{-1}B) \end{bmatrix}$$

所以其逆为：

$$= \begin{bmatrix} A^{-1}[I+B^{-1}(D-CA^{-1}B)^{-1}CA^{-1}B] & -A^{-1}[B^{-1}(D-CA^{-1}B)^{-1}] \\ -(D-CA^{-1}B)^{-1}CA^{-1}B & (D-CA^{-1}B)^{-1}(I-CA^{-1}B) \end{bmatrix}$$

4 Givens矩阵（旋转矩阵）对向量操作：

4.1 二维例子

顺时针旋转 θ角(在极坐标系中角度增加$-\theta$)后空间中的向量$\boldsymbol{x}$为：因为极坐标中的$\boldsymbol{x}$为:

$$x_1 = r \cdot cos(\varphi), \quad x_2 = r \cdot sin(\varphi)$$

$$\begin{bmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{bmatrix} \boldsymbol{x} = T_{12}\boldsymbol{x} \tag{10}$$

用c表示$cos(\theta)$，s表示$sin(\theta)$

4.2 高维时的旋转矩阵

高纬时候其仍然一次只旋转两个维度，因为任意两个维度可以组成一个平面，然后在该平面上有上面二维的结论，故当旋转将 i 向着 j 旋转时：

$$\boldsymbol{x} = [x_1, x_2, ..., x_n]^T, \quad \boldsymbol{y} = T_{ij}\boldsymbol{x} = [y_1, y_2, ..., y_n]^T$$

$$\begin{cases} y_i = cx_i + sx_j \\ y_j = -sx_i + sx_j \\ y_k = x_k, \quad (if\ k \neq i, j) \end{cases}$$

T_{ij}被叫做GIVENS矩阵

4.3 旋转定理

1.存在有限个Givens矩阵的乘积，记作T可以使$T\boldsymbol{x} = |\boldsymbol{x}|\boldsymbol{e}_1$，$\boldsymbol{e}_i$为仅仅只有第i维为1的单位列向量。
2.同时也存在有限个Givens矩阵的乘积，记作T可以使$T\boldsymbol{x} = |\boldsymbol{x}|\boldsymbol{z}$，$z$为任意单位向量

5 Householder变换（初等反射变换）

5.1 Householder矩阵定义与推导

$$H = I - 2\boldsymbol{u}\boldsymbol{u}^T \tag{11}$$

$\boldsymbol{u}$是$\boldsymbol{x}$要对称的超平面的法向量（单位向量，方向垂直于超平面），$\boldsymbol{\mu}$为图上与法向量方向相同的向量,推导：

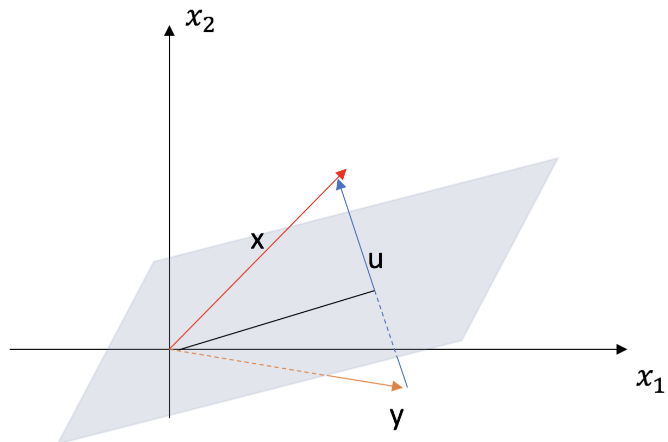

Figure 1: 反射变换图

$$\begin{aligned}
According\ to\ the\ defination: \quad H\boldsymbol{x} &= \boldsymbol{y} \\
\boldsymbol{x} - \boldsymbol{y} &= 2\boldsymbol{\mu} \\
\Rightarrow (I - H)\boldsymbol{x} &= 2\boldsymbol{\mu} \\
Becasus\ \boldsymbol{u}^T\boldsymbol{x} = \|\boldsymbol{\mu}\|,\ So\ \ \boldsymbol{u}\boldsymbol{u}^T\boldsymbol{x} &= \boldsymbol{\mu} \\
\Rightarrow H &= I - 2\boldsymbol{u}\boldsymbol{u}^T
\end{aligned}$$

