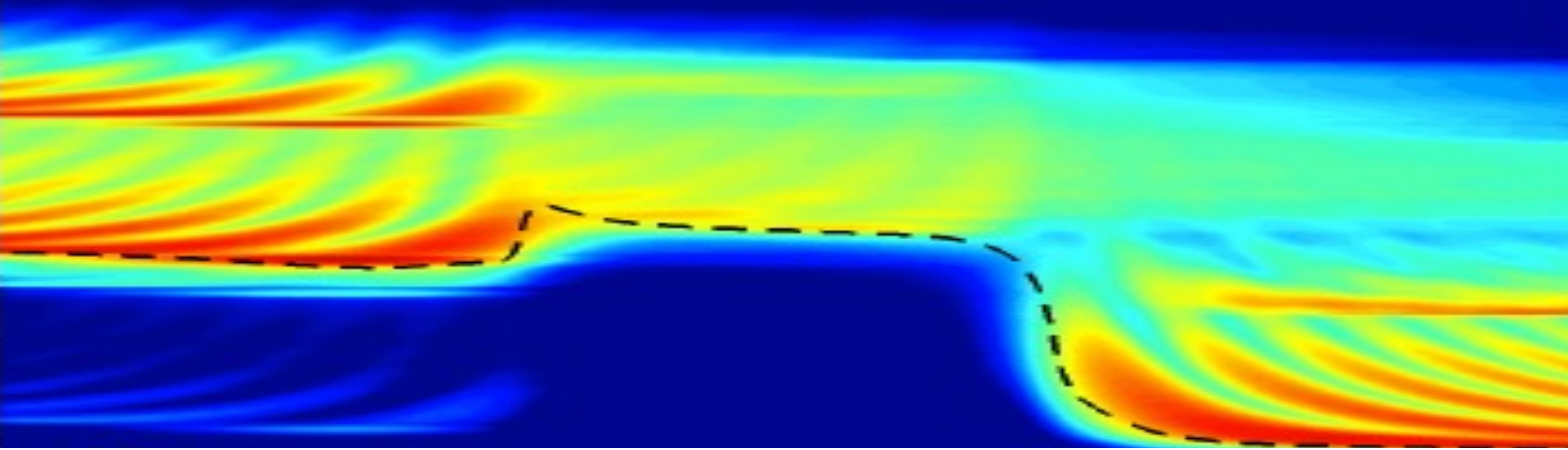


EECS 16B
Designing Information Devices and Systems II
Lecture 23

Prof. Yi Ma

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

Outline

- Singular Value Decomposition (SVD)
  - Geometric Interpretation of SVD
- Applications of SVD: unifying
  - Matrix (Pseudo) Inverse
  - Least Squares
  - Minimum Norm Solution

$$y = Ax$$

$$\text{“}A^{-1}\text{”} \qquad x = \text{“}A^{-1}\text{”}\, y$$

Singular Value Decomposition (SVD)

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

$V = [\vec{v}_1, \ldots, \vec{v}_n]$ orthonormal e.v.'s for $A^\top A$ eigenvalues of $A^\top A$ (or $AA^\top$) : $\lambda_1 \geq \cdots \geq \lambda_r > 0 \cdots 0$

$U = [\vec{u}_1, \ldots, \vec{u}_m]$ orthonormal e.v.'s for $AA^\top$ $\Sigma_r = \mathrm{diag}\{\sigma_1 = \sqrt{\lambda_1}, \ldots, \sigma_r = \sqrt{\lambda_r}\} > 0$

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

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

Full SVD: $A = U\Sigma V^\top = [\vec{u}_1, \ldots, \vec{u}_r, \vec{u}_{r+1}, \ldots, \vec{u}_n] \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & 0 & \ddots & \vdots \\ 0 & 0 & \sigma_r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \ddots & 0 & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \vec{v}_1^\top \\ \vdots \\ \vec{v}_r^\top \\ \vec{v}_{r+1}^\top \\ \vdots \\ \vec{v}_n^\top \end{bmatrix}$

