Mathematics for AI

Lecture 12

Properties of Positive Definite Matrices, Cholesky Decomposition Introduction to Singular Value Decomposition

Reminder: Positive definite and positive semi-definite matrices

MA12 (I)

Positive definite

$\boxed{A^T = A}$

$A \in \mathbb{R}^{n\times n}$ is positive definite if $\forall x \in \mathbb{R}^n_{\neq \emptyset}$ $x^T A x > 0$

~~sem~~ positive semi-definite $\forall x \in \mathbb{R}^n$ $x^T A x \geq 0$

$A \in \mathbb{C}^{n\times n}$ is positive-definited $A^H = \bar{A}^T = A$

$\forall x \in \cancel{\mathbb{R}}\,\mathbb{C}^n \setminus \{0\}$ $\bar{x}^T A x > 0$

$A \succ 0$: A is positive-definit

$A \succeq 0$ A is positive semi-definite

negative (semi) definite $\prec, \preceq$

Positive definite

Here, by positive-definite we mean symmetric positive definite

Positive definiteness and singularity

K. N. Toosi
University of Technology

A is positive-definite $\Rightarrow$ A is non-singular

Proof:

Assume A is singular $\Rightarrow \exists\, v \in \mathbb{R}^n,\ v \neq 0 \quad Av = \vec{0}$

$\Rightarrow v^T A v = 0$ تناقض

Positive definiteness and eigenvalues

K. N. Toosi
University of Technology

A is positive-definite ⇒ all eigenvalues are (real and) positive.

~~A~~ A is positive-~~d~~semi-definite ⇒ all eigenvalues are (real and) nonnegative

$U*U^T$

$$\underset{n\times n}{A} = \underset{n\times p}{U} U^T \text{ for } U \in \mathbb{R}^{n\times p} \Rightarrow x^T A x = x^T U U^T x = (\underbrace{U^T x}_{y})^T \underbrace{U^T x}_{y}$$

$$= y^T y = \|y\|^2 \geq 0$$

$\Rightarrow UU^T$ is always positive semi definite

$$U = [u_1\ u_2 \cdots u_n] \qquad UU^T = [u_1\ u_2 \cdots u_n]\begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix}$$

$$= u_1 u_1^T + u_2 u_2^T + \cdots + u_n u_n^T$$

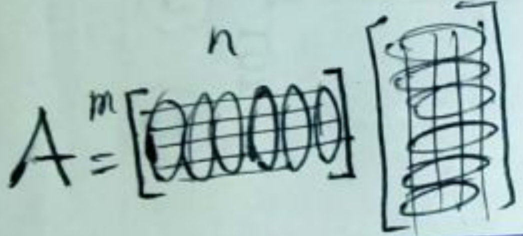
$$(UU^T)_{ij} = \sum_k u_k[i]\, u_k[j]$$

$U*U^T$

$A = UU^T$ $\quad U$ full-row-rank

$\boxed{A = U^T U \quad U \text{ full-column-rank}} \Rightarrow$ independent columns

$x \neq 0 \qquad x^T U^T U x = (Ux)^T(\underbrace{Ux}_{y \neq 0})$ $\qquad x \neq 0 \Rightarrow Ux \neq 0$



$= y^T y > 0$

$y \neq 0 \Rightarrow A$ positive-definite

Correlation matrix

K. N. Toosi
University of Technology

$$\text{data} = D = \begin{bmatrix} d_1^T \\ d_2^T \\ d_3^T \\ \vdots \\ d_M^T \end{bmatrix}$$

Correlation Matrix

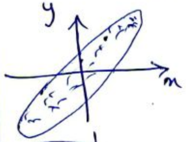
$$C = \frac{1}{M}\sum_{i=1}^{M} \underbrace{d_i\, d_i^T}_{n\times n} = \frac{1}{M} D^T D$$

$$D^T D = \begin{bmatrix} d_1 & d_2 & \cdots & d_M \end{bmatrix} \begin{bmatrix} d_1^T \\ d_2^T \\ \vdots \\ d_M^T \end{bmatrix}$$

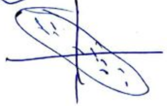
$$d_i^T = \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}$$

$$\mathrm{Corr}(x,y) = \frac{1}{M}\sum_{i=1}^{M} x_i\, y_i$$

$$\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_M & y_M & z_M \end{bmatrix}$$



