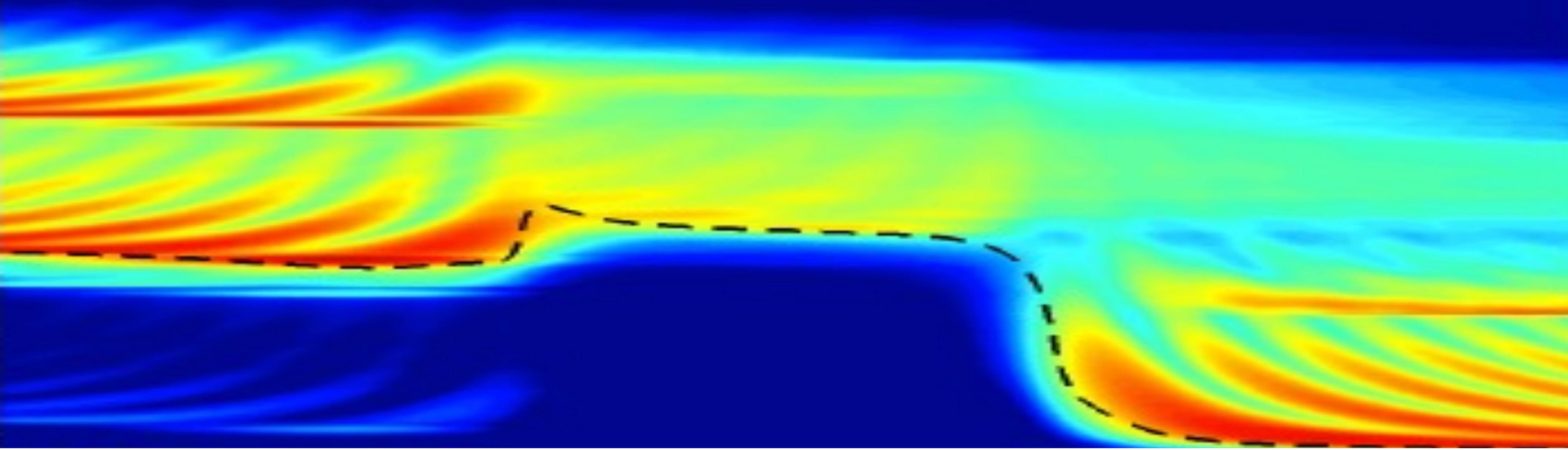


EECS 16B
Designing Information Devices and Systems II
Lecture 22

Prof. Yi Ma
Department of Electrical Engineering and Computer Sciences, UC Berkeley,
yima@eecs.berkeley.edu

Outline

- Singular Value Decomposition (SVD)
  - Theorem (with proof)
  - Examples of SVD
  - Full SVD
  - Geometric Interpretation of SVD

Singular Value Decomposition (SVD)

Given $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rank}(A) = r$, we like to decompose it into a special matrix form:

$U_r = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r]$ orthogonal

$V_r = [\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r]$ orthogonal

$\Sigma_r = \operatorname{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_r\} > 0$

$U_r \quad U_n = U$

$$A = U_r \Sigma_r V_r^\top = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_r \end{bmatrix} \begin{bmatrix} \vec{v}_1^\top \\ \vec{v}_2^\top \\ \vdots \\ \vec{v}_r^\top \end{bmatrix}$$

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

$\Sigma_r \quad V_r^T$

$\mathbb{R}^{n\times n}$

$$\begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix} A^T A [\underbrace{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r}_{V_r \text{ - orthogonal}}, \vec{v}_{r+1}, \ldots, \vec{v}_n] = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_r & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}$$

$(n-r)\times(n-r)$

$\lambda_1 \geq \lambda_2, \cdots \geq \lambda_r > 0$

$\sigma_i = \sqrt{\lambda_i}, \ i = 1, \ldots, r.$

Singular Value Decomposition (Theorem)

Theorem: given $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rank}(A) = r$, let $A^\top A = \sum_{i=1}^{r} \lambda_i \vec{v}_i \vec{v}_i^\top$ and $\sigma_i = \sqrt{\lambda_i}$, $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i \in \mathbb{R}^m$, $i = 1, \ldots, r$. Then we have ① $U_r = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r]$ orthogonal, and