5.2 Householder变换性质

1. H矩阵性质：

$$\begin{cases}
H^T = H \quad , \ (\lambda \ \ is \ \ real\ number) \\
H^T H = I, (Thus\ H^T = H^{-1}) \quad , \ \ H\ is\ Orthogonal\ Matrix \\
H^2 = I \\
H^{-1} = H \\
|H| = -1
\end{cases} \tag{12}$$

2. 对于任意$\boldsymbol{x} \in R^n$机器单位列向量$z \in R^n$存在Householder矩阵H，使得$Hx = |x|\boldsymbol{z}$，z为任意单位向量。
3. 任意Givens矩阵（旋转矩阵）可以由两个Householder矩阵相乘得到，但是Householder矩阵不可以由若干Givens矩阵相乘得到，因为$|G| = 1, |H| = -1$。

6 矩阵的QR分解

6.1 定义

矩阵的QR分解是将满秩矩阵A分解为$A=QR$其中Q为正交矩阵（酉矩阵$Q^TQ=I$）,而R则为上三角矩阵。

6.2 方法一：Schmidt正交化方法

因为A为满秩矩阵所以A的所有列是线性无关的，所以其所有列可以正交化得到一组正交基，将正交化的列放在一起便成了正交矩阵B。具体过程如下：

1. 对于矩阵$A=(\boldsymbol{a}_1,\boldsymbol{a}_2,...,\boldsymbol{a}_n)$作正交化：Schmidt正交化过程可以想象一个下三角矩阵，其每个下标是k_{ij}而其第n行（每一行）都是$\boldsymbol{b}_n$对于$\boldsymbol{a}_n$的线性组合,而每行的$\boldsymbol{b}_k$表示列的下标。

   $$\begin{aligned} &\text{C的行为下面的列、第一行、第二行......第n 行}\\ \boldsymbol{a}_1 &= \boldsymbol{b}_1\\ \boldsymbol{a}_2 &= k_{21}\boldsymbol{b}_1+\boldsymbol{b}_2\\ \boldsymbol{a}_3 &= k_{31}\boldsymbol{b}_1+k_{32}\boldsymbol{b}_2+\boldsymbol{b}_3\\ &...\\ \boldsymbol{a}_n &= k_{n1}\boldsymbol{b}_1+k_{n2}\boldsymbol{b}_2+....+\boldsymbol{b}_n \end{aligned}$$

   几何意义想象一下，第k行为$\boldsymbol{a}_k$在$\boldsymbol{b}_1,...,\boldsymbol{b}_k$的空间内的坐标，其中$k_{ij}$为第j个坐标的值为：$\boldsymbol{a}_k$在$\boldsymbol{b}_j$上的投影乘以单位化的$\boldsymbol{b}_j$即：

2. 其中$\boldsymbol{a}_k$是已知要被分解的矩阵A的第k列，而k_{ij}为：

   $$k_{ij}=\frac{(\boldsymbol{a}_i,\boldsymbol{b}_j)}{(\boldsymbol{b}_j,\boldsymbol{b}_j)} \tag{13}$$

   所以$k_{ii}=1$所以对角线上的k为1。于是可以得到b_i的公式：

   $$\begin{aligned} \boldsymbol{b}_1 &= \boldsymbol{a}_1\\ \boldsymbol{b}_2 &= \boldsymbol{a}_2-k_{21}\boldsymbol{b}_1\\ \boldsymbol{b}_3 &= \boldsymbol{a}_3-k_{31}\boldsymbol{b}_1-k_{32}\boldsymbol{b}_2\\ &...\\ \boldsymbol{b}_n &= \boldsymbol{a}_n-k_{n1}\boldsymbol{b}_1-k_{n2}\boldsymbol{b}_2-....-k_{nn-1}\boldsymbol{b}_{n-1} \end{aligned}$$

   所以$A=BC$，C为上面所述。
   我们求出所有$\boldsymbol{b}_i$之后我们将其对角化之后便是Q 而对角化需要乘以对角矩阵:
   $diag(1/\|\boldsymbol{b}_1\|,1/\|\boldsymbol{b}_2\|,...,1/\|\boldsymbol{b}_n\|)$
   矩阵$diag(|\boldsymbol{b}_1\|,\|\boldsymbol{b}_2\|,...,\|\boldsymbol{b}_n\|)C$即为$R$