V_r^T $\mathbb{R}^n$ V_{n-r}^T

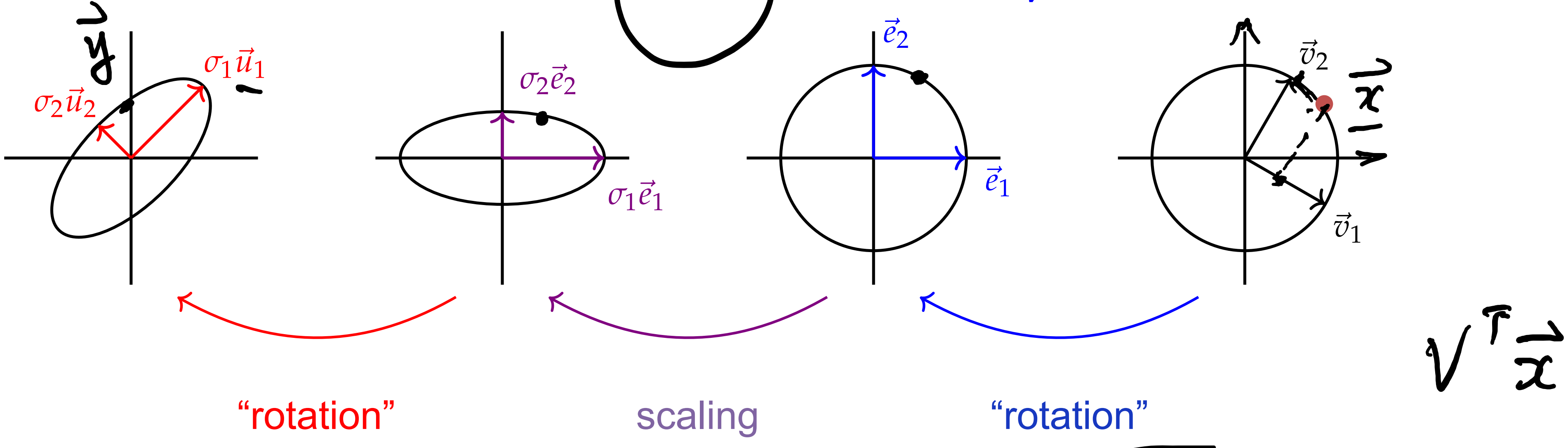
$A: \mathbb{R}^n \to \mathbb{R}^m$ U_r U_{m-r} $\mathbb{R}^m$

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

Geometric Interpretation of SVD

$$\underline{\vec{y} = A\vec{x}} : \quad A = U\Sigma V^\top = \underbrace{[U_r, U_{m-r}]}_{U} \begin{bmatrix} \Sigma_r & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{bmatrix} \underbrace{\begin{bmatrix} V_r^\top \\ V_{n-r}^\top \end{bmatrix}}_{V^\top}$$

