Lecture 25 Singular Value Decomposition (SVD)

Recall for linear map $T: V \to W$.

$\dim(V) = n$, $\dim W = m$. If $Tu = Au$, $A \in \mathbb{R}^{m\times n}$

wrt the standard basis of $V = \mathbb{R}^n$, $W = \mathbb{R}^m$

T has A as its matrix representation.

If we do a basis change for V and W with new bases $B_{in} \in \mathbb{R}^{n\times n}$ and $B_{out} \in \mathbb{R}^{m\times m}$, both nonsingular

The matrix representation for T becomes

$$B_{out}^{-1} A B_{in}$$

SVD says for any $A \in \mathbb{R}^{m\times n}$, we can find orthogonal B_{in} and B_{out} s.t.

$$B_{out}^{-1} A B_{in} \text{ is (quasi)-diagonal}$$

<u>Th</u>. (SVD). Let. $A \in \mathbb{R}^{m\times n}$, there are orthogonal matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$, s.t.

$$AV = U\Sigma, \quad \text{where } \Sigma = \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right) \in \mathbb{R}^{m\times n}$$

where $r = \text{rank}(A)$, and $\sigma_1 \geq \cdots \geq \sigma_r > 0$.

We given two different proofs, one for A nonsingular, and one for general A.

First, we look at a $n=2$ example

Example. Consider $A \in \mathbb{R}^{m\times 2}$, SVD $\Leftrightarrow$

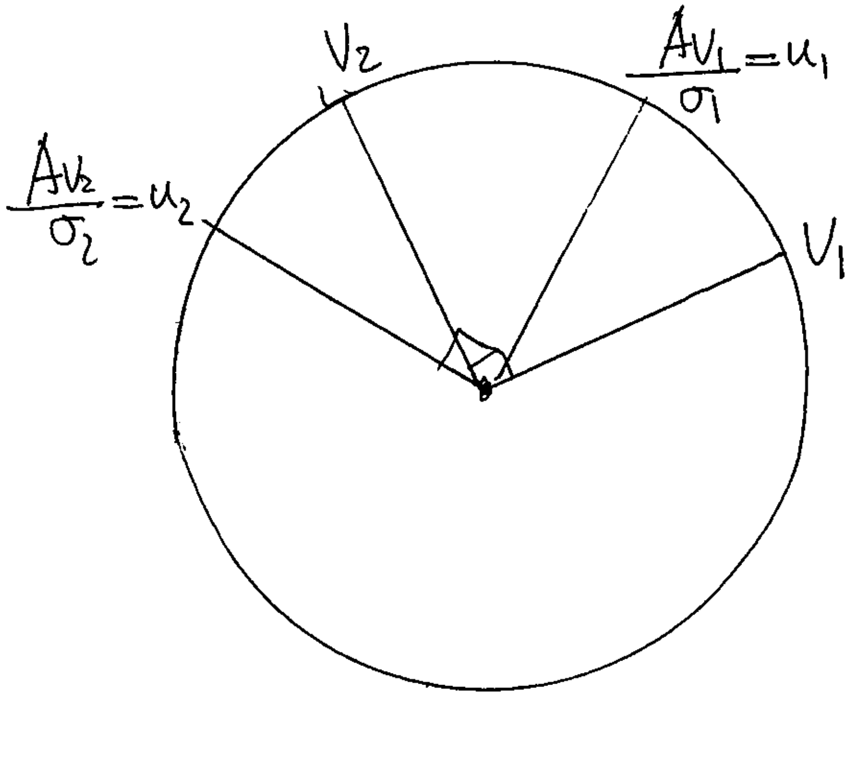
find orthonormal $v_1, v_2 \in \mathbb{R}^2$, s.t.

Av_1, Av_2 are still orthogonal to each other

$Av_1 = \sigma_1 \underline{u_1}$, $Av_2 = \sigma_2 \underline{u_2}$, $\quad u_1, u_2$ orthonormal

such v_1 and v_2 must be eigenvectors of.

