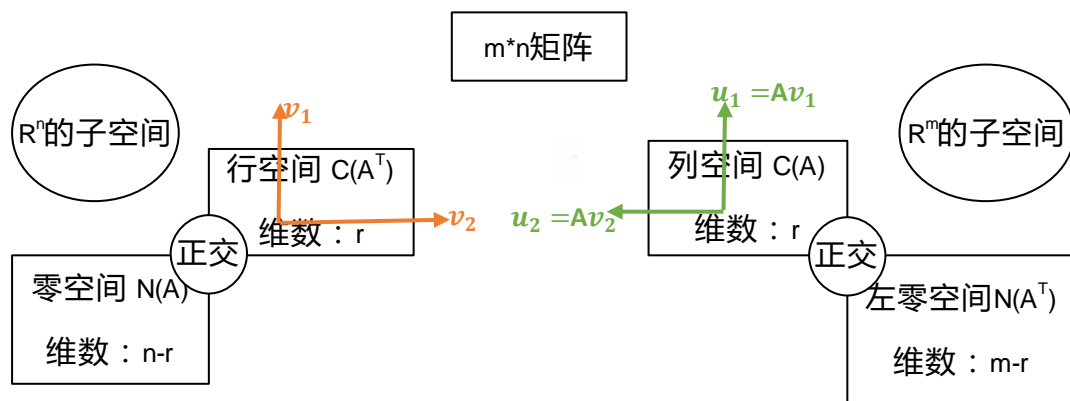


## 一. 知识概要

本节介绍矩阵的奇异解分解，本质就是将行空间的一组正交基变换成列空间里的一组正交基。示意如下图



## 二. 奇异值分解 SVD

### 2.1 基变换

对矩阵做奇异值分解，首先如上图，先在行空间中找到一组正交基  $v_1, v_2, \dots, v_r$ ，然后通过矩阵  $A$  对其进行线性变换，得到列空间中一组正交基  $u_1, u_2, \dots, u_r$ 。

$$Av_i = \sigma_i u_i$$

注：(1)  $v_i$  是单位的，将其变换之后的  $u_i$  也是单位的，对应的  $\sigma_i$  就是伸缩因子

(2) 找行空间的一组正交基方法：我们之前学习的格拉姆-施密特正交化方法就可以在行空间中得到一组正交基  $v_1, v_2, \dots, v_r$ 。

### 2.2 矩阵形式

接下来将我们这个过程用矩阵形式表示出来，

$$A[v_1 \ v_2 \ \dots \ v_r] = [u_1 \ u_2 \ \dots \ u_r] \begin{bmatrix} \sigma_1 & & & \\ & \sigma_1 & & \\ & & \dots & \\ & & & \dots \end{bmatrix}$$

$$\Leftrightarrow AV = U\Sigma$$

- $v_i$  对应的是行空间中的基向量；
- $u_i$  对应的是列空间中的基向量；
- $\sigma_i$  对应的是伸缩因子；

我们的任务就是由给定矩阵  $A$  得到其余三者。

即：

$$AV = U\Sigma$$

简单变形：

$$A = U\Sigma V^{-1} = U\Sigma V^T$$

(因为  $V$  与  $U$  都是正交基构成的正交矩阵，转置即为逆)

这个形式就是 SVD 分解： $A = U\Sigma V^{-1}$ 。任意矩阵都可以被分解为这个形式。

即两个正交矩阵  $U, V$  以及一个对角矩阵  $\Sigma$ 。

### 2.3 计算方法

但是在这种变形方式中， $U$  与  $V$  并不好同时寻找，我们考虑使用变形技巧来简化求解过程，如：

假设现在希望做些处理，将  $U$  消去。

使用  $A^T A$  运算，利用  $U$  是正交阵的性质：

$$A^T A = V \Sigma U^T U \Sigma V^T = V \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \dots & \\ & & & \dots \end{bmatrix} V^T$$

这样变形之后，我们就只需要选择  $V$  就可以了。同样，我们经过相似处理，使用  $AA^T$  运算，这样剩下的就是  $U$  矩阵。

我们在看看这个  $A^T A$  矩阵形式，回顾之前对称矩阵分解 (26 课) 形式，发现满足以下形式：

$$A^T A = Q \Lambda Q^T$$

也就是说，这个式子就是  $A^T A$  这个对称矩阵 (也是正交矩阵) 的分解， $V$  由  $A^T A$  这个矩阵的线性无关的特征向量标准化后构成 (就是我们选取的行空间中的基)，而中间的对角阵则是  $\Sigma$  则是由  $A^T A$  的特征向量  $\sigma_i^2$  构成。