$$A = \quad U \quad \Sigma \quad V^\top$$

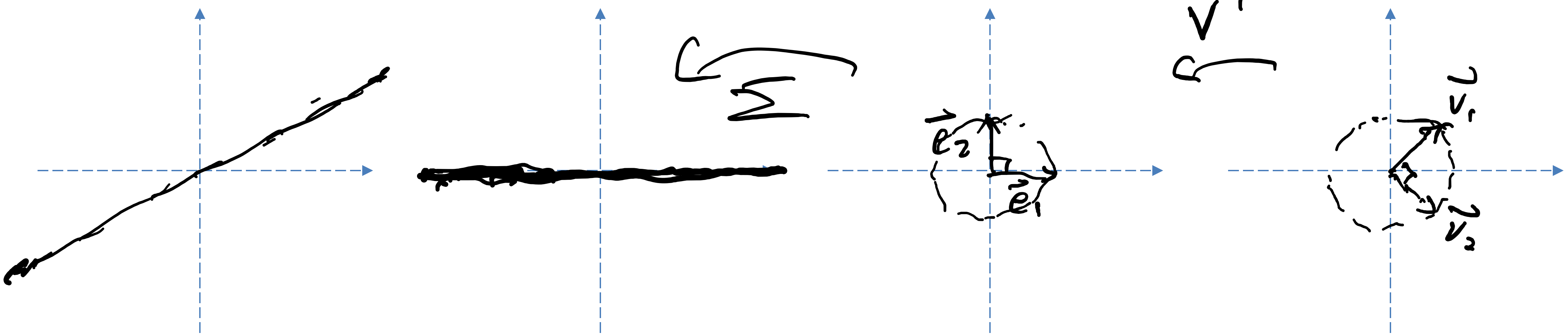


$V^\top \vec{x}$

Geometric Interpretation of SVD (Example)

$$A = \begin{bmatrix} 4 & 4 \\ 3 & 3 \end{bmatrix} = \underbrace{5\sqrt{2}}_{\sigma_1} \left(\underbrace{\frac{1}{5}\begin{bmatrix} 4 \\ 3 \end{bmatrix}} \right) \left(\frac{1}{\sqrt{2}} [1,\ 1] \right) \quad V^T$$

$$= \underbrace{\left[\frac{1}{5}\begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad \underline{\frac{1}{5}\begin{bmatrix} -3 \\ 4 \end{bmatrix}} \right]}_{U} \underbrace{\begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 0 \end{bmatrix}}_{\Sigma} \begin{bmatrix} \frac{1}{\sqrt{2}} [1,\ 1] \\ \frac{1}{\sqrt{2}} [1,\ -1] \end{bmatrix} \leftarrow$$



Algebraic Interpretation of SVD

$$A = U\Sigma V^\top = [U_r, U_{m-r}] \begin{bmatrix} \Sigma_r & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{bmatrix} \begin{bmatrix} V_r^\top \\ V_{n-r}^\top \end{bmatrix} = \underline{U_r \Sigma_r V_r^\top}$$

① $A V_{n-r} = 0 \Rightarrow \text{Null}(A) = \underline{\text{col}(V_{n-r})}$

② $U_{m-r}^T A = 0 \Rightarrow \text{Null}(A^T) = \text{col}(U_{m-r})$

③ $\text{col}(U_r) = \text{col}(A) \perp \text{Null}(A^T)$

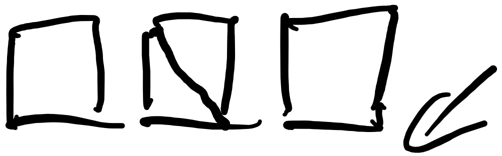
④ $\text{Col}(V_r) = \text{col}(A^T) = \text{row}(A)$

$\perp \text{col}(V_{n-r})$ $\qquad\qquad \perp \text{Null}(A)$

Applications of SVD: Matrix Inverse

Given $A \in \mathbb{R}^{m\times n}$ with $\underline{\mathrm{rank}(A) = r = m = n}$: $A = U\Sigma V^\top$. What is its inverse?

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



$$\begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix} \rightarrow \begin{bmatrix} \frac{1}{\sigma_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_r} \end{bmatrix}$$

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

$$A A^{-1} = U \Sigma \underline{V^T V} \Sigma^{-1} U^T = I$$

Applications of SVD: Matrix Pseudo Inverse

Definition: Given $A \in \mathbb{R}^{m \times n}$ with $\operatorname{rank}(A) = r$ and SVD:

$$A = \underline{U \Sigma V^\top} = U \begin{bmatrix} \Sigma_r & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{bmatrix} V^\top \leftarrow \in \mathbb{R}^{m\times n}$$