② $A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^\top = U_r \Sigma_r V_r^\top$ $\quad \Sigma_r = \operatorname{diag}\{\sigma_1, \ldots, \sigma_r\} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_r \end{bmatrix}$

Proof:

① $\vec{u}_i$ orthonormal?

$$\vec{u}_i^T \vec{u}_j = \left(\frac{1}{\sigma_i} A \vec{v}_i\right)^T \left(\frac{1}{\sigma_j} A \vec{v}_j\right)$$

$$= \frac{1}{\sigma_i \sigma_j} \vec{v}_i^T \underbrace{A^T A \vec{v}_j}_{\lambda_j \vec{v}_j} = \frac{\lambda_j}{\sigma_i \sigma_j} \vec{v}_i^T \vec{v}_j$$

$$= \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \qquad \frac{\lambda_i}{\sigma_i \sigma_i} \qquad \blacksquare$$

② $A \overset{?}{=} \sum_{i=1}^{r} \sigma_i \vec{u}_i v_i^T$

$$\text{right} = \sum_{i=1}^{r} \sigma_i \left(\frac{1}{\sigma_i} A \vec{v}_i\right) \vec{v}_i$$

$$= \sum_{i=1}^{r} A \vec{v}_i v_i^T$$

$$= A\left(\sum_{i=1}^{r} v_i v_i^T\right)$$

Singular Value Decomposition (Theorem)

Theorem: given $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rank}(A) = r$, let $A^\top A = \sum_{i=1}^{r} \lambda_i \vec{v}_i \vec{v}_i^\top$ and $\sigma_i = \sqrt{\lambda_i}$, $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i \in \mathbb{R}^m$, $i = 1, \ldots, r$. Then we have $U = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r]$ orthogonal, and

$$A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^\top = U_r \Sigma_r V_r^\top$$

Proof:

$V^T A^T A V$ $\quad V_{n\times n}$ - orthogonal $\quad V = [\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r, \vec{v}_{r+1}, \ldots, \vec{v}_n]$

$$V^T V = \underline{V V^T} = I \qquad [\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n] \begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix} = I$$

$$= \sum_{i=1}^{n} \vec{v}_i \cdot \vec{v}_i^T = \sum_{i=1}^{r} \vec{v}_i \cdot \vec{v}_i^T + \sum_{i=r+1}^{n} \vec{v}_i \cdot \vec{v}_i^T$$

$$A = A \cdot I = A(VV^T) = A\left(\sum_{i=1}^{r} \vec{v}_i \cdot \vec{v}_i^T + \sum_{i=r+1}^{n} \vec{v}_i \vec{v}_i^T\right)$$

$$= A\left(\sum_{i=1}^{r} \vec{v}_i \vec{v}_i^T\right) + \underline{A\left(\sum_{i=r+1}^{n} \vec{v}_i \vec{v}_i^T\right)}$$

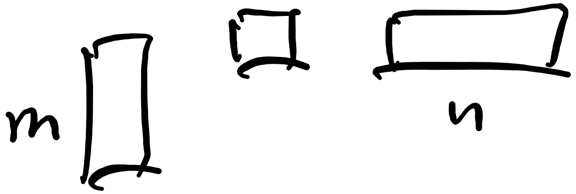
$$A \sum_{i=r+1}^{n} \vec{v}_i \vec{v}_i^T \qquad \underline{A^T A v_i} = 0 \quad i = r+1, \dots, n.$$

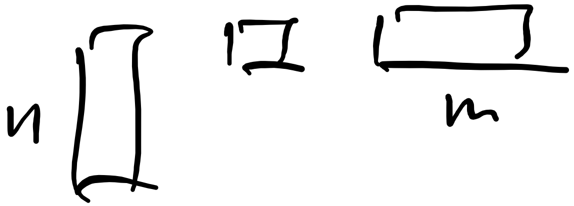
$$= \sum_{i=r+1}^{n} \underline{(A \vec{v}_i)} \vec{v}_i^T \qquad Nu(A^T A) = Nu(A) \quad (v_i^T A^T A v_i)$$

$$\Rightarrow A v_i = 0 \quad i = r+1, \dots, n.$$

$$= 0. \qquad \square$$

$$\rightarrow A = U_r \Sigma_r V_r^T \qquad \rightarrow A^T = V_r \Sigma_r U_r^T$$





