Lec 29. Warm up.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Find $\vec{u} \in \mathbb{R}^5$, $\vec{v} \in \mathbb{R}^2$ s.t. $A = \vec{u}\,\vec{v}^T$

$$\vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\vec{u}\,\vec{v}^T = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} = A.$$

$A \in \mathbb{R}^{m \times n}$ if there exists $\vec{u} \in \mathbb{R}^m$,

$\vec{v} \in \mathbb{R}^n$ s.t.

$$A = \vec{u}\,\vec{v}^T \qquad A\vec{x} = \vec{u}\left(\vec{v}^T\vec{x}\right)$$

A is rank-1 matrix.

storage cost of A : mn

$\Downarrow$

of $\underbrace{\vec{u}, \vec{v}}_{}$: $m+n$

reduced storage cost

singular value decomposition (SVD)

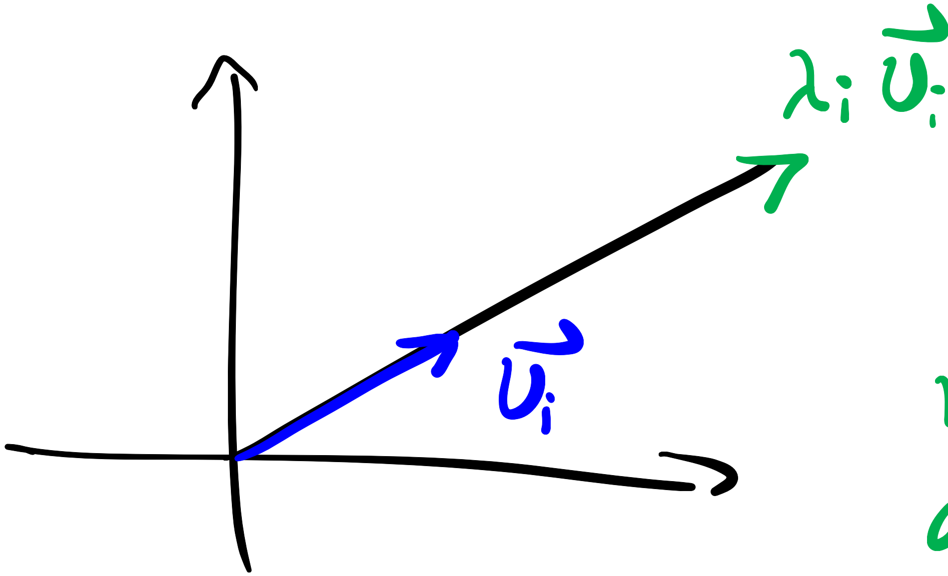
$\updownarrow$

eigenvalue decomposition $\Longleftrightarrow$ diagonalization

$A \in \mathbb{R}^{n \times n}$ square matrix

$AV = VD$. V invertible. $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

$A\vec{v_i} = \lambda_i \vec{v_i}$



only scalar multiplication along $\vec{v_i}$

Two drawbacks :

① A may not diagonalizable e.g. $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

② $A \in \mathbb{R}^{m \times n}$ $(m \neq n)$. useless.

SVD overcomes the problems above.

Applicable to ANY matrix.

$A \in \mathbb{R}^{m \times n}$ WLOG. $m \geq n$

otherwise just consider A^T.

Consider

$A^T A \in \mathbb{R}^{n \times n}$ real sym.

$\Rightarrow$ spectral decomp.

① $A^T A V = V D$

Fact

$V^T V = I_n$

$$D = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix} \geq 0.$$

$$\Sigma = \text{“} D^{\frac{1}{2}} \text{”} = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$$

