Lecture 30

奇异值分解简称 SVD，是对矩阵最好的分解，对于任意矩阵都有：

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

其中 U、V 均是正交矩阵，Σ是对角矩阵。

正定矩阵中学过，对称矩阵可分解为：$\mathbf{A}=\mathbf{Q\Lambda}\boldsymbol{Q}^{T}$，这里只需要一个矩阵 Q 就可以将对称矩阵分解，这个也是正定矩阵的奇异值分解。

在对角化中，可对角化的矩阵可分解为：$\mathbf{A}=\ \boldsymbol{S}\Lambda\boldsymbol{S}^{-1}$，这里的 S 不是正交阵，所以这不是奇异值分解。

现在我们希望在矩阵 A 的列空间中选取一组标准正交基，$\boldsymbol{v}_1$，$\boldsymbol{v}_2\cdots\boldsymbol{v}_r$，这组基在矩阵 A 的映射下，转化为行空间的一组标准正交基$\boldsymbol{u}_1$，$\boldsymbol{u}_2\cdots\boldsymbol{u}_r$。即

$$\mathbf{A}\boldsymbol{v}_1=\boldsymbol{\sigma}_1\boldsymbol{u}_1$$

....

$$\mathbf{A}\boldsymbol{v}_r=\boldsymbol{\sigma}_r\boldsymbol{u}_r$$

写成矩阵的形式：

$$A[\boldsymbol{v}_1\quad \boldsymbol{v}_2\quad \dots\quad \boldsymbol{v}_r]=[\boldsymbol{\sigma}_1\boldsymbol{u}_1\quad \boldsymbol{\sigma}_2\boldsymbol{u}_2\quad \dots\quad \boldsymbol{\sigma}_r\boldsymbol{u}_r]=[\boldsymbol{u}_1\quad \boldsymbol{u}_2\quad \dots\quad \boldsymbol{u}_r]\begin{bmatrix}\boldsymbol{\sigma}_1 & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \boldsymbol{\sigma}_r\end{bmatrix}$$

其中$\boldsymbol{\sigma}$称为缩放因子，在转换的过程中，有可能有伸缩或者拉伸。如果加上零空间和左零空间，有：

$$A[\boldsymbol{v}_1\quad \boldsymbol{v}_2\quad \boldsymbol{v}_r\quad \dots\quad \boldsymbol{v}_m]=[\boldsymbol{u}_1\quad \boldsymbol{u}_2\quad \boldsymbol{u}_r\quad \dots\quad \boldsymbol{u}_n]\left[\begin{array}{ccc|c}\boldsymbol{\sigma}_1 & & & \\ & \ddots & & \\ & & \boldsymbol{\sigma}_r & \\ \hline & & & \mathbf{0}\end{array}\right]$$

此时的 V 是 n×n 矩阵,U 是 m×m 矩阵,Σ是 m×n 的对角矩阵$\boldsymbol{\sigma}_1\cdots\boldsymbol{\sigma}_r$在对角线上，其余元素均为 0。

根据矩阵形式，写为：AV= $\mathbf{U\Sigma}$，右乘V^{-1},有 A= $\mathbf{U\Sigma V}^{-1}$，又因为 V 是正交矩阵，所以有：

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

下面我们来看一下 U 和 V 分别的含义(假定 A 是可逆方阵)：

1) $\mathbf{A}=\mathbf{U\Sigma V}^{T}$，左乘$\boldsymbol{A}^{T}$，有：

$$\boldsymbol{A}^{T}\mathbf{A}=\ \mathbf{V\Sigma}^{T}\boldsymbol{U}^{T}\,\mathbf{U\Sigma V}^{T}=\ \mathbf{V\Sigma}^{T}\boldsymbol{I}\mathbf{\Sigma V}^{T}=\mathbf{V\Sigma}^{T}\mathbf{\Sigma V}^{T}=\mathbf{V\Sigma}^{2}\mathbf{V}^{T}$$

可以将 V 看作是$\boldsymbol{A}^{T}\mathbf{A}$的特征向量，$\Sigma^2$为特征值的平方。

2) $\mathbf{A}=\mathbf{U\Sigma V}^{T}$，右乘$\boldsymbol{A}^{T}$，有：

$$\mathbf{A}\boldsymbol{A}^{T}=\ \mathbf{U\Sigma V}^{T}\,\mathbf{V\Sigma}^{T}\boldsymbol{U}^{T}=\ \mathbf{U\Sigma}^{T}\boldsymbol{I}\mathbf{\Sigma U}^{T}=\mathbf{U\Sigma}^{T}\mathbf{\Sigma U}^{T}=\mathbf{U\Sigma}^{2}\mathbf{U}^{T}$$

可以将 U 看作是$\mathbf{A}\boldsymbol{A}^{T}$的特征向量，$\Sigma^2$为特征值的平方。

例 1：

A= $\begin{bmatrix}4 & 4\\ -3 & 3\end{bmatrix}$，计算矩阵 A 的 SVD 分解。

1) $\boldsymbol{A}^{T}$A=$\begin{bmatrix}25 & 7\\ 7 & 25\end{bmatrix}$，可以计算出特征值分别为 32、18，特征向量分别为：

$\begin{bmatrix}1\\ 1\end{bmatrix}$、$\begin{bmatrix}1\\ -1\end{bmatrix}$，化为单位向量为$\boldsymbol{v}_1=\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}$、$\boldsymbol{v}_2=\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{-1}{\sqrt{2}}\end{bmatrix}$，$\boldsymbol{\sigma}_1=32$，$\boldsymbol{\sigma}_2=18$

2) AA^T=$\begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$，可以计算出特征值分别为 32、18，特征向量分别为：$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$、$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$。但是我们不能直接使用这一组特征向量，因为 UV 必须满足式子 AV =UΣ。A$v_1 = \begin{bmatrix} 0 \\ -3\sqrt{2} \end{bmatrix}$，所以我们取$u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$、$u_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$。

3) 带入上面的结果：

$$\begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$$

当 A 中存在零空间和左零空间时：

v_1，$v_2 \cdots v_r$是行空间的一组标准基

u_1，$u_2 \cdots u_r$是列空间的一组标准基

v_r，$v_{r+1} \cdots v_n$是零空间的一组基

u_{r+1}，$u_{r+1} \cdots u_m$是左零空间的一组基