$\mathrm{Corr}(x,y) > 0$



$\mathrm{Corr}(x,y) < 0$

$\mathrm{Corr}(x,y) = 0$

Correlation matrix

$\Leftrightarrow \|Ax\| > 0$ for all $x \neq 0$

$\cancel{Ax \neq 0} \Leftrightarrow$ for all $x \neq 0$ we have $Ax \neq 0$

$(A \in \mathbb{R}^{m \times n}) \Leftrightarrow \operatorname{rank}(A) = n \Leftrightarrow$ A has full column rank

Covariance matrix

K. N. Toosi
University of Technology

Decomposition

K. N. Toosi
University of Technology

Every positive $^{\text{semi}}$-definite matrix $A \in \mathbb{R}^{n\times n}$ can be factorized

as $A = \underline{\underline{U}}^T U$. for some $U \in \mathbb{R}^{n \times n}$

Square root

Every positive semi-definite matrix $A \in \mathbb{R}^{n\times n}$ can be factorized

as $A = U^T U$. for some $U \in \mathbb{R}^{n\times n}$

For a (symmetric) positive semi-definite matrix A there is a unique positive semi-definite matrix P such that $A = P\,P$ $(= P^H P)$. P is called the square root of A and is denoted by $A^{-\frac{1}{2}}$.

Cholskey Decomposition

K. N. Toosi
University of Technology

Cholskey Decomposition

Every positive semi definite matrix $A \in \mathbb{R}^{n \times n}$ can be decomposed as $A = L L^T$ where L is ~~lower~~ lower-triangular.

$A \in \mathbb{C}^{n \times n}$ $\quad A = L L^H = L L^*$.

Singular Value Decomposition (SVD)

K. N. Toosi
University of Technology

Singular Value Decomposition

M12 (III)