$A^T A \in \mathbb{R}^{2\times 2}$



let $A \in R^{n \times n}$ be nonsingular, consider

$A^T A$ which is symetric and if u is

an eigenvector, we have

$A^T A u = \lambda u$. dot-prod with u^T

on both sides

$$u^T A^T A u = \lambda u^T u.$$

$$\lambda = \frac{(Au)^T Au}{u^T u}$$

Since $u \neq 0$ as an eigenvector, $u^T u = \|u\|^2 > 0$

$u \neq 0 \Rightarrow Au \neq 0$ since A is nonsingular

$$(Au)^T Au = \|Au\|^2 > 0 \Rightarrow \lambda > 0$$

Therefore all eigenvalues of A^TA are positive, let them be $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2 > 0$

<u>Spectral</u> decomposition of A^TA gives an <u>orthogonal</u> matrix $V \in \mathbb{R}^{n \times n}$

$$V^T(A^TA)V = \begin{pmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \\ & & & \sigma_n^2 \end{pmatrix}$$

let $U = AV \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_n} \end{pmatrix}$ $(*)$

Then $$U^TU = \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_n} \end{pmatrix} \underbrace{V^TA^TAV} \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_1} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_n} \end{pmatrix} \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_1} \end{pmatrix} = I_n$$

i.e, U is an <u>orthogonal</u> matrix.

$(*)$ implies $\quad AV = U \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} \quad r(A) = n.$

For general $A \in \mathbb{R}^{m \times n}$, we use induction

(σ_1^2, v_1) with $\|v_1\| = 1$

let ~~v_1~~ be an eigenpair of $A^T A$, and $\sigma_1 > 0$

$A^T A v_1 = \sigma_1^2 v_1$, define $u_1 = \dfrac{A v_1}{\sigma_1}$

$\|u_1\| = \dfrac{\|A v_1\|}{\sigma_1} = 1.$

$$= \sigma_1^2 \|v_1\|^2$$

$\Big($ since $v_1^T A^T A v_1 = \sigma_1^2 v_1^T v_1 = \sigma_1^2$

$$\underbrace{\left(\frac{A v_1}{\sigma_1}\right)^T}_{u_1^T} \underbrace{\left(\frac{A v_1}{\sigma_1}\right)}_{u_1} = 1 \Big)$$

(*) $A v_1 = \sigma_1 u_1.$, $\underbrace{A^T A v_1}$ $= \sigma_1 A^T u_1$

$\sigma_1^2 v_1 = \sigma_1 A^T u_1 \Rightarrow A^T u_1 = \sigma_1 v_1.$ (**)

v_1 eigenvector of $A^T A$

expand u_1, v_1 into orthogonal matrices

$\mathcal{U}_1 = [u_1, *] \in \mathbb{R}^{m\times m}$ orthogonal

$V_1 = [v_1, *] \in \mathbb{R}^{n\times n}$ orthogonal

From $\mathcal{U}_1^T \mathcal{U}_1 = I_m$, we obtain $\mathcal{U}_1^T u_1 = e_1 \in \mathbb{R}^m$

$V_1^T V_1 = I_n$ " " $V_1^T v_1 = e_1 \in \mathbb{R}^n$

Now, $AV_1 = A[v_1, *] = [Av_1, *]$

$$\overset{\text{from}(*)}{=} [\sigma_1 u_1, *] = \underbrace{\mathcal{U}_1 \mathcal{U}_1^T}_{=I_m} [\sigma_1 u_1, *]$$

$$= \mathcal{U}_1 [\sigma_1 \mathcal{U}_1^T u_1, *] = \mathcal{U}_1 [\sigma_1 e_1, *]$$

Similarly, $A^T \mathcal{U}_1 = A^T [u_1, *] = [A^T u_1, *]$

$$\overset{\text{from}(**)}{=} [\sigma_1 v_1, *] = \underbrace{V_1 V_1^T}_{=I_n} [\sigma_1 v_1, *]$$

$$= V_1 [\sigma_1 V_1^T v_1, *] = V_1 [\sigma_1 e_1, *]$$

from $A^T U_1 = V_1 [\sigma_1 e_1, *]$

take transpose of both sides

$$U_1^T A = \begin{bmatrix} \sigma_1 e_1^T \\ * \end{bmatrix} V_1.$$

$$\Rightarrow \quad A V_1 = U_1 \begin{bmatrix} \sigma_1 e_1^T \\ * \end{bmatrix}$$

