# Linear Algebra for Computer Science

## Lecture 28

## SVD (cont.)

# Remember: SVD

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

$A$: $m \times n$

$U$: orthogonal $m \times m$

$\Sigma$: diagonal $m \times n$

$V^T$: orthogonal $n \times n$

$$UU^T = U^TU = I_m$$

$$VV^T = V^TV = I_n$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_{\min(m,n)} \end{bmatrix}$$

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_{\min(m,n)} \geq 0$$

$U = [u_1 \ u_2 \cdots u_m] \rightarrow$ left singular vectors (منفرد)

$V = [v_1 \ v_2 - v_n] \rightarrow$ right singular vectors

$\sigma_1 \ \sigma_2 \cdots \sigma_{\min(m,n)}$ singular values

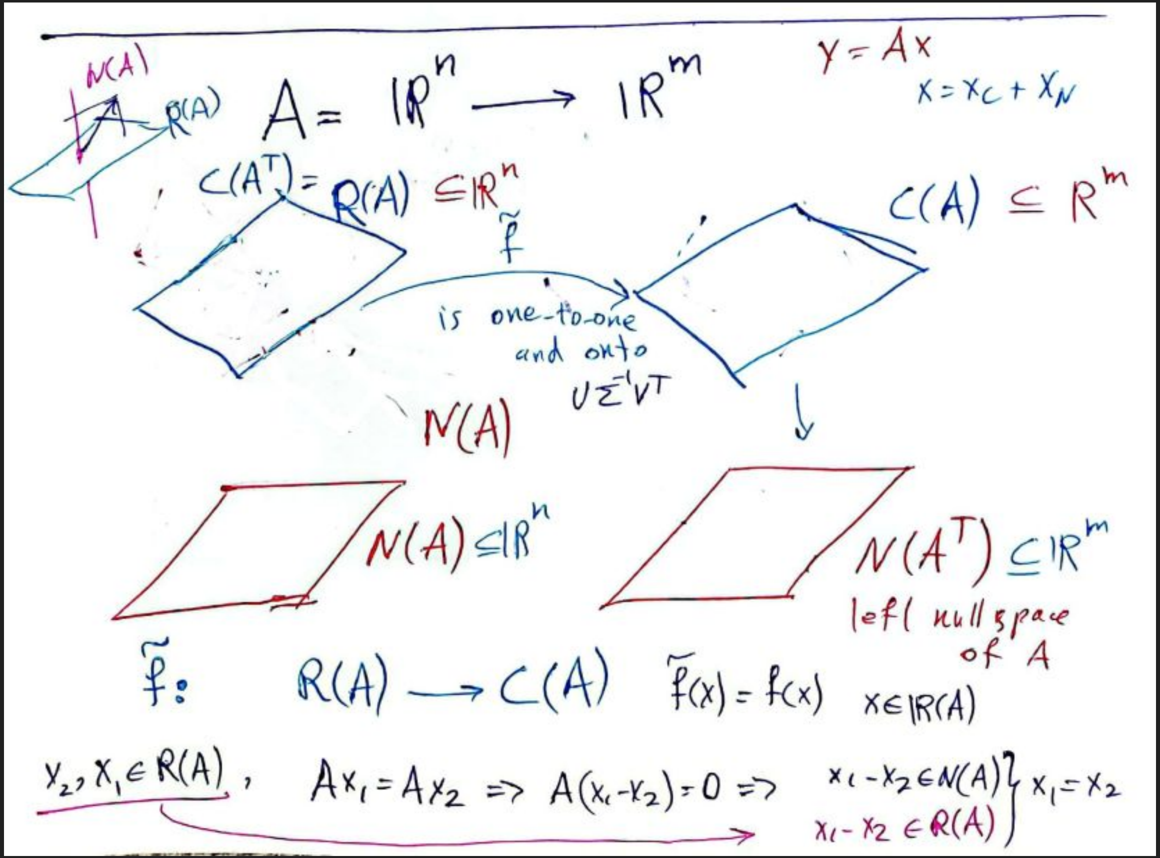
$$A v_i = \sigma_i u_i$$

$$A = [u_1 \ u_2 \cdots u_m] \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_l \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}$$