这意味着：我们通过矩阵  $A^T A$  的特征值与特征向量就可以确定  $v_i$  和  $\sigma_i$ 。同样，计算  $AA^T$  就能确定  $u_i$ 。

下面我们来通过一个例子来熟悉上述流程：

【例】已知：  $A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$

解：

先计算  $A^T A$ ，为得到  $V$  与  $\sigma$  做准备。

$$A^T A = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix}$$

对应求其特征值，特征向量，并将特征向量标准化得到：

$$\lambda_1 = 32, \quad x_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\lambda_2 = 18, \quad x_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

接下来，我们来求  $AA^T$  来确定  $u_i$

由：

$$AA^T = U \sum V^T V \sum^T U^T = U \sum \sum^T U^T$$

又：

$$AA^T = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$$

求得其对应的特征值与特征向量为：

$$\lambda_1 = 32, \quad x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\lambda_2 = 18, \quad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

**注：**如果观察得仔细我们不难发现这里得到的特征值与之前的是一样的。事实上，存在性质：BA 与 AB 的特征值相同。

完成了各个矩阵求解的工作，对应之前提到的分解式：  $A = U \sum V^{-1} = U \sum V^T$

如果直接代入我们会得到：

$$\begin{bmatrix} 4 & 4 \\ 3 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{32} & 0 \\ 0 & \sqrt{18} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$$

这就出现了问题，之前的矩阵明明是  $\begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$ ，为什么这里计算出来的符号有问题呢？

### 出错的原因：

我们无法确定特征向量的符号，或者说，求出的一组特征向量  $U$  的反向依然满足我们求解  $U$  的过程。原本  $V$  与  $U$  的方向时有关联的，但是我们上面求解过程无法考虑这种关联。

### 解决办法：

使用最原始的公式：

$$\sigma_i u_i = A v_i$$

如果确定了  $v$  与  $\sigma$ ，直接代入，就能得到  $u$  了。这样就将  $u$  与  $v$  的关系连接了起来。

注意到，上述得到  $u$ ， $v$ ， $\sigma$  的过程都是通过变形计算得到。事实上，对于部分内容我们可以从其内容出发直接得到，下面我们通过例 2 来进行说明。

**【例 2】**  $A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$ ，对其进行奇异值分解。

分析：

易知  $A$  秩为 1，对应的行空间是一条直线，即  $(4, 3)$  的倍数，零空间是与之垂直的一条直线。

而同时， $A$  对应的列空间也是一条直线，即  $(4, 8)$  的倍数，左零空间是与之垂直的一条直线。

于是，由奇异值分解式的意义，行空间方向上的基： $\begin{bmatrix} 0.8 \\ 0.6 \end{bmatrix}$ ，零空间方向上的基： $\begin{bmatrix} -0.6 \\ 0.8 \end{bmatrix}$ 。列空间方向： $u_1 = 1/\sqrt{5} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ ，左零空间方向： $1/\sqrt{5} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ 。我们要寻找的 SVD，就是一种转换方式，将行空间与零空间上的基转换为列空间与左零空间上的基。

就可以很快地得到矩阵  $U$ :

$$U = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$$

类似的，我们可以得到：

$$V^T = \begin{bmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix}$$

余下过程与例 1 类似，通过求矩阵  $A^T A$  的特征值来得到对角矩阵  $\Sigma$ 。代入公式  $A = U \Sigma V^T$  完成最后的分解步骤。

### 【总结】

回忆这个奇异值分解的过程，我们所做的就是在线性代数的四个子空间中选出合适的基。

- 从  $v_1$  到  $v_r$  是矩阵行空间的标准正交基；
- 从  $u_1$  到  $u_r$  是矩阵列空间的标准正交基；
- 从  $v_{r+1}$  到  $v_n$  是矩阵零空间的标准正交基；
- 从  $u_{r+1}$  到  $u_n$  是对应的转置矩阵的零空间的标准正交基。

### 【进一步理解】

最后，让我们再来理解一下等式： $Av_i = \sigma_i u_i$

这表明：

- (1) 左式与右式是倍数关系
- (2)  $A$  乘以每一个  $v$  对应一个  $u$  的方向

## 三. 学习感悟

SVD 很重要，与最小二乘法也有一定的联系。这部分内容主要是将之前学习的内容进行应用与联系。其实 SVD 的用处远不止我们介绍的这些而已。它还可以减小一些计算量，进行图像变换等。