3. 综上所述可得$A=QR$

6.3 方法二：Givens矩阵变换得到QR分解

Givens变换T_1可以使得矩阵$A = (\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n)$的第一列变为 $[\|\boldsymbol{a}_1\|, 0, 0, ..., 0]^T$,于是将矩阵A乘以$T_1$可以得到：($T_1$变换时最接近$A$,即将$A$用许多Givens变换变为上三角矩阵(R)时$T_1$为最接近$A$的那个左乘矩阵)

$$T_1 A = \begin{bmatrix} \|\boldsymbol{a}_1\| & * & .. & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}_{n \times n} \tag{14}$$

再对A_1的第一列进行Givens变换$T_2 A_1$

$$T_2 A_1 = \begin{bmatrix} \|\boldsymbol{a}_1^{(1)}\| & * & .. & * \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{bmatrix}_{(n-1) \times (n-1)} \tag{15}$$

以此类推：

$$T = \begin{bmatrix} I_{n-2} & O \\ O & T_{n-1} \end{bmatrix} \begin{bmatrix} I_{n-3} & O \\ O & T_{n-2} \end{bmatrix} \cdots \begin{bmatrix} 1 & O \\ O & T_2 \end{bmatrix} T_1 \tag{16}$$

$$TA = R \tag{17}$$

于是：

$$A = T^{-1} R = QR \tag{18}$$

6.4 方法三：利用Householder变换得到QR分解

思路与用Givens矩阵得到QR分解一样也是先用变换矩阵T得到上三角矩阵R
同样先对$A = (\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n)$的第一列找到变换矩阵$H_1$使得$\boldsymbol{a}_1$变为：$\|\boldsymbol{a}_1\| \boldsymbol{e}_1$,然后以此类推。

$$T = \begin{bmatrix} I_{n-2} & O \\ O & H_{n-1} \end{bmatrix} \begin{bmatrix} I_{n-3} & O \\ O & H_{n-2} \end{bmatrix} \cdots \begin{bmatrix} 1 & O \\ O & T_2 \end{bmatrix} T_1 \tag{19}$$

$$TA = R \tag{20}$$

于是：

$$A = T^{-1} R = T^T R = QR \tag{21}$$

对于H的求解方法：$H = I - 2\boldsymbol{u}\boldsymbol{u}^T$而对于$\boldsymbol{u}$的求解：

$$\boldsymbol{\mu} = \boldsymbol{a}_1 - \|\boldsymbol{a}_1\| \boldsymbol{e}_1$$
$$\boldsymbol{u} = \frac{\boldsymbol{\mu}}{\|\mu\|}$$

7 矩阵的满秩分解：FG分解

用于分解一些不是方阵（非方阵）的矩阵。
分解方法:

1. 用高斯消去法将待分解的矩阵转换为Hermite标准型（不是Hermite矩阵）
   Hermite标准型：每一行第一个非零元素为1，前r行为非零行。

2. 取A的每个Hermite（标准型）对应的首1列（行的第一个非零的元素也就是1对应的列）组成F，而Hermite标准型的所有非零行构成G。

满秩分解重要性质：$A_r^{m\times n} = F_r^{m\times r} G_r^{r\times n}$

8 奇异值分解

8.1 前序知识与回顾

8.1.1 谱分解

如果A是实对称矩阵那么A满足:

$$Q^T AQ = diag(\lambda_1, \lambda_2, ..., \lambda_n)\text{，其中Q为正交矩阵} \tag{22}$$

如果A不是实数对称矩阵，但是A仍然非奇异，那么A可以分解为:

$$P^T AQ = diag(\lambda_1, \lambda_2, ..., \lambda_n)\text{，其中P与Q为正交矩阵} \tag{23}$$

而奇异值分解为无任何条件的矩阵A可以分解为:

$$U^H AV = \begin{bmatrix} \Sigma_{r\times r} & O \\ O & O \end{bmatrix} \tag{24}$$

