第7章 奇异值分解 (Singular Value Decomposition, SVD)

§7.2 Bases and Matrices in the SVD

相似对角化　$A \in M_n$.　$\mathbb{R}^n \longrightarrow \mathbb{R}^n$.

$AX = X\Lambda$, i.e. $A[\underbrace{\vec{x}_1\ \vec{x}_2\ \cdots\ \vec{x}_n}_{\mathbb{R}^n\text{一组基}}] = [\underbrace{\vec{x}_1\ \vec{x}_2\ \cdots\ \vec{x}_n}_{\mathbb{R}^n\text{一组基}}]\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}$

$A\vec{x}_i = \lambda_i \vec{x}_i$, $i=1,2,\cdots,n$.　λ_i为特征值, $\vec{x}_i$为属于λ_i的特征向量.

X可逆, $AX = X\Lambda \Rightarrow A = X\Lambda X^{-1}$,　$X^{-1}AX = \Lambda$.

① A必须为方阵. a square matrix

② 不是一定能相似对角化. not always enough eigenvectors.

③ 对于一般的A, $\vec{x}_1, \cdots, \vec{x}_n$未必正交.

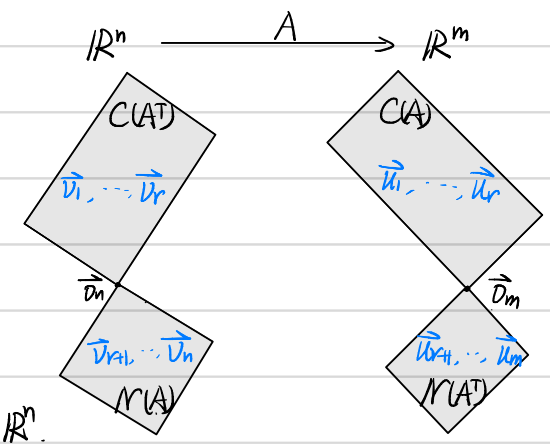
④ 若A为实对称阵, $\exists$正交阵Q s.t. $AQ = Q\Lambda \Rightarrow A = Q\Lambda Q^{-1} = Q\Lambda Q^T$.

SVD　$A \in M_{m,n}$.　$\mathbb{R}^n \longrightarrow \mathbb{R}^m$.　$\operatorname{rank}(A) = r$.

$A: \mathbb{R}^n \to \mathbb{R}^m$.　$n \neq m$时, 不可能仍有$A\vec{x} = \lambda\vec{x}$.

一般情形, 希望有: $A[\underbrace{\vec{v}_1\ \vec{v}_2\ \cdots\ \vec{v}_n}_{\mathbb{R}^n\text{一组基}}] = [\underbrace{\vec{u}_1\ \vec{u}_2\ \cdots\ \vec{u}_m}_{\mathbb{R}^m\text{一组基}}]\begin{bmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \end{bmatrix}_{m\times n}$

进一步地, $\mathbb{R}^n = C(A^T) \oplus N(A)$　$\mathbb{R}^m = C(A) \oplus N(A^T)$.



what we want:

① $\vec{v}_1, \cdots, \vec{v}_r$: orthonormal (标准正交) basis for the row space $C(A^T)$.

② $\vec{v}_{r+1}, \cdots, \vec{v}_n$: orthonormal basis for the nullspace $N(A)$.

③ $\vec{u}_1, \cdots, \vec{u}_r$: orthonormal basis for the column space $C(A)$.

④ $\vec{u}_{r+1}, \cdots, \vec{u}_m$: orthonormal basis for the left nullspace $N(A^T)$.

$C(A^T)$与$N(A)$互为$\mathbb{R}^n$中正交补 $\xRightarrow{①②}$ $\vec{v}_1, \cdots, \vec{v}_n$: orthonormal basis for $\mathbb{R}^n$.

$C(A)$与$N(A^T)$互为$\mathbb{R}^m$中正交补 $\xRightarrow{③④}$ $\vec{u}_1, \cdots, \vec{u}_m$: orthonormal basis for $\mathbb{R}^m$.