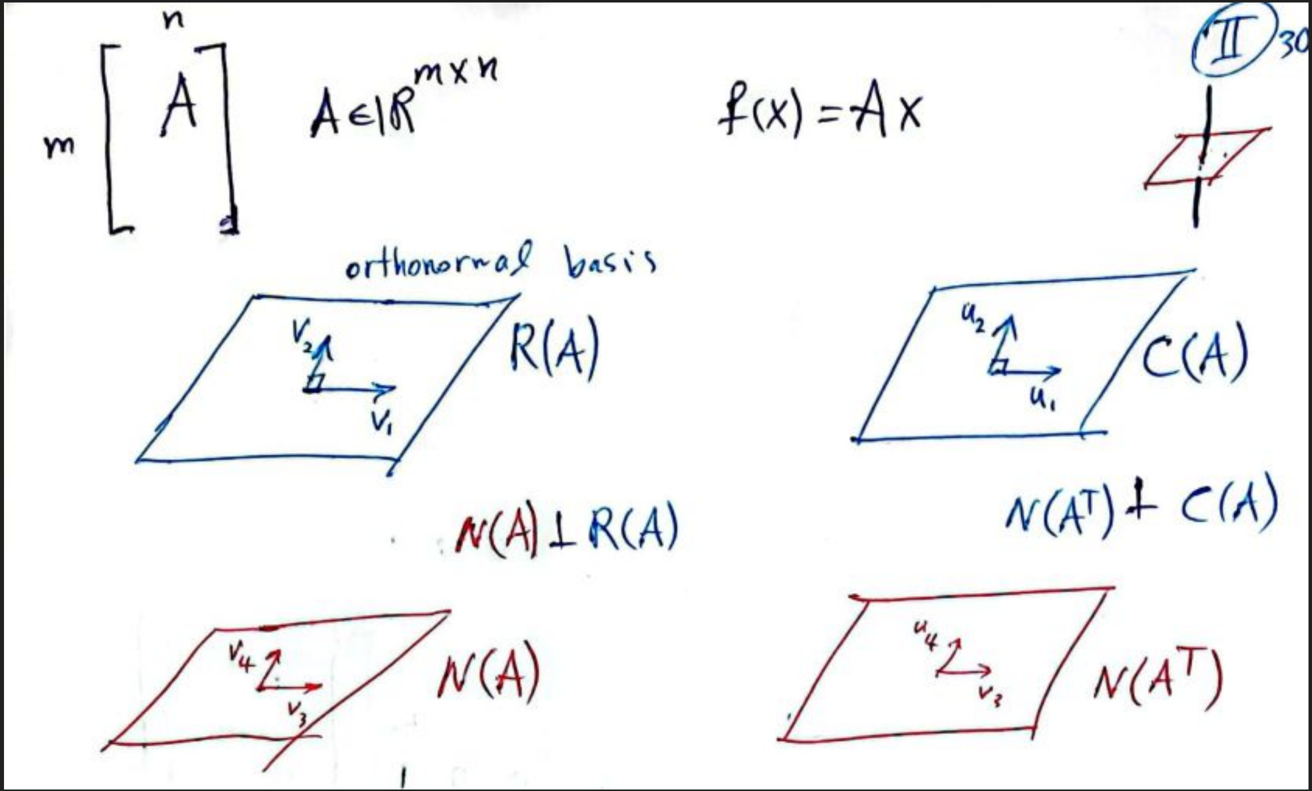
$$= \sum_{i=1}^{\min(m,n)} \sigma_i u_i v_i^T = U \Sigma V^T$$