即：V是左乘A的所以对于运算时用$A^T A$还是AA^T，可以先求出V或者U

$$A = U \begin{bmatrix} \Sigma_{r\times r} & O \\ O & O \end{bmatrix} V^H \tag{25}$$

8.2 奇异值分解

8.2.1 思路与引理：其与前序的方法不同也不要求A是非奇异

1. 由于A虽然不是实对称矩阵,A为$\forall A \in \boldsymbol{C}^{m\times n}$，但是$A^T A$是Hermite矩阵（对称矩阵）所以其是半正定的。

2. 且$rank(A^T A) = rank(A)$

3. $A^T A = O$的冲要条件是$A = O$

4. 对A^TA进行谱分解可以得到：$V^TA^TAV = diag(\lambda_1, \lambda_2, ..., \lambda_n)$，其中V为正交(酉)矩阵其中其中$\lambda_1, \lambda_2, ..., \lambda_n$为矩阵$A^TA$的特征值,而$\sigma_i = \sqrt{\lambda_i}$，$(i = 1, 2, ..., n)$被称为其奇异值。
   对于V取其每一列$\boldsymbol{v}_i$可以发现：
   $$A^TA\boldsymbol{v}_i = \boldsymbol{v}_i\lambda_i \tag{26}$$
   所以求出每个λ_i对应的特征向量$\boldsymbol{\xi}_i$,(零特征值的也求，重根也求，求n个嘛)（每个特征值只求一个即可），可得$V = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_n)$,

8.2.2 奇异值分解证明

1. 上一小节得到的矩阵$B = A^TA$的特征值对角矩阵$diag(\lambda_1, \lambda_2, ..., \lambda_n)$其中有r个非零项，所以将其写为：
   $$diag(\lambda_1, \lambda_2, ..., \lambda_n) = \begin{bmatrix} \Sigma^2_{r\times r} & O \\ O & O \end{bmatrix} \tag{27}$$

2. 求出每个λ_i对应的特征向量,(零特征值的也求，重根也求，求n个嘛)（每个特征值只求一个即可），可得$(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_n)$,由于谱分解的定律我们可以得到：
   $$V^HA^TAV = \begin{bmatrix} \Sigma^2_{r\times r} & O \\ O & O \end{bmatrix}$$
   $$V = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_n)$$
   取$V_{n\times n}$的前r列构成矩阵V_1可以得到：$V = [V_{1(n\times r)}|V_{2(n\times n-r)}]$
   可知：
   $$V_1^HA^TAV_1 = \Sigma^2_{(r\times r)} \tag{28}$$

3. 可以将上式写成：
   $$A^TA = V_1\Sigma^2V_1^H\text{，因为V正交}$$
   $$\Rightarrow (AV_1\Sigma^{-1})^H(AV_1\Sigma^{-1}) = I$$
   所以：令 $U_1 = AV_1\Sigma^{-1}$可得
   $$\begin{cases} U_1^TU_1 = I\text{，得到}U_1\text{为正交矩阵} \\ U_1^HAV_1 = \Sigma \end{cases} \tag{29}$$
   让$U = [U_1|U_2]$其中U_2为随意列只要保证U正交即可：因为：
   $$\begin{aligned} U^HAV =& [U_1|U_2]^HA[V_1|V_2] \\ =& [U_1A|U_2A]^H[V_1|V_2] \\ =& \begin{bmatrix} U_1^HAV_1 & U_1^HAV_2 \\ U_2^HAV_1 & U_2^HAV_2 \end{bmatrix} \end{aligned}$$

又因为

$$
\begin{aligned}
AV_1 = U_1\Sigma \Rightarrow U_2^H AV_1 = U_2^H U_1 \Sigma = O,\text{（}U_2, U_1\text{正交）} \\
AV_2 = O \\
U_1^H AV_1 = \Sigma \\
\Rightarrow U^H AV = \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix}
\end{aligned}
$$

8.2.3 用A^TA奇异值分解使用方法

记住VA^TAV在一起，即V为n阶方阵

1. 得到的矩阵$B = A^TA$的特征值对角矩阵$diag(\lambda_1, \lambda_2, ..., \lambda_n)$其中有r个非零项，所以将其写为:

$$diag(\lambda_1, \lambda_2, ..., \lambda_n) = \begin{bmatrix} \Sigma^2_{r\times r} & O \\ O & O \end{bmatrix} \tag{30}$$