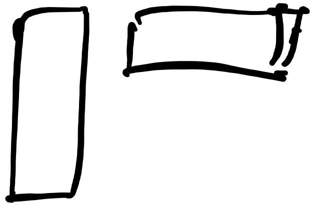
its (Moore-Penrose) **pseudo inverse** is defined to be:

$$A^\dagger = \underline{V} \begin{bmatrix} \Sigma_r^{-1} & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{bmatrix} \underline{U^T} = \underline{V_r \Sigma_r^{-1} U_r^\top} \quad \in \mathbb{R}^{n\times m}$$

Example: $A = \begin{bmatrix} 4 & 4 \\ 3 & 3 \end{bmatrix}, \quad A^\dagger = ?$ $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{5\sqrt{2}} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{5} \begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix}$

Applications of SVD: Matrix Pseudo Inverse

① $m=n=r \quad A^\dagger = V\Sigma^{-1}U^T = A^{-1}$



② $r < \min\{m,n\}$

$$AA^\dagger = U_r\Sigma_r V_r^T\,(V_r\Sigma_r^{-1}U_r^T) = U_rU_r^T = \sum_{i=1}^{r}\vec{u}_i\vec{u}_i^T$$

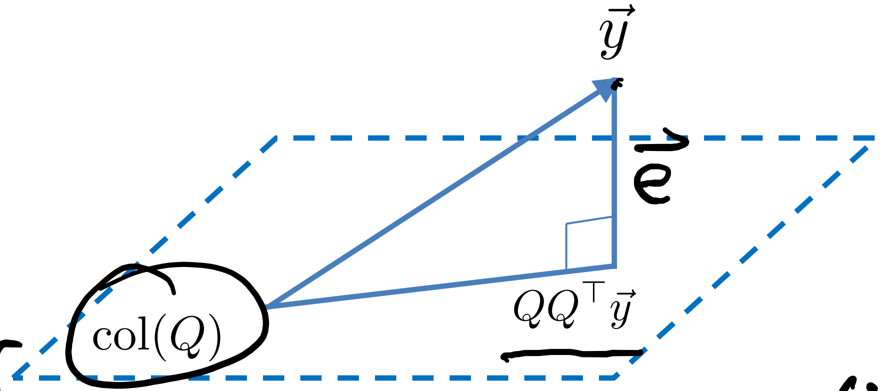
③ $$A^\dagger A = (V_r\Sigma_r^{-1}U_r^T)\,U_r\Sigma_r V_r^T = V_rV_r^T = \sum_{i=1}^{r}\vec{v}_i v_i^T$$

projection matrix

Applications of SVD: Matrix Pseudo Inverse

Geometric interpretation of $AA^\dagger$ or $A^\dagger A$.

$\underline{AA^\dagger} = U_r U_r^T \qquad \underline{A^\dagger A} = V_r V_r^T$



Q - orthogonal $\quad Q = [\vec{q}_1, \vec{q}_2, \ldots \vec{q}_k]$

$\in \mathbb{R}^{n \times k} \qquad k \le n \qquad Q^TQ = I_{k\times k} \quad \underbrace{QQ^T}_{n\times n} = \sum_{i=1}^{k} \vec{q}_i \vec{q}_i^T$

projection matrix

$\vec{y} \in \mathbb{R}^n \qquad (QQ^T)\vec{y} = \underline{Q(Q^T\vec{y})} \in \mathrm{col}(Q)$