# Bijective Map between Row space and Column Space

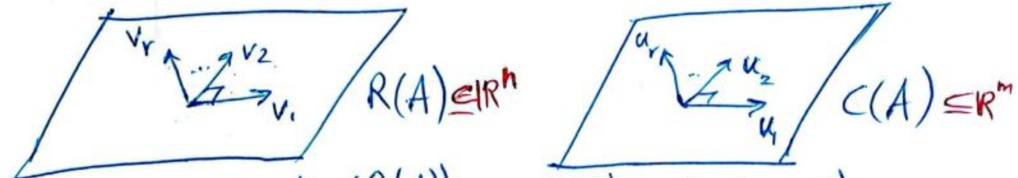
K. N. Toosi
University of Technology



# Choose a basis



# Choose a basis




$$r = \dim(R(A)) = \operatorname{rank}(A) = \dim(C(A))$$


$v_1, v_2, \dots, v_r$ is an orthonormal basis for $R(A)$

$u_1, u_2, \dots, u_1$ form an orthonormal bassis for $C(A)$

$v_{r+1}, \dots, v_n$ form an orthonormal basis for $N(A)$

$u_{r+1}, \dots, u_m$ form an " " " $N(A^T)$

$v_1, v_2, \dots, v_n$ " " " " " $\mathbb{R}^n$

$u_1, u_2, \dots, u_m$ " " " " " $\mathbb{R}^m$

# Singular Values and Singular Vectors



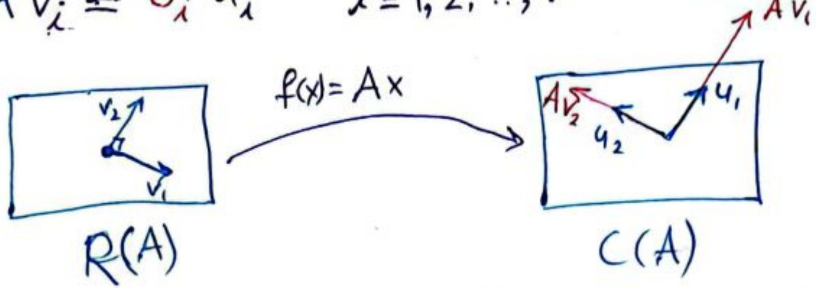
for a matrix $A \in \mathbb{R}^{m \times n}$ choose $v_1, v_2, \dots v_r$ and (III) 30

$u_1, u_2, \dots, u_r$ such that

$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_r$

$\sigma_i \geq 0$

$$A v_i = \sigma_i u_i \qquad i = 1, 2, \dots, r$$




$\sigma_1, \sigma_2, \dots, \sigma_r$ are called singular values

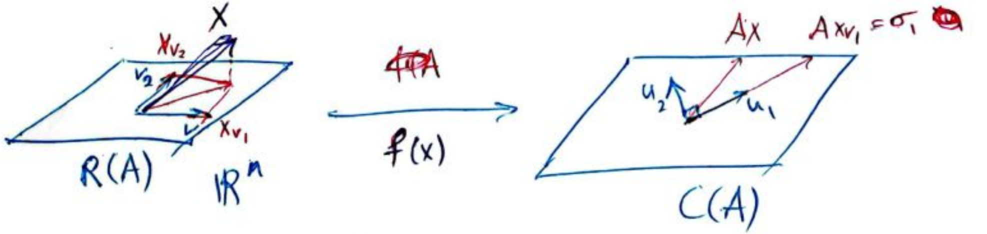
مقادیر منفرد

$v_1, v_2, \dots, v_r$ : right singular vectors

$u_1, u_2, \dots, u_r$ : left singular vectors

# Singular Values and Sigular Vectors






How to compute $v_1, v_2 — v_r$, $u_1 — u_r$, $\sigma_1 — \sigma_r$?