⑤ $A\vec{v}_1 = \sigma_1\vec{u}_1$, $A\vec{v}_2 = \sigma_2\vec{u}_2$, $\cdots$, $A\vec{v}_r = \sigma_r\vec{u}_r$.　$r = \operatorname{rank}(A) \leq \min(m,n)$

$(A\vec{v}_i = \vec{0}_m$ for $i = r+1, \cdots, n)$

⑥ $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$　in order of importance

remark: $\|A\vec{v}_i\| = \sigma_i \underset{1}{\underset{\uparrow}{\|\vec{u}_i\|}} = \sigma_i$, $i = 1, 2, \cdots, r$

1°. $A[\vec{v}_1 \cdots \vec{v}_r\ \vec{v}_{r+1} \cdots \vec{v}_n] = [\vec{u}_1 \cdots \vec{u}_r\ \vec{u}_{r+1} \cdots \vec{u}_m]\begin{bmatrix} \begin{matrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{matrix} & \vdots & \\ \cdots & & \cdots \\ & \vdots & \end{bmatrix}_{m\times n}$ i.e. $AV = U\Sigma$ (*)

V, U, Σ

$A: m\times n$. $V: n\times n$. $U: m\times m$, $\Sigma: m\times n$.

$\vec{v}_1, \cdots, \vec{v}_n$ 标准正交 $\Rightarrow$ V 为正交阵. $V^{-1} = V^T$.

$\vec{u}_1, \cdots, \vec{u}_m$ 标准正交 $\Rightarrow$ U 为正交阵. $U^{-1} = U^T$

$AV = U\Sigma \Rightarrow A = U\Sigma V^{-1} = U\Sigma V^T = [\vec{u}_1 \cdots \vec{u}_m]\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & \end{bmatrix}\begin{bmatrix}\vec{v}_1^T \\ \vdots \\ \vec{v}_n^T\end{bmatrix} = \sigma_1 \vec{u}_1 \vec{v}_1^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$

SVD: $A = U\Sigma V^T = \sigma_1 \vec{u}_1 \vec{v}_1^T + \sigma_2 \vec{u}_2 \vec{v}_2^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$. (full SVD)

$\vec{u}_1, \cdots, \vec{u}_m$: left singular vectors. $\vec{v}_1, \cdots, \vec{v}_n$: right singular vectors. $\sigma_1, \cdots, \sigma_r$: singular values.

(A拆分成 r 个秩一矩阵之和. Application: §7.1).

2°. (*) $\Rightarrow A[\vec{v}_1 \cdots \vec{v}_r] = [\vec{u}_1 \cdots \vec{u}_r]\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix}_{r\times r}$, i.e. $AV_r = U_r\Sigma_r$.

$V_r = [\vec{v}_1 \cdots \vec{v}_r] \in M_{n,r}$. $U_r = [\vec{u}_1 \cdots \vec{u}_r] \in M_{m,r}$, $\Sigma_r = diag(\sigma_1, \cdots, \sigma_r) \in M_r$.

$\vec{v}_1 \cdots \vec{v}_r$ are orthonormal: $V_r^T V_r = I$

$\vec{u}_1, \cdots, \vec{u}_r$ - - - - - - - : $U_r^T U_r = I$

reduced SVD: $AV_r = U_r\Sigma_r$

SVD的计算

we want: $A = U\Sigma V^T$

$A^TA = (V\Sigma^T U^T)(U\Sigma V^T) = V\Sigma^T(\underbrace{U^TU}_{I})\Sigma V^T = V(\Sigma^T\Sigma)V^T$ (**)

where $\begin{cases}\Sigma^T\Sigma = diag(\sigma_1^2, \cdots, \sigma_r^2, 0, \cdots, 0) \in M_n & \text{对角阵} \\ V \text{为正交阵}. \quad V^T = V^{-1}\end{cases}$

(**) 即 A^TA 的相似对角化. $A^TA = V(\Sigma^T\Sigma)V^{-1}$.

A^TA: 实对称, 一定可由正交阵 V 相似对角化. V: eigenvector matrix for A^TA.

$\left.\begin{array}{l} A^TA: \text{至少是半正定的, 特征值全} \geq 0. \\ rank(\Sigma^T\Sigma) = rank(A^TA) = rank(A) = r. \end{array}\right\} \Rightarrow$ A^TA 有 r 个正的特征值. $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_r^2 > 0$. 可能有重复的.

(当 $r=n$ 时, A列满秩, A^TA 为正定矩阵)

Step 1 求 A^TA 的特征值 $\overset{\text{可能有重复}}{\sigma_1^2}, \sigma_2^2, \cdots, \sigma_r^2, \overbrace{0, \cdots, 0}^{n-r\text{个}}$，并按降序排列：$\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_r^2 > 0$.