$\sigma_i \geq 0.$ for simplicity $\sigma_i > 0.$

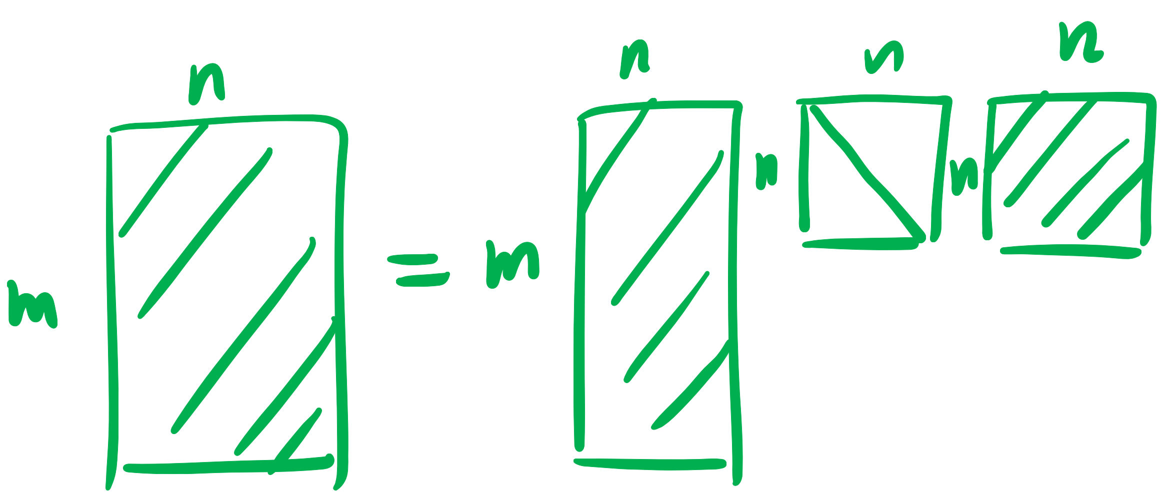
② Define $U = A V \Sigma^{-1}$

$$\Rightarrow \boxed{A = U \Sigma V^T} \quad \text{SVD.}$$

(a) $V^T V = I_n$. $V \in \mathbb{R}^{n \times n}$

(b) $\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$, $\sigma_i \geq 0$.

(c) $U^T U = I_n$. $U \in \mathbb{R}^{m \times n}$



why (c) holds

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

$$= \Sigma^{-1} \cancel{V^T (V} \Sigma^2 \cancel{V^T) V} \Sigma^{-1}$$

$$= \Sigma^{-1} \cdot \Sigma^2 \Sigma^{-1}$$

$$= I_n$$

why $D = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix}$.

quadratic form. for any $\vec{x} \in \mathbb{R}^n$

$$\vec{x}^T A^T A \vec{x} = (A\vec{x})^T A\vec{x}$$

$$= \|A\vec{x}\|^2 \geq 0$$

Fact: $B \in \mathbb{R}^{n\times n}$. sym.

$$\vec{x}^T B \vec{x} \geq 0. \text{ Then}$$

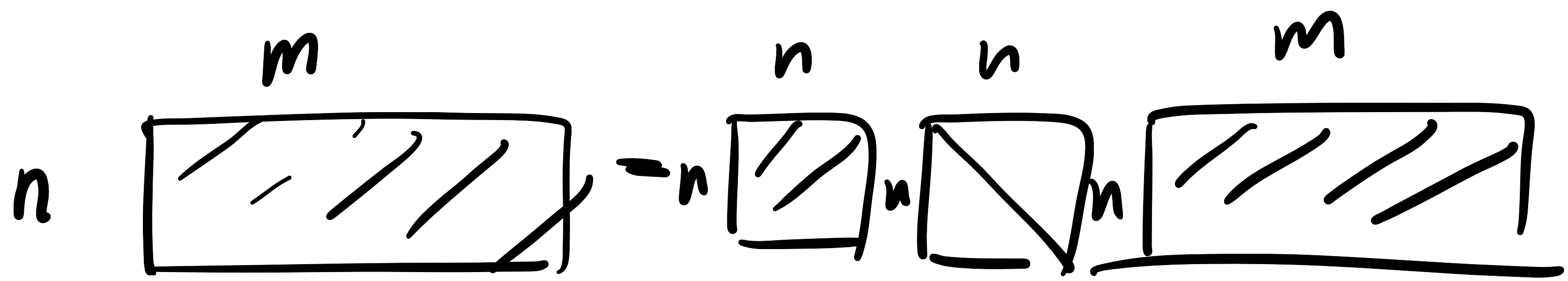
$$\lambda_i(B) \geq 0. \quad i = 1, \cdots, n.$$

Take $\vec{x} = \vec{v_i}$

$$\lambda_i(B) = \vec{v_i}^T B \vec{v_i} \geq 0.$$

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

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



Compare w. eigen-decomp.