$$A v_i = \sigma_i u_i \qquad i = 1 \cdots r$$

$$A \underbrace{\begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix}}_{V} = \underbrace{\begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}}_{U} \underbrace{\overbrace{\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_r \end{bmatrix}}^{\mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)}}_{\Sigma}$$

$$\underset{m \times n}{A} \; \underset{n \times r}{V} = \underset{m \times r}{U} \; \underset{r \times r}{\Sigma}$$

$$V^T V = I$$

$$U^T U = I$$

$$\begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

# Singular Values and Sigular Vectors



$$A\begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}$$

JV

30

$$\underset{m\times n}{A}\ \underbrace{\begin{bmatrix} v_1 & v_2 & \cdots & v_r & v_{r+1} & v_{r+2} & \cdots & v_n \end{bmatrix}}_{V\in\mathbb{R}^{n\times n}} = \underset{m\times r}{\begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}}\begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_r & 0 & 0 & 0 \end{bmatrix}$$

$$\underbrace{\qquad\qquad}_{r}\ \underbrace{\qquad}_{n-r}$$

$v_{r+1},\dots,v_n \in N(A)$

$V^TV = VV^T = I$ $\quad$ $V$: orthogonal

# Singular Values and Sigular Vectors

$$A = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & 0 & 0 & 0 \\ & \sigma_2 & & & 0 & 0 & 0 \\ & & \ddots & & 0 & 0 & 0 \\ & & & \sigma_r & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$\underset{m \times n}{A} = \underbrace{\begin{bmatrix} u_1 & u_2 & \cdots & u_r & u_{r+1} & \cdots & u_m \end{bmatrix}}_{U \,:\, m \times m} \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{\Sigma \,:\, m \times n} \underbrace{\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \\ v_{r+1}^T \\ v_{r+2}^T \\ \vdots \\ v_n^T \end{bmatrix}}_{V^T \,:\, n \times n}$$

# Singular Values and Sigular Vectors

K. N. Toosi
University of Technology

$$\underbrace{\underset{m\times n}{A} = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T}}_{}$$

Singular Value decomposition
(SVD)

$U, V$ orthogonal

$U^T U = U U^T = I$

$V^T V = V V^T = I$

$\Sigma$ diagonal

# Singular Values and Sigular Vectors

K. N. Toosi
University of Technology

$$m\left[\;A\;\right]_{n} = \underset{m\times m}{\left[\;U\;\right]}\;\underset{m\times n}{\begin{bmatrix}\sigma_1 & & & 0 & 0\\ & \sigma_2 & & & \\ & & \ddots & & \\ 0 & & & \sigma_m & \end{bmatrix}}\;\underset{n\times n}{\left[\;V^T\;\right]}$$

$$m \leqslant n \qquad \sigma_1 \geqslant \sigma_2 \geqslant \sigma_3 \geqslant \cdots \geqslant \sigma_r > 0$$

$$\mathrm{rank}(A) = r \qquad \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_m = 0$$

# Singular Values and Sigular Vectors

K. N. Toosi
University of Technology

$$m\overset{n}{\begin{bmatrix} A \end{bmatrix}}_{m \geqslant n} = \underset{m \times m}{\begin{bmatrix} U \end{bmatrix}} \underset{m \times n}{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 \ddots & 0 \\ 0 & 0 & \sigma_n \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}} \underset{n \times n}{\begin{bmatrix} V^T \end{bmatrix}}$$

$$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$$

$$\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$$

$$\underset{m \times n}{A} = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_k u_k v_k^T \qquad k = \min(m,n)$$

$$= \sum_{i=1}^{\min(m,n)} \sigma_i u_i v_i^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

Check: $A v_j = \sigma_j u_j \qquad u_j^T A = \sigma_j v_j^T$

# Pseudo-inverse



LA28

$f(x) = Ax$

$$A = U\Sigma V^T = U\begin{bmatrix} \sigma_1 & & 0 & 0\,0 \\ 0 & \sigma_2 \ddots & & 0\,0 \\ & & \sigma_\ell & 0\,0 \end{bmatrix} V^T$$