从而解得 A 的 singular value：$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ (全为正).

Step 2 求 A^TA 的标准正交 (orthonormal) 特征向量：$\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n$.

s.t. $(A^TA)\vec{v}_i = \sigma_i^2 \vec{v}_i$, $i=1,\cdots,r$. $\quad (A^TA)\vec{v}_i = 0\vec{v}_i = \vec{0}$, $i=r+1,\cdots,n$. (如有)

$\Rightarrow \vec{v}_i = \frac{1}{\sigma_i^2} A^T(A\vec{v}_i) \in C(A^T)$, $i=1,\cdots,r$. $\quad$ orthonormal.

$\vec{v}_i \in N(A^TA) = N(A)$, $i=r+1,\cdots,n$ $\quad$ orthonormal. $\quad A\vec{v}_i = \vec{0}$, $i=r+1,\cdots,n$.

Step 3 令 $A\vec{v}_i = \sigma_i \vec{u}_i$，得 $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$, $i=1,\cdots,r$.

$\vec{u}_i \in C(A)$, $i=1,\cdots,r$.

$$\vec{u}_i^T \vec{u}_j = \frac{1}{\sigma_i\sigma_j}(A\vec{v}_i)^T(A\vec{v}_j) = \frac{1}{\sigma_i\sigma_j}\vec{v}_i^T(A^TA)\vec{v}_j = \frac{\sigma_j^2}{\sigma_i\sigma_j}\overset{\delta_{ij}\ (\vec{v}_1,\cdots,\vec{v}_r\text{ orthonormal})}{\vec{v}_i^T\vec{v}_j} = \delta_{ij}. \quad i,j=1,2,\cdots,r.$$

$\therefore \vec{u}_1, \cdots, \vec{u}_r$ are orthonormal.

Step 4 求 $N(A^T)$ 的一组标准正交基 $\vec{u}_{r+1}, \cdots, \vec{u}_m$.

Remark: (1) 可验证以上求出的解满足①～⑥.

(2). $A\vec{v}_i = \sigma_i \vec{u}_i$, $i=1,\cdots,r \Rightarrow (A^TA)\vec{v}_i = \sigma_i(A^T\vec{u}_i)$, $i=1,\cdots,r$

$\Rightarrow \sigma_i^2 \vec{v}_i = \sigma_i(A^T\vec{u}_i) \Rightarrow A^T\vec{u}_i = \sigma_i\vec{v}_i$, $i=1,2,\cdots,r$

(3) $(AA^T)\vec{u}_i = A(A^T\vec{u}_i) = \sigma_i(A\vec{v}_i) = \sigma_i^2\vec{u}_i$, $i=1,\cdots,r$.

σ_i^2, $i=1,\cdots,r$ 也为 AA^T 的特征值. (P317 $A\in M_{m,n}$, $B\in M_{n,m}$, AB 与 BA 有相同的非零特征值)

$\vec{u}_i$ 为 AA^T 的属于 σ_i^2 的特征向量. $i=1,\cdots,r$.

(4) $(AA^T)\vec{u}_i = A(A^T\vec{u}_i) = A\vec{0} = \vec{0}$, $i=r+1,\cdots,m$.

$\vec{u}_i$ 为 AA^T 的属于 0 的特征向量 $i=r+1,\cdots,m$ (如有)

矩阵形式

(5) $A = U\Sigma V^T \Rightarrow A^T = V\Sigma^T U^T$

$\xRightarrow{U^T=U^{-1}} A^TU = V\Sigma^T$, $\quad A^T[\vec{u}_1 \cdots \vec{u}_r\ \vec{u}_{r+1} \cdots \vec{u}_m] = [\vec{v}_1 \cdots \vec{v}_r\ \vec{v}_{r+1} \cdots \vec{v}_n]\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}_{n\times m}$

$(A^T\vec{u}_i = \sigma_i\vec{v}_i,\ i=1,\cdots,r.)$

(6) $AA^T = U\Sigma(V^TV)\Sigma^TU^T = U(\Sigma\Sigma^T)U^T = U\overset{\text{对角阵}}{(\Sigma\Sigma^T)}U^{-1}$

(σ_i^2, $i=1,\cdots,r$ 为 AA^T 的非零特征值，U 为 AA^T 的 eigenvector matrix).