Singular Value Decomposition

$A = \left[\ \right]$

Given $A \in \mathbb{R}^{m\times n}$ with $\mathrm{rank}(A) = r$, two (equivalent) ways to find SVD:

$A^\top A \in \mathbb{R}^{n\times n}$

① Find orthogonal $V \in \mathbb{R}^{n\times n}$

$$V^T A^T A V = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}_{n\times n}$$

$\lambda_1 \geq \lambda_2 \cdots \geq \lambda_r > 0$

② $\sigma_i = \sqrt{\lambda_i}$, $i = 1, \ldots, r$.

$\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i$, $i = 1, \ldots, r$.

$V_r = [\vec{v}_1, \ldots, \vec{v}_r]$, $U_r = [\vec{u}_1, \ldots, \vec{u}_r]$

$$A = U_r \Sigma_r V_r^T$$

$AA^\top \in \mathbb{R}^{m\times m}$

① Find orthogonal $U \in \mathbb{R}^{m\times m}$

$$U^T A A^T U = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & 0 & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

$\lambda_1 \geq \lambda_2 \cdots \lambda_r > 0$

② $\sigma_i = \sqrt{\lambda_i}$

$\vec{v}_i = \frac{1}{\sigma_i} A^T \vec{u}_i$

$V_r = [\vec{v}_1, \ldots, \vec{v}_r]$ $U_r = [\vec{u}_1, \ldots, \vec{u}_r]$

$$A = U_r \Sigma_r V_r^T$$

Singular Value Decomposition (example)

$$A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}, \quad A^\top = \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix}$$

$$A^T A \leftarrow V$$

$$A^T A = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix} \qquad AA^T = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$$

$\lambda_1 = 32, \ \lambda_2 = 18$

$\sigma_1 = 4\sqrt{2}, \ \sigma_2 = 3\sqrt{2}$

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\vec{v}_1 = \frac{1}{4\sqrt{2}} A^T \vec{u}_1 = \frac{1}{4\sqrt{2}} \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\vec{v}_2 = \frac{1}{3\sqrt{2}} A^T \vec{u}_2 = \frac{1}{3\sqrt{2}} \begin{bmatrix} -3 \\ 3 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} 1 & \frac{1}{\sqrt{2}} 1 \\ \frac{1}{\sqrt{2}} -1 & \frac{1}{\sqrt{2}} 1 \end{bmatrix}$$

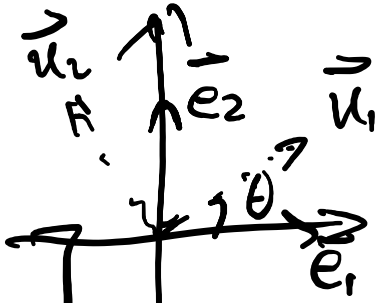
Singular Value Decomposition (example)

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad A^\top = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad -A = A^T \qquad A\vec{v} = \lambda \vec{v} \text{ (previous lecture)}$$

① $AA^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\quad \vec{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$\sigma_1 = \sigma_2 = 1$ $\quad \vec{v}_1 = A^T \vec{u}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ \vec{v}_2 = A^T \vec{u}_2 \begin{bmatrix} -1 \\ 0 \end{bmatrix}$

$$A = U \Sigma V^T = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_{V^T} = A.$$



② $\vec{u}_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \quad \vec{u}_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$

$\vec{v}_1 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} -\cos\theta \\ -\sin\theta \end{bmatrix}$

$$A = \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{U \cdot I} \underbrace{\begin{bmatrix} -\sin\theta & \cos\theta \\ -\cos\theta & -\sin\theta \end{bmatrix}}_{V^T} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Compact versus Full SVD