$m \times n$

$$\Sigma^+ = \begin{bmatrix} 1/\sigma_1 & & 0 \\ 0 & 1/\sigma_2 \ddots & \\ & & 1/\sigma_\ell \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \text{if not zero}$$

$n \times m$

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

$$A^+A = V\Sigma^+\underbrace{U^TU}\Sigma V^T = V\Sigma^+\Sigma V^T$$

$$= V\begin{bmatrix} \sigma_1^{-1} & & 0 \\ 0 & \sigma_2^{-1} & \\ & & 0 \end{bmatrix}\begin{bmatrix} \sigma_1 & & 0 & 0\,0 \\ 0 & \sigma_2 & & 0\,0 \\ & & 0 & 0\,0 \end{bmatrix} V^T$$

# Pseudo-inverse

$$A^{+} = V \Sigma^{+} U^{T} \quad n \times m$$

$$A^{+}A = V \Sigma^{+} \underbrace{U^{T} U} \Sigma V^{T} = V \Sigma^{+} \Sigma V^{T}$$

$$= V \begin{bmatrix} \sigma_1^{-1} & \emptyset & \\ \emptyset & \sigma_2^{-1} & \\ & & 0 \\ & & \end{bmatrix} \begin{bmatrix} \sigma_1 & \emptyset & 0 & 0 \\ \emptyset & \sigma_2 & 0 & 0 \\ & 0 & 0 & 0 \end{bmatrix} V^{T}$$

$$= V \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix} V^{T} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix} \begin{bmatrix} v_1^{T} \\ v_2^{T} \\ v_3^{T} \end{bmatrix}$$

$$= \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} \begin{bmatrix} v_1^{T} \\ v_2^{T} \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} v_1^{T} \\ v_2^{T} \end{bmatrix}$$

# Pseudo-inverse

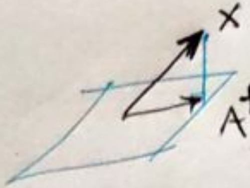
$$= V\begin{bmatrix}1&&\\&1&\\&&0\end{bmatrix}V^T = \begin{bmatrix}v_1 & v_2 & v_3\end{bmatrix}\begin{bmatrix}1&&\\&1&\\&&0\end{bmatrix}\begin{bmatrix}v_1^T\\v_2^T\\v_3^T\end{bmatrix}$$

$$= \begin{bmatrix}v_1 & v_2\end{bmatrix}\begin{bmatrix}1&\\&1\end{bmatrix}\begin{bmatrix}v_1^T\\v_2^T\end{bmatrix} = \begin{bmatrix}v_1 & v_2\end{bmatrix}\begin{bmatrix}v_1^T\\v_2^T\end{bmatrix}$$

$$x \in R(A) \Rightarrow \alpha v_1 + \beta v_2$$

$$x \in R(A) \Rightarrow A^+Ax = \begin{bmatrix}v_1 & v_2\end{bmatrix}\begin{bmatrix}v_1^T\\v_2^T\end{bmatrix}(\alpha v_1 + \beta v_2) = \begin{bmatrix}v_1 & v_2\end{bmatrix}\begin{bmatrix}\alpha\\\beta\end{bmatrix}$$

$$= \alpha v_1 + \beta v_2 = x$$




$A^+Ax$

$A^T(AA^T)^{-1}A$

A full ~~column~~ row rank $\Rightarrow A^+ = A^T(AA^T)^{-1}$

# Pseudo-inverse

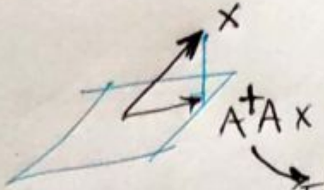
$$= V \begin{bmatrix}1&&\\&1&\\&&0\end{bmatrix} V^T = [v_1\ v_2\ v_3]\begin{bmatrix}1&&\\&1&\\&&0\end{bmatrix}\begin{bmatrix}v_1^T\\v_2^T\\v_3^T\end{bmatrix}$$