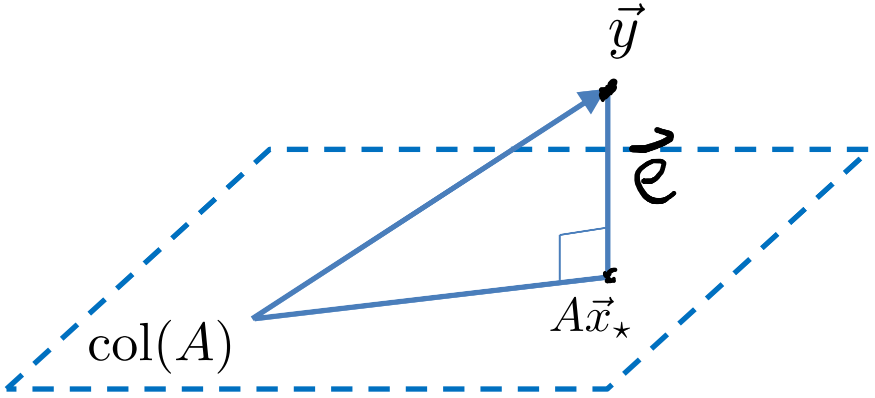
$\vec{e} = \underline{\vec{y} - QQ^T\vec{y}} \qquad \langle \vec{e}, QQ^T y \rangle = \langle \vec{y} - QQ^T\vec{y}, QQ^T\vec{y} \rangle$

$= \vec{y}^T QQ^T \vec{y} - \vec{y}^T QQ^T \vec{y} = 0.$

Applications of SVD: Least Squares

$\min_{\vec{x}} \|\vec{y} - A\vec{x}\|_2^2$, with $A \in \mathbb{R}^{m\times n}$ and $\text{rank}(A) = n$: $\quad \vec{x}_\star = (A^\top A)^{-1} A^\top \vec{y}$ ✓

an alternative way.



$A\vec{x}_*$ - orthogonal projection onto col(A) = col(U_r)

$AA^\dagger = U_r U_r^\top$ - proj. onto col(U_r)

$AA^\dagger \vec{y}$ - proj. of $\vec{y}$ onto col(A)

$A\vec{x}_* = AA^\dagger\vec{y} \implies \vec{x}_* = A^\dagger \vec{y}$. $\boxtimes$

Applications of SVD: Least Squares

Show: Given $A \in \mathbb{R}^{m\times n}$ and $\operatorname{rank}(A) = n$: $A^\dagger = (A^\top A)^{-1}A^\top$?

$$(A^T A)^{-1} A^T \leftarrow A = U_r \Sigma_r V^T \quad (r = n)$$

$$= (V \Sigma_r U_r^T U_r \Sigma_r V^T)^{-1} V \Sigma_r U_r^T$$

$$= (V \Sigma_r^2 V^T)^{-1} V \Sigma_r U_r^T$$

$$= V \Sigma_r^{-2} V^T V \Sigma_r U_r^T$$

$$= V \Sigma^{-1} U_r^T = A^\dagger \quad \boxtimes$$

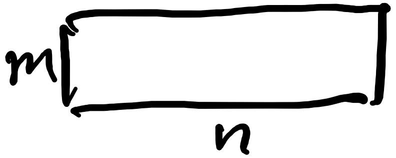
$$\frac{1}{\sigma_i^2} \sigma_i = \frac{1}{\sigma_i}$$

$$\Sigma_r^{-2} \Sigma_r$$

Applications of SVD: Minimum Norm Solution

$\min_{\vec{x}} \|\vec{x}\|_2^2$ s.t. $\vec{y} = A\vec{x}$, with $A \in \mathbb{R}^{m\times n}$ and $\operatorname{rank}(A) = m$: $\vec{x}_\star = A^\top(AA^\top)^{-1}\vec{y}$

Show: $\vec{x}_\star = A^\dagger \vec{y}\ (= A^\top(AA^\top)^{-1}\vec{y})$.



$\vec{y} = A(\vec{x} + \vec{s}) \quad \vec{s} \in Nu(A)$

$\vec{y} = A\vec{x} = U_r \Sigma_r V_r^T \vec{x} \Rightarrow U_r^T \vec{y} = \Sigma_r V_r^T \vec{x}$

$\Rightarrow \Sigma_r^{-1} U_r^T \vec{y} = V_r^T \vec{x