تجزیه مقادیر منفرد

Every matrix $A \in \mathbb{R}^{m\times n}$ can be decomposed as

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

U, V orthogonal & Σ diagonal

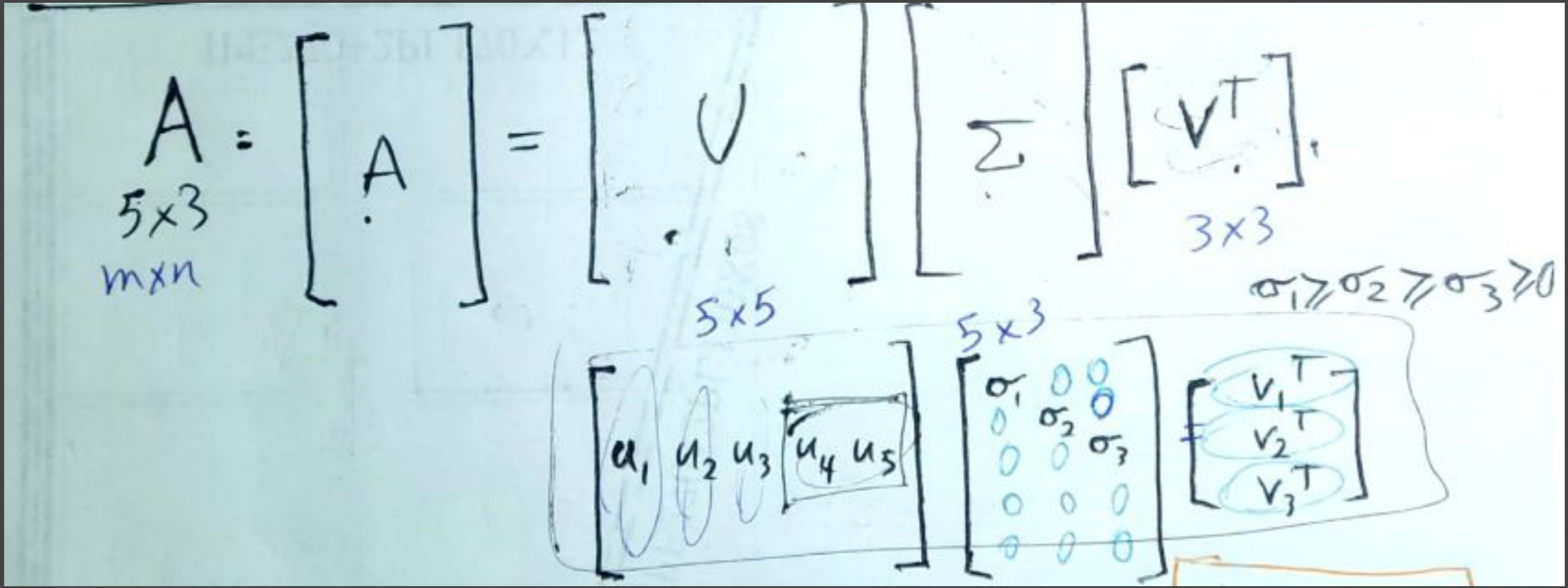
$U^TU = UU^T = I$

$V^TV = VV^T = I$

$$\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \end{bmatrix} \qquad \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \\ & & \\ & & \end{bmatrix}$$

$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n \geqslant 0$

Singular Value Decomposition (SVD)



Singular Value Decomposition (SVD)

K. N. Toosi
University of Technology

$$\underset{\substack{5\times 3\\ m\times n}}{A} = \begin{bmatrix} A \end{bmatrix} = \underset{5\times 5}{\begin{bmatrix} U \end{bmatrix}} \underset{5\times 3}{\begin{bmatrix} \Sigma \end{bmatrix}} \underset{3\times 3}{\begin{bmatrix} V^T \end{bmatrix}}$$

$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0$

$$\begin{bmatrix} u_1 & u_2 & u_3 & \boxed{u_4 \; u_5} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ v_3^T \end{bmatrix}$$

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

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

$\sigma_1, \sigma_2, \dots, \sigma_n$: singular values.

$$\boxed{\begin{array}{l} \|u_i\| = 1 \\ \|v_i\| = 1 \\ \\ u_i \perp u_j \quad i \neq j \\ v_i \perp v_j \quad i \neq j \end{array}}$$

$$\underset{5\times 3}{A} = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 & u_5 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ v_3^T \end{bmatrix}$$

$$= \underset{5\times 3}{\begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}} \underset{3\times 3}{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}} \underset{3\times 3}{\begin{bmatrix} v_1^T \\ v_2^T \\ v_3^T \end{bmatrix}} = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \sigma_3 u_3 v_3^T$$