$$= [v_1\ v_2]\begin{bmatrix}1&\\&1\end{bmatrix}\begin{bmatrix}v_1^T\\v_2^T\end{bmatrix} = [v_1\ v_2]\begin{bmatrix}v_1^T\\v_2^T\end{bmatrix}$$

$$x \in R(A) \Rightarrow \alpha v_1 + \beta v_2$$

$$x \in R(A) \Rightarrow A^+Ax = [v_1\ v_2]\begin{bmatrix}v_1^T\\v_2^T\end{bmatrix}(\alpha v_1 + \beta v_2) = [v_1\ v_2]\begin{bmatrix}\alpha\\\beta\end{bmatrix}$$

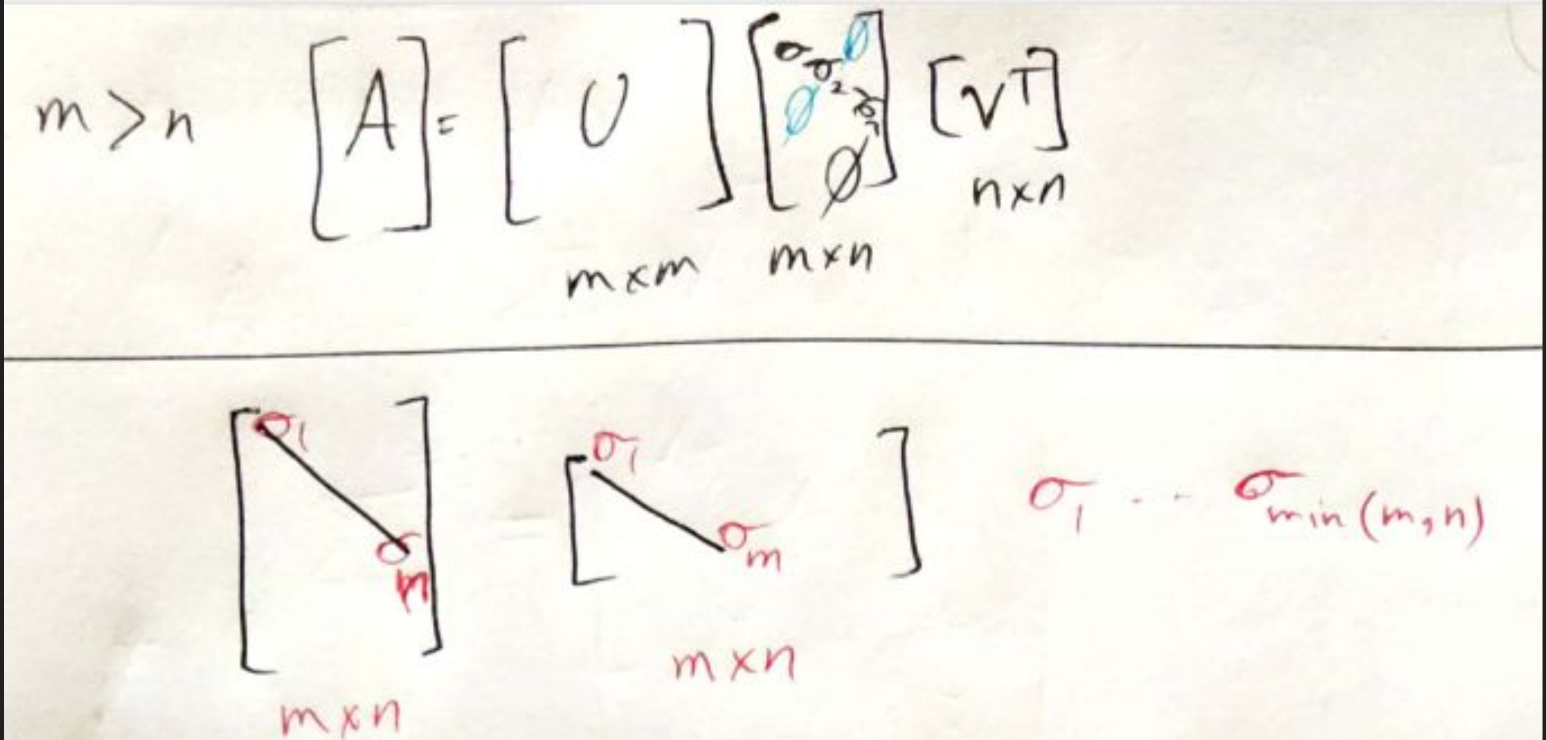
$$= \alpha v_1 + \beta v_2 = x$$




$A^T(AA^T)^{-1}A$

A full ~~column~~ row rank $\Rightarrow A^+ = A^T(AA^T)^{-1}$

# SVD for tall and fat matrices



# How to find singular values and singular vectors



The bad way!

$$\begin{bmatrix}\sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3\end{bmatrix} = \begin{bmatrix}\sigma_1 & & & \\ & & & \\ & & \sigma_3 & \emptyset\end{bmatrix}\begin{bmatrix}\sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \\ & \emptyset & \end{bmatrix}$$

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

$$\underbrace{A^TA}_{\text{symmetric}} = V \Sigma^T \underbrace{U^T U}_{I} \Sigma V^T = V \Sigma^T \Sigma V^T = \underset{n\times n}{V}\,\underset{n\times n}{\Lambda}\,\underset{n\times n}{V^T}$$

$$= V \Lambda V^{-1}$$

$$\lambda_i = \sigma_i^2$$

$v_1, v_2, \dots v_n,\ \sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ = eigenvector & eigenvalues of $A^TA$

$u_1, —, u_m \leftarrow \text{eig}(AA^T)$

# Obtain SVD from Eigendecomposition

K. N. Toosi
University of Technology

$$A = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T}$$

$$AA^T = U\Sigma \underbrace{V^T V}\Sigma^T U^T$$

