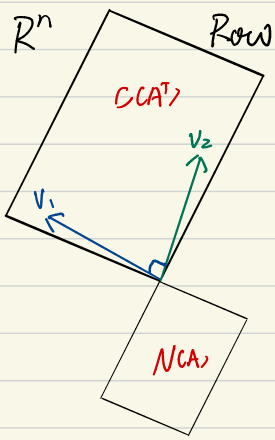
Lecture 29. Singular value decomposition. ☆ Super Important

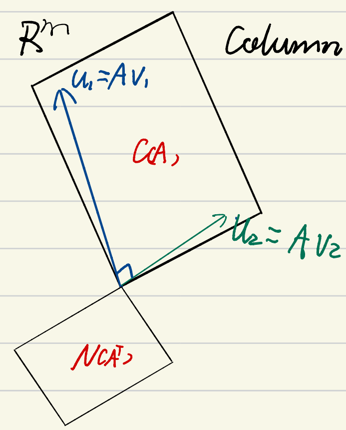
1. Overview

SVD，任意矩阵都可以分解为 $A=U\Sigma V^T$，U 为 Orthogonal Matrix 正交矩阵，Σ 是对角矩阵，V 是正交矩阵。

如果 A 正定，它的 SVD 就是 $A=Q\Lambda Q^T$，一个 Q 就搞定，这也是正定最大的优点。对于可对角化矩阵，$A=S\Lambda S^{-1}$，但特征向量矩阵 S 并不是正交矩阵，而 SVD 中 U 和 V 是正交矩阵。

2. 定义





矩阵 A 可以理解一种线性变换，行空间中一个向量 v_1 变到列空间中 $u_1=Av_1$。奇异值分解就是在行空间中找到一组正交基，将其通过矩阵 A 线性变换生成列空间中的一组正交基：

找到行空间中的一组正交基很容易，Gram-Schmidt 正交化即可。但随便一组正交基经过 A 变换后不一定正交。因此要满足此要求的行空间的正交基非常特殊。

$Av_i=\sigma_i u_i$，σ_i 是使 u 归一化因子。$\sigma_i>0$

设在行空间中找到一组特殊正交基：

$$\left.\begin{aligned} A[v_1\ v_2\ \cdots\ v_r] &= [\sigma_1 u_1\ \ \sigma_2 u_2\ \cdots\ \sigma_r u_r] \\ &= [u_1\ u_2\ \cdots\ u_r]\begin{bmatrix}\sigma_1 & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r\end{bmatrix}\end{aligned}\right\}\ AV=U\Sigma$$

如果加入 Null Space，相当

$$A[v_1\ v_2 \cdots v_r\ v_{r+1} \cdots v_n] = [\sigma_1 u_1\ \sigma_2 u_2 \cdots \sigma_r u_r\ \ \sigma_{r+1}u_{r+1} \cdots \sigma_n u_n]$$

$$= [u_1\ u_2 \cdots u_r\ \ u_{r+1} \cdots u_n]\begin{bmatrix}\sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & 0 \\ & & & & & & 0\end{bmatrix}$$

现在 $AV=U\Sigma$

已知. $AV = U\Sigma$

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

我们想求 U和V, 但无法同时求, 于是,

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

因为 A^TA 在 A 满秩时正定, 上式其实就是 A^TA 的正交分解。
V是 A^TA 的 eigenvector, σ^2 是 A 的 eigenvalue.
用同样的方法求 U:
U是 AA^T 的 eigenvector, eigenvalue 不变。

例1. $A = \begin{bmatrix}4 & 4\\ -3 & 3\end{bmatrix}$, 求SVD.

$(25-\lambda)^2 - 49 = 0$

求V. $A^TA = \begin{bmatrix}4 & -3\\ 4 & 3\end{bmatrix}\begin{bmatrix}4 & 4\\ -3 & 3\end{bmatrix} = \begin{bmatrix}25 & 7\\ 7 & 25\end{bmatrix}$, $\begin{vmatrix}25-\lambda & 7\\ 7 & 25-\lambda\end{vmatrix} = 0$

$\lambda_1 = \sigma_1^2 = 32$
$\lambda_2 = \sigma_2^2 = 18$
$x_1 = \begin{bmatrix}1\\ 1\end{bmatrix}$, $x_2 = \begin{bmatrix}1\\ -1\end{bmatrix}$ x_1 x_2 正交, 因为 A^TA 为对称.

标准化后: $v_1 = \begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$, $v_2 = \begin{bmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{bmatrix}$, $\sigma_1^2 = 32$, $\sigma_2^2 = 18$.

$$V = \begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}$$

求U $AA^T = \begin{bmatrix}32 & 0\\ 0 & 18\end{bmatrix}$, $x_1 = \begin{bmatrix}1\\ 0\end{bmatrix}$, $x_2 = \begin{bmatrix}0\\ 1\end{bmatrix}$,

标准化后, $u_1 = \begin{bmatrix}1\\ 0\end{bmatrix}$, $u_2 = \begin{bmatrix}0\\ 1\end{bmatrix}$, $\sigma_1^2 = 32$, $\sigma_2^2 = 18$.

$$U = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

但是

$$U\Sigma V^T = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\sqrt{32} & 0\\ 0 & \sqrt{18}\end{bmatrix}\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix} = \begin{bmatrix}4 & 4\\ 3 & -3\end{bmatrix} \neq A$$

因为在计算 AA^T 的 eigenvector 时, 其方向正反皆可, 无法确定其符号。但在 $A = U\Sigma V^T$ 中, V确定后, U的符号也确定了, 所以为了避免此问题, 可以把 V 直接代入 $AV = U\Sigma$ 来计算 U。

例2. $A=\begin{bmatrix}4 & 3\\ 8 & 6\end{bmatrix}$，求 SVD.

A是Singular Matrix，rank = 1，行空间/列空间是1维的。

可以确定 $v_1=\begin{bmatrix}0.8\\ 0.6\end{bmatrix}$，$v_2=\begin{bmatrix}0.6\\ -0.8\end{bmatrix}$，$u_1=\frac{1}{\sqrt{5}}\begin{bmatrix}1\\ 2\end{bmatrix}$，$u_2=\frac{1}{\sqrt{5}}\begin{bmatrix}2\\ -1\end{bmatrix}$

$A^TA=\begin{bmatrix}80 & 60\\ 60 & 45\end{bmatrix}$，$\sigma_1^2=125$，$\sigma_2^2=0$

$$\Rightarrow A=\frac{1}{\sqrt{5}}\begin{bmatrix}1 & 2\\ 2 & -1\end{bmatrix}\begin{bmatrix}\sqrt{125} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0.8 & 0.6\\ 0.6 & -0.8\end{bmatrix}$$

做SVD就是在矩阵的四个子空间中寻找到合适的基：

$v_1, v_2, \cdots v_r$ 为行空间的标准正交基.
$v_{r+1}, v_{r+2}, \cdots v_n$ 为 $N(A)$ 的标准正交基.
$u_1, u_2 \cdots u_r$ 为列空间的标准正交基.
$u_{r+1}, u_{r+2}, \cdots u_m$ 为 $N(A^T)$ 的标准正交基.

SVD的应用：因为实际问题中A通常不是满秩的，因此 A^TA 不可逆，无法分解为 $Q\Lambda Q^T$。另外，A^TA 计算量太大，用SVD会降低计算量。

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