矩阵的奇异值分解(SVD)

1、回顾特征值和特征向量

首先，特征值和特征向量定义：A为实对称矩阵

$$Ax = \lambda x$$

我们求出特征值和特征向量之后，我们就可以对矩阵A进行特征分解，如下：

$$A = W\Sigma W^{-1}$$

其中，Σ=diag([λ1,...,λn])，W是n个特征向量张成的n×n维矩阵。我们一般会把n个特征向量标准化，既满足：

$$w_i^T w_i = 1$$

此时，W的n个特征向量为标准正交基(详见：高等代数，欧几里得空间)，满足：

$$W^T W = I$$

因此，W为酉矩阵。*(有关酉矩阵介绍详见：高等代数，线性变换)*。

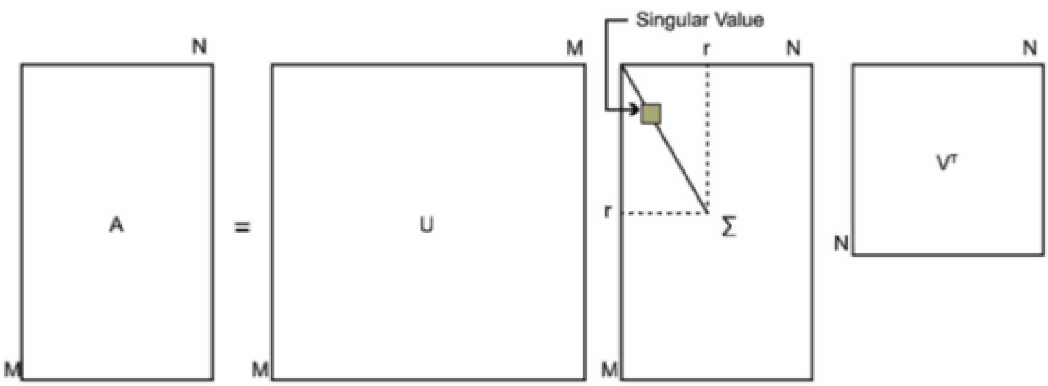
那么当A不是方阵时，怎么办？引出SVD

2、奇异值分解SVD

SVD也是对矩阵进行分解，但是和特征分解不同，SVD不要求被分解的矩阵为方阵。假设我们的矩阵A是一个m×n的矩阵，那么我们定义矩阵A的SVD为：

$$A = U\Sigma V^T$$

其中U是m×m的矩阵，Σ是m×n的矩阵，除了主对角线上的元素以外全为0，主对角线上的每个元素都称为奇异值，V是一个n×n的矩阵。U和V都是酉矩阵。



现在解决如何求出分解后的矩阵。

我们先解决如何求解奇异值的问题：

$$A = U\Sigma V^T$$

右乘V得到

$$AV = U\Sigma$$

写成分块矩阵：

$$A[v_1, \ldots, v_n] = [u_1, \ldots, u_m]\Sigma$$
$$Av_i = u_i\sigma_i$$

注意到σi是奇异值，是数。因此可以在上式右(左)乘一个uiT得到：

$$Av_iu_i^T = \sigma_i$$

我们得到了奇异值的与U、V矩阵的关系。但是现在还不知道U和V如何求，下面介绍：

$$A^TA = V\Sigma^TU^TU\Sigma V^T = V\Sigma^T\Sigma V^T$$

由分块矩阵定理：(假设m>n)

$$\Sigma^T\Sigma = \begin{bmatrix} B^T & O \end{bmatrix} \begin{bmatrix} B \\ O \end{bmatrix} = B^TB$$

这是一个对角矩阵，且是方阵。**由特征值分解可以知道：A^TA可以进行特征值分解，V则是A^TA的特征向量张成的矩阵。同理，U是AA^T的特征向量张成的矩阵。**

继续由上式，得到奇异值与A^TA(AA^T)特征值的关系：

$$\sigma_i = \sqrt{\lambda_i}$$

这样，我们就解决了SVD中的所有需求量。

3、SVD计算举例