2. 求出每个λ_i对应的特征向量,(零特征值的也求，重根也求，求n个嘛)（每个特征值只求一个即可），可得$(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_n)$,由于谱分解的定律我们可以得到:

$$
\begin{aligned}
V^H A^T AV = \begin{bmatrix} \Sigma^2_{r\times r} & O \\ O & O \end{bmatrix} \\
V = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_n)
\end{aligned}
$$

   取$V_{n\times n}$的前r列构成矩阵V_1可以得到：$V = [V_{1(n\times r)}|V_{2(n\times n-r)}]$

3. 令 $U_1 = AV_1\Sigma^{-1}$(用$U_1^H AV = \Sigma$记忆。)可得$U = [U_1|U_2]$让其中U_2为随意列只要保证U正交即可:

4. 则:

$$U^H AV = \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix} \tag{31}$$

8.2.4 用AA^T奇异值分解使用方法

记住UAA^TU在一起，即U为m阶方阵

1. 得到的矩阵$C = AA^T$的特征值对角矩阵$diag(\lambda_1, \lambda_2, ..., \lambda_n)$其中有r个非零项，所以将其写为:

$$diag(\lambda_1, \lambda_2, ..., \lambda_n) = \begin{bmatrix} \Sigma^2_{l\times l} & O \\ O & O \end{bmatrix} \tag{32}$$

2. 求出每个λ_i对应的特征向量,(零特征值的也求，重根也求，求n个嘛)（每个特征值只求一个即可），可得$(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, ..., \boldsymbol{\eta}_n)$,由于谱分解的定律我们可以得到:

$$
\begin{aligned}
U^H AA^T U = \begin{bmatrix} \Sigma^2_{l\times l} & O \\ O & O \end{bmatrix} \\
U = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, ..., \boldsymbol{\eta}_n)
\end{aligned}
$$

   取$U_{n\times n}$的前r列构成矩阵U_1可以得到：$U = [U_{1(n\times l)}|U_{2(n\times n-l)}]$

3. 令 $V_1 = (U_1^H A\Sigma^{-1})^H$可得$V = [V_1|V_2]$让其中$V_2$为随意列只要保证$V$正交即可：
4. 则：
$$U^H AV = \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix} \tag{33}$$

8.3 奇异值分解的定理与应用

8.3.1 定理

1. $A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^H + \sigma_2 \boldsymbol{u}_2 \boldsymbol{v}_2^H + ... + \sigma_n \boldsymbol{u}_n \boldsymbol{v}_n^H$
2. $N(A) = L(\boldsymbol{v}_{r+1}, \boldsymbol{v}_{r+2}, ..., \boldsymbol{v}_{r+n})$
3. $R(A) = L(\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_n)$

8.3.2 应用

数据复原上我们有了一些数据，就是数据矩阵上的一些点是已知X_{ij}，$(i,j) \in \mathscr{A}$的但是矩阵中存在未知的元素在里面。
我们要求已知点被还原时满足

Part II

矩阵的特征值的重要性质：

1. 其中$\lambda(A)$代表A的特征值
2. $$\lambda(B) = 1/\lambda(B^{-1}), \lambda(A^T A) = \lambda(AA^T)$$
3. 即A的特征值是λ，则A^{-1}的特征值是$\frac{1}{\lambda}$
4. 且A^k的特征值为λ^k
5. AA^T的非零特征值为$A^T A$的非零特征值，AB与BA的非零特征值一样的，又因为$tr(AB) = tr(BA)$所以非零值不仅一样重数也一样。

$$\begin{aligned} &\text{推导} \\ A\boldsymbol{x} &= \lambda \boldsymbol{x}, \quad B\boldsymbol{x} = \beta \boldsymbol{x} \\ BA\boldsymbol{x} &= B\lambda \boldsymbol{x} = \lambda B\boldsymbol{x} = \lambda\beta\boldsymbol{x} \\ AB\boldsymbol{x} &= A\beta\boldsymbol{x} = \beta A\boldsymbol{x} = \beta\lambda\boldsymbol{x} \end{aligned}$$