Skinny SVD

full_matrices = False $\longrightarrow$ skinny

m 12 (I

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

full svd

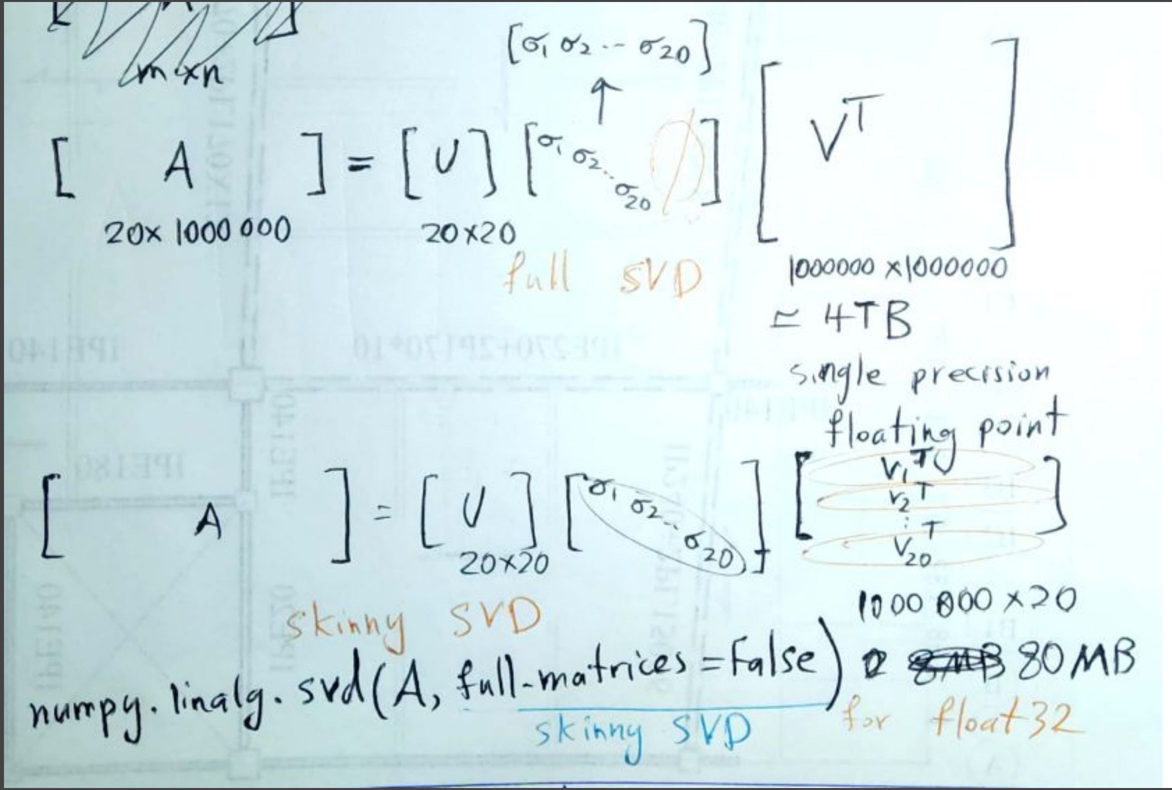
.m

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_m & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$\underset{m\times m}{\begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix}} \underset{m\times m}{\begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \; \sigma_m \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_m^T \end{bmatrix}}$$

skinny SVD

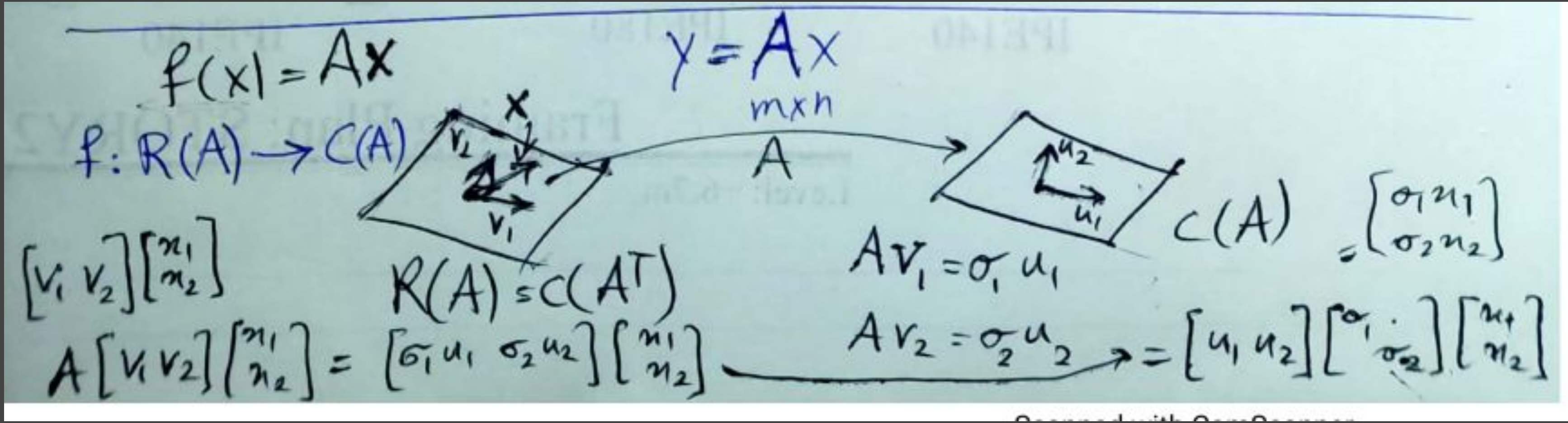
Skinny SVD & Memory usage



K. N. Toosi
University of Technology

Geometric Interpretation

K. N. Toosi
University of Technology



SVD and matrix rank

M12 IV

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

$u_1\, u_2 \cdots u_r \quad u_{r+1} - u_m$

$v_1 - v_r$

$v_{r+1}, \ldots, v_n$

$\mathrm{rank}(A) = \#$ non-zero singular values

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

ran

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

$\downarrow$

$\min(m,n)$

SVD, row space, column space, null space

M12 IV

$$\underset{m\times n}{\begin{bmatrix} A \end{bmatrix}} = \begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} \begin{matrix} \} \; v_1 - v_r \\ \} \; v_{r+1}, \dots, v_n \end{matrix}$$