$$= U\Sigma\Sigma^T U^T$$

$\boxed{m=3}$

$$= \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix}\sigma_1 & 0 & 0 & \boxed{\emptyset}\\ 0 & \sigma_2 & 0 & \\ 0 & 0 & \sigma_3 & \end{bmatrix} \begin{bmatrix}\sigma_1 & 0 & 0\\ 0 & \sigma_2 & 0\\ 0 & 0 & \sigma_3\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \begin{bmatrix}u_1^T\\ u_2^T\\ u_3^T\end{bmatrix}$$

$$AA^T = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix}\sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \sigma_3^2\end{bmatrix} \longleftrightarrow \begin{bmatrix}u_1^T\\ u_2^T\\ u_3^T\end{bmatrix}$$

$$\underbrace{AA^T}_{\text{symmetric \& positive semi def}} = \underbrace{U \begin{bmatrix}\sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \sigma_3^2\end{bmatrix} U^T}_{\text{eigen decomposition}}$$

$$U^T \to U^{-1}$$

# Obtain SVD from Eigendecomposition

$$\underbrace{AA^T}_{\text{symmetric \& positive semi def}} = \underbrace{U \begin{bmatrix}\sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \sigma_3^2\end{bmatrix} U^T}_{\text{eigen decomposition for symmetric matrices}}$$

$$U^T \to U^{-1}$$

$$\Rightarrow (AA^T)u_i = \sigma_i^2 u_i$$

$u_1, \dots, u_m$ eigenvectors of $AA^T$

$\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2$ eigenvalues of $AA^T$

$AA^T =$

$v_1, v_2, \dots, v_n$ eigenvectors of $A^TA$

# SVD, Rank, Column Space, Row Space, Null Sapce

$$A = \begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n \end{bmatrix}$$

VI

Compute svd (e.g. using np.linalg.svd)