Combined with $A V_1 = U_1 [\sigma_1 e_1, *]$

$$\Rightarrow \quad A V_1 = U_1 \left(\begin{array}{c|c} \sigma_1 & 0 \\ \hline 0 & A_1 \end{array}\right), \quad A_1 \in R^{(m-1)\times(n-1)}$$

since $\sigma_1 > 0$, and U_1, V_1 nonsingular

$$\operatorname{rank}(A_1) = \operatorname{rank}(A) - 1 = r - 1.$$

By induction assumption, there are orthogonal $U_2 \in \mathbb{R}^{(m-1)\times(m-1)}$, $V_2 \in \mathbb{R}^{(n-1)\times(n-1)}$,

$$A_1 V_2 = U_2 \left(\begin{array}{ccc|c} \sigma_2 & & & \\ & \ddots & & \\ & & \sigma_r & 0 \\ \hline & 0 & & 0 \end{array}\right)$$

Now

$$A V_1 = U_1 \left(\begin{array}{c|c} \sigma_1 & 0 \\ \hline 0 & U_2 \left(\begin{array}{ccc|c} \sigma_2 & & & \\ & \ddots & & \\ & & \sigma_r & 0 \\ \hline & 0 & & 0 \end{array}\right) V_2^T \end{array}\right)$$

$$= U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & 0 \\ \hline & 0 & & 0 \end{array}\right) \begin{pmatrix} 1 & 0 \\ 0 & V_2^T \end{pmatrix}$$

$$\underbrace{A V_1 \begin{pmatrix} 1 & 0 \\ 0 & V_2 \end{pmatrix}}_{V} = \underbrace{U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}}_{U.} \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & 0 \\ \hline & 0 & & 0 \end{array}\right)$$

U and V are orthogonal, and

$$AV = U \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & 0 \\ \hline & 0 & & 0 \end{array}\right). \quad \blacksquare$$

Example Consider an rank-1 matrix

$$A = uv^T \qquad u \in R^m, \; v = R^n$$

$u \neq 0$ and $v \neq 0$.

Now $A^TA = (vu^T)(uv^T) = (u^Tu)\, vv^T$

Notice that $u^Tu \in R$ is a scalar, and

$vv^T \in R^{n\times n}$ is a matrix

let $v_1 = v/\|v\|$, then $\|v_1\| = 1$, and we

expand it into an orthogonal matrix $\cancel{U = [u_1, *]}$

$$V = [v_1, \cancel{*}\, V_1] \in R^{n\times n} \qquad V_1 \in R^{n\times(n-1)}$$

because $V^TV = I$. $\quad V^Tv_1 = e_1$ and $v_1^TV = e_1^T$

or $v^TV = \|v\| e_1^T$

$$V^TA^TAV = (u^Tu)\, V^Tv\, v^TV$$
$$= u^Tu(\|v\|e_1^T)^T\|v\|e_1^T$$

$$= (u^T u)(v^T v) e_1 e_1^T$$

$$= \begin{pmatrix} (u^T u)(v^T v) & & \\ & 0 & \\ & & \ddots \\ & & & 0 \end{pmatrix} \in R^{n\times n}$$

Therefore, the eigenvalues of $A^T A$ are

$$(u^T u)(v^T v), \underbrace{0, \cdots, 0}_{n-1}$$

Similarly, the eigenvalues of AA^T are

$$(u^T u)(v^T v), \underbrace{0, \cdots, 0}_{m-1}.$$

$$U = [u_1, U_1] \in R^{m\times m}, \quad u_1 = u/\|u\|$$

$$U^T (AA^T) U = \begin{pmatrix} (u^T u)(v^T v) & & \\ & 0 & \\ & & \ddots \\ & & & 0 \end{pmatrix} \in R^{m\times m}$$

and

$$U^T A V = U^T (uv^T) V = (u^T U)^T (v^T V)$$

$$= (\|u\| e_1^T)^T \|v\| e_1^T = \|u\| \|v\| e_1 e_1^T.$$

$$= \begin{pmatrix} \|u\| \|v\| & & \\ & 0 & \\ & & \ddots & \\ & & & 0 \end{pmatrix} \qquad \text{rank}(A) = 1.$$

gives the SVD of $A = uv^T$