$$A\vec{v}_i = \lambda_i \vec{v}_i$$

Now $A \in \mathbb{R}^{m \times n}$

$$U = [\vec{u}_1, \dots, \vec{u}_n]$$

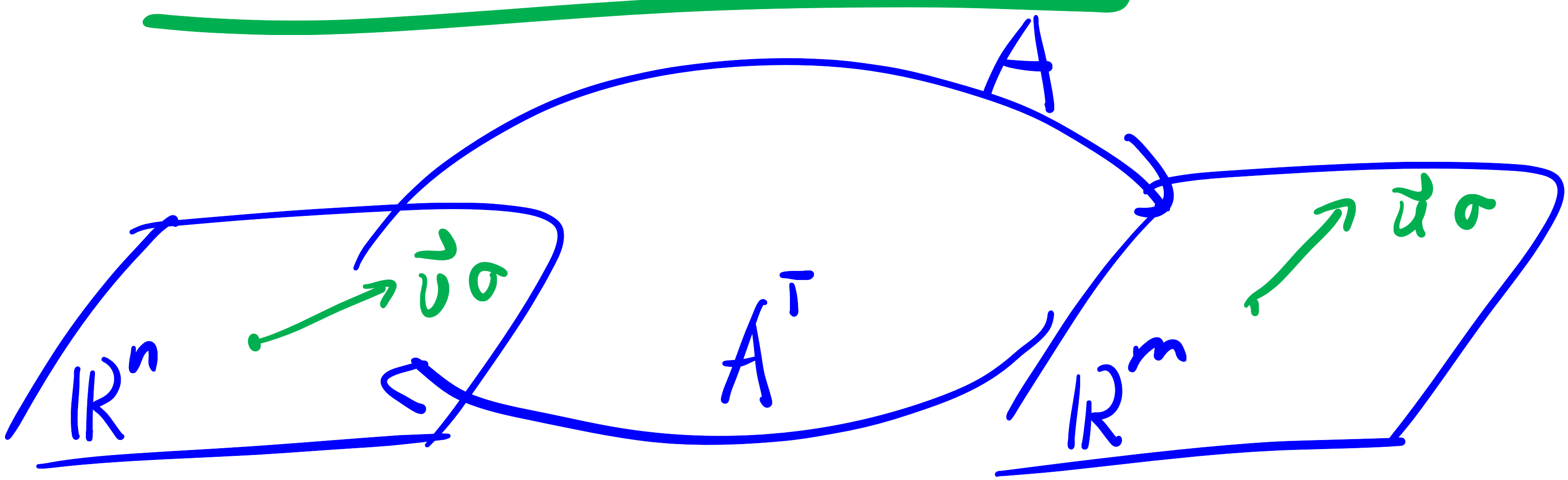
$$V = [\vec{v}_1, \dots, \vec{v}_n]$$

$$AV = U\Sigma$$

$$\Rightarrow \boxed{A\vec{v}_i = \sigma_i \vec{u}_i}$$

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

$$\Rightarrow \boxed{A^T\vec{u}_i = \sigma_i \vec{v}_i}$$



Compression. Use SVD.

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

$$= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix}$$

$$= \sum_{i=1}^{n} \vec{u}_i \, \sigma_i \, \vec{v}_i^T$$

order $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$

In many applications,

$$\sigma_i < \epsilon, \quad \text{for all } i \geq k.$$

$$A_k := \sum_{i=1}^{k} \vec{u}_i \, \sigma_i \, \vec{v}_i^{\,T}$$

rank k-compression.

$A - A_k$ is "small"

"best" rank-k approx. to A.

Ex. Compute SVD of

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

① $A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$

$= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

$A^TAV = V\Sigma^2$

$\xrightarrow{\text{ch. 5}} \quad V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \ \Sigma^2 = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$

② $U = AV\Sigma^{-1}$

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

$$= \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \sqrt{\frac{2}{3}} & 0 \\ \sqrt{\frac{1}{6}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$U^T U = I_2$

$$A = U \Sigma V^T$$ "thin-SVD"

3×2 $\quad 3 \times 2 \quad 2 \times 2 \quad 2 \times 2$

$\uparrow$

U is not an orthogonal matrix (has orthogonal columns, but w. missing cols).

How to make U an ortho. matrix? In this case

$3 \times 2 \longrightarrow 3 \times 3$

$$\tilde{U} = [\vec{u}_1 \ \vec{u}_2 \ \ \vec{u}_3]$$

use G-S.

$$A = \tilde{U} \tilde{\Sigma} V^T$$

$$\underset{3\times 2}{A} \qquad \underset{3\times 3}{\tilde{U}} \quad \underset{3\times 2}{\tilde{\Sigma}} \quad \underset{2\times 2}{V^T}$$

$$= \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \end{bmatrix}$$

$$= \vec{u}_1 \sigma_1 \vec{v}_1^T + \vec{u}_2 \sigma_2 \vec{v}_2^T$$

"full-SVD" $m \geq n$

$$A = \tilde{U} \tilde{\Sigma} V^T$$

$m \times n$ $\quad$ $m \times m$ $\quad$ $m \times n$ $\quad$ $n \times n$

$$\tilde{U} = [\vec{u}_1, \cdots, \vec{u}_n, \vec{u}_{n+1}, \cdots, \vec{u}_m] \qquad \tilde{\Sigma} = \begin{array}{c} \\ n \\ m-n \end{array} \begin{array}{c} n \\ \left[\begin{array}{c} \Sigma \\ \hline 0 \end{array}\right] \end{array}$$