$r = \text{rank}(A)$: no. of non-zero $\sigma_i = \max_i$ s.t. $\sigma_i > 0$

column space : $\text{span}\underbrace{(u_1, u_2, \ldots, u_r)}_{\text{orthonormal basis for } C(A)}$

row space : $\text{span}\underbrace{(v_1, v_2, \ldots, v_r)}_{\text{orthonormal basis for } R(A)}$

null space : $\text{span}\underbrace{(v_{r+1}, \ldots, v_n)}_{\text{orthonormal basis for } N(A)}$

left null space : $\text{span}\underbrace{(u_{r+1}, \ldots, u_m)}_{\text{orthonormal basis for } N(A^T)}$

# Problems with full SVD

K. N. Toosi
University of Technology

Assume that we have a lot of data $x_1, x_2, \dots, x_n \in \mathbb{R}^d$ LA28 VI

$n = 1000000$

$d = 100$

arrange them in a matrix

$$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix} \in \mathbb{R}^{n \times d}$$

np.linalg.svd(A)

$$X = \underbrace{U}_{n \times n} \underbrace{\Sigma}_{n \times d} \underbrace{V^T}_{d \times d} \longrightarrow \text{Full SVD}$$

$1000000 \times 1000000$

$100 \times 100$

$n \times d$ → represented as $\vec{\sigma} = \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_d \end{bmatrix}$ → 100

$\sigma_d$ → $\min(n,d) = d$

float32 $\Rightarrow 4 \times 10^{12}$ bytes

4000 GB

# Skinny (thin) SVD

4000 GB

$$X = U \Sigma V^T = \begin{bmatrix} \underbrace{u_1\, u_2 \cdots u_d} & u_{d+1} \cdots u_n \end{bmatrix} \begin{bmatrix} \ddots \\ 0\,0\,0 \\ 0\,0\,0 \\ \vdots \end{bmatrix} [V^T]$$

$m \times n$

$m \gg n$

$$= \begin{bmatrix} u_1\, u_2 \cdots u_d \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_d \end{bmatrix} V^T$$

$4 \times 1000000 \times 100 = 400\ MB$

$l = \min(m, n)$

$$A = U \Sigma V^T = U_{1:l} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_l \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_l^T \end{bmatrix}$$

skinny SVD, thin SVD

# Thin (Skinny) SVD

$$\underset{m \times n}{\begin{array}{c|c} & n \\ \hline m & \begin{bmatrix} A \end{bmatrix} \end{array}} = \underbrace{\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}}_{m \times n} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}}_{n \times n} \underbrace{\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}}_{n \times n}$$

$$\begin{array}{c|c} & n \\ \hline m & \begin{bmatrix} A \end{bmatrix} \end{array} = \underbrace{\begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}}_{m \times m} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \end{bmatrix}}_{m \times m} \underbrace{\begin{bmatrix} v_1^T \\ \vdots \\ v_m \end{bmatrix}}_{m \times n}$$

thin svd

skinny svd

# Compact SVD

$$\begin{bmatrix} \\ A \\ \\ \end{bmatrix} = \underbrace{\begin{bmatrix} u_1 & u_2 & \cdots & u_r \\ & & & \end{bmatrix}}_{m\times r} \underbrace{\underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}}_{r\times r}}_{\text{invertible}} \underbrace{\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{bmatrix}}_{r\times n}$$

compact svd

$\text{rank}(A) = r$

$A:\ m\times n$

$$\Sigma^{-1} = \begin{bmatrix} 1/\sigma_1 & & \\ & 1/\sigma_2 & \\ & & \ddots \\ & & & 1/\sigma_r \end{bmatrix}$$