Compact SVD: $A = \underbrace{\sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^\top = U_r \Sigma_r V_r^\top}_{\text{outer products}}$ $\quad A = \underbrace{[\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r]}_{U_r} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_r \end{bmatrix}}_{\Sigma_r} \underbrace{\begin{bmatrix} \vec{v}_1^\top \\ \vec{v}_2^\top \\ \vdots \\ \vec{v}_r^\top \end{bmatrix}}_{V_r^T}$

$A \in \mathbb{R}^{m \times n} : \mathbb{R}^n \longrightarrow \mathbb{R}^m \quad \vec{y} = A\vec{x}$

$$\begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_r^T \\ \vec{v}_{r+1}^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix} A^T A \, [\underbrace{\vec{v}_1, \ldots, \vec{v}_r}_{V_r}, \underbrace{\vec{v}_{r+1}, \ldots, \vec{v}_n}_{V_{n-r}}] = \begin{bmatrix} \lambda_1 & & & & 0 \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{bmatrix}$$

Compact versus Full SVD

Full SVD: $$A = U\Sigma V^\top = \underbrace{[\overbrace{\vec{u}_1,\dots,\vec{u}_r}^{U_r} \mid \overbrace{\vec{u}_{r+1},\dots,\vec{u}_m}^{U_{n-r}}]}_{U_{m\times m}} \underbrace{\left[\begin{array}{ccc|ccc} \sigma_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & 0 & & \vdots \\ 0 & 0 & \sigma_r & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & & 0 & 0 & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right]}_{m\times n} \underbrace{\begin{bmatrix} \vec{v}_1^\top \\ \vdots \\ \vec{v}_r^\top \\ \hline \vec{v}_{r+1}^\top \\ \vdots \\ \vec{v}_n^\top \end{bmatrix}}_{V^\top_{n\times n}}$$

Σ_r, $O_{m\times n-r}$, $O_{m-r\times r}$, $O_{m-r\times n-r}$

$V_r^\top$, $V_{n-r}^\top$

$$= U_r \Sigma_r V_r^\top$$

Full SVD for Full-rank Matrices

① $m = n = r \qquad A = U \Sigma V^T \leftarrow$

② $m > n = r \qquad A = U_r \Sigma_r V_r^T = U \underbrace{\begin{bmatrix} \Sigma_r \\ 0 \end{bmatrix}}_{\Sigma} V^T$

③ $n > m = r \qquad A = U_r \Sigma_r V_r^T$

$$= U \begin{bmatrix} \Sigma_r & 0 \end{bmatrix} V^T$$

Geometric Interpretation of SVD

$$A = U_r \Sigma_r V_r^T = U \Sigma V^T : \mathbb{R}^n \to \mathbb{R}^m$$

$$\vec{y} = A\vec{x} = (U(\underline{\Sigma}(\underline{V}^T \vec{x})))$$

① orthogonal Q $\quad Q^TQ = QQ^T = I$

$$\vec{y} = Q\vec{x} \qquad \|Q\vec{x}\|_2^2 = (Qx)^T(Qx) = x^T \underbrace{Q^TQ}_{I} x = x^Tx = \|x\|_2^2$$

$$\langle Q\vec{z}, Q\vec{x} \rangle = \langle \vec{z}, \vec{x} \rangle$$

② Σ - pseudo diagonal $\quad \vec{y} = \Sigma \vec{x}$, $\quad y_i = \sigma_i x_i$, $i = 1, \ldots, r$.

$y_i = 0$, $i = r+1, \ldots$

Geometric Interpretation of SVD

$$\vec{y} = A_{m\times n}\vec{x} = U\Sigma V^T\vec{x} \qquad U^TU = I$$

$$\underbrace{(U^T\vec{y})}_{\vec{y}'} = \Sigma\underbrace{(V^T\vec{x})}_{\vec{x}'} \qquad U^T\vec{y}\in\mathbb{R}^m,\ V^T\vec{x}\in\mathbb{R}^n$$

$$\vec{y}' = \Sigma\vec{x}'$$

$$A = \quad U \quad \underline{\Sigma} \quad V^\top$$