另一种解法：求 AA^T 特征值 $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_r^2 > 0$. 及对应特征向量 $\vec{u}_1, \cdots, \vec{u}_r$, $\overbrace{\vec{u}_{r+1}, \cdots, \vec{u}_m}^{(AA^T)\vec{u}_i = \vec{0},\ i=r+1,\cdots,m\ (\Rightarrow A^T\vec{u}_i = \vec{0})}$, orthonormal.

令 $A^T\vec{u}_i = \sigma_i\vec{v}_i$，即 $\vec{v}_i = \frac{1}{\sigma_i}A^T\vec{u}_i$, $i=1,\cdots,r$.

取 $N(A)$ 的标准正交基 $\vec{v}_{r+1}, \cdots, \vec{v}_n$.

例. 求 $A=\begin{bmatrix}1&0&1&0\\0&1&0&1\end{bmatrix}$ 的奇异值分解.

解. $A: 2\times4$. $r=\text{rank}(A)=2$,

$AA^T=\begin{bmatrix}2&0\\0&2\end{bmatrix}$, $A^TA=\begin{bmatrix}1&0&1&0\\0&1&0&1\\1&0&1&0\\0&1&0&1\end{bmatrix}$

① AA^T 的特征值为 $\lambda_1=\lambda_2=2$. (从而 A^TA 的特征值为 2 和 0, 代数重数均为 2).

A 的奇异值为 $\sigma_1=\sqrt{\lambda_1}=\sqrt2$. $\sigma_2=\sqrt{\lambda_2}=\sqrt2$.

② 求 AA^T 的标准正交特征向量 $\vec u_1,\vec u_2\in N(2I-AA^T)=R^2$,

这里可取 R^2 中任意一组标准正交基. 如 $\vec u_1=\begin{bmatrix}1\\0\end{bmatrix}$, $\vec u_2=\begin{bmatrix}0\\1\end{bmatrix}$, 令 $U=[\vec u_1\ \vec u_2]=\begin{bmatrix}1&0\\0&1\end{bmatrix}$

③ 利用公式求得 $\vec v_1=\frac{1}{\sigma_1}A^T\vec u_1=\frac{1}{\sqrt2}\begin{bmatrix}1\\0\\1\\0\end{bmatrix}$, $\vec v_2=\frac{1}{\sigma_2}A^T\vec u_2=\frac{1}{\sqrt2}\begin{bmatrix}0\\1\\0\\1\end{bmatrix}$.

④ 求 $N(A)$ 的一组标准正交基

$\vec v_3=\frac{1}{\sqrt2}\begin{bmatrix}-1\\0\\1\\0\end{bmatrix}$, $\vec v_4=\frac{1}{\sqrt2}\begin{bmatrix}0\\-1\\0\\1\end{bmatrix}$, 令 $V=[\vec v_1\ \vec v_2\ \vec v_3\ \vec v_4]=\frac{1}{\sqrt2}\begin{bmatrix}1&0&-1&0\\0&1&0&-1\\1&0&1&0\\0&1&0&1\end{bmatrix}$

奇异值分解: $A=U\Sigma V^T$, 其中 $\Sigma=\begin{bmatrix}\sqrt2&0&0&0\\0&\sqrt2&0&0\end{bmatrix}$.

注: 也可通过先将 A^TA 对角化来得到 V, 再求 U.

例. When is $A=U\Sigma V^T=U\Sigma V^{-1}$ the same as $X\Lambda X^{-1}$?

解. $X=U=V \Rightarrow A$ needs orthonormal eigenvectors. $\Rightarrow X=Q$. $Q^{-1}=Q^T$.

$A=Q\Lambda Q^{-1}=Q\Lambda Q^T \Rightarrow A^T=A$. A is symmetric. real.

$\Lambda=\Sigma \Rightarrow A$ needs eigenvalues $\lambda\geq0$. $\Rightarrow A$ is positive semidefinite (or definite)

例. P373 – P374. 自学例题.

$A=\begin{bmatrix}0&1&0&0\\0&0&2&0\\0&0&0&3\\0&0&0&0\end{bmatrix}=U\Sigma V^T$. $U=\begin{bmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{bmatrix}$, $\Sigma=\begin{bmatrix}3&&&\\&2&&\\&&1&\\&&&0\end{bmatrix}$, $V=\begin{bmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{bmatrix}$

$=3\vec u_1\vec v_1^T+2\vec u_2\vec v_2^T+1\vec u_3\vec v_3^T$, $3\vec u_1\vec v_1^T$: picks out the biggest number $A_{34}=3$