$\underbrace{u_1\, u_2 \cdots u_r}$ $\underbrace{u_{r+1} - u_m}$

$\text{rank}(A) = \#$ non-zero singular values

$$\sigma_1 > \sigma_2 > \cdots > \sigma_r > 0 \qquad \text{ran}$$

$$\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_{\underset{\min(m,n)}{\downarrow}{n}} = 0$$

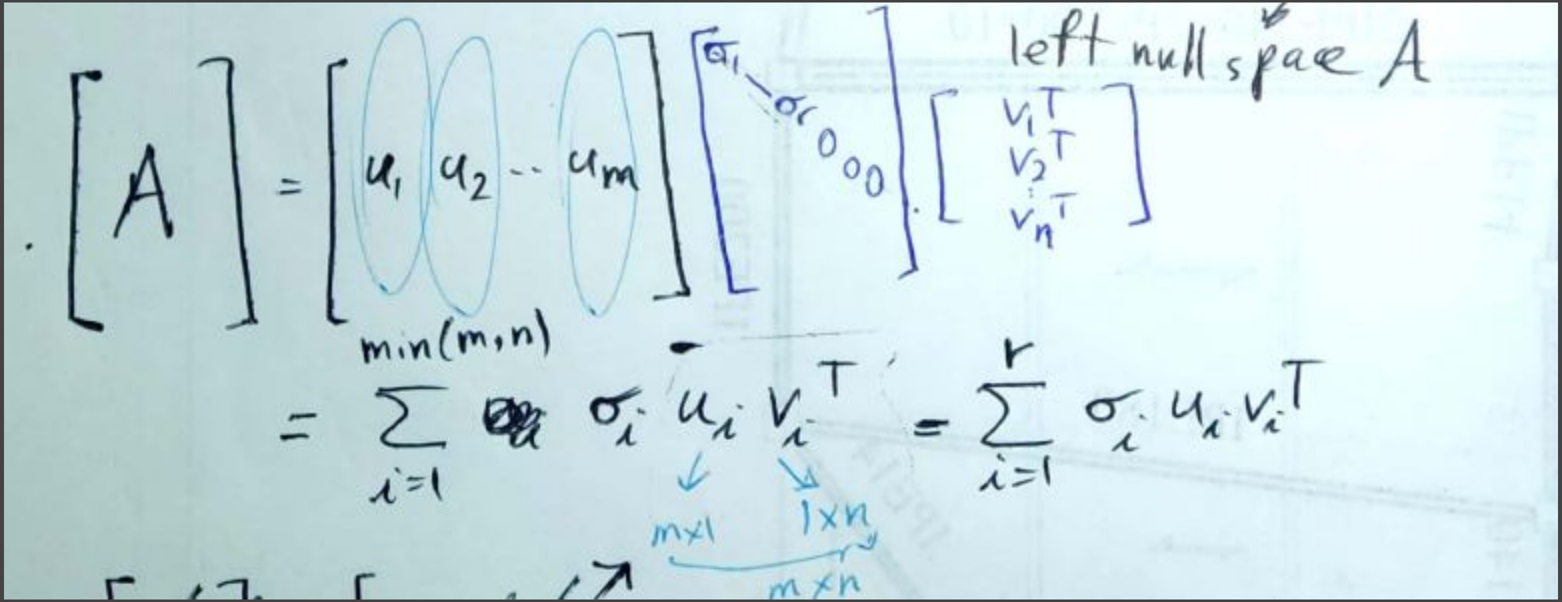
$u_1, u_2, \dots u_r$: form a $\underset{\text{orthonormal}}{\wedge}$ basis for $C(A)$

$v_1, v_2, \dots, v_r$ form an orthonormal basis for $C(A^T)$ row-spaces of A

$v_{r+1}, v_{r+2}, \dots, v_n$ form an " basis for $N(A)$

$u_{r+1}, u_{r+2}, \dots, u_m$ form " " basis for $N(A^T)$

Unrolling SVD



Compact SVD

K. N. Toosi
University of Technology

left null space A

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

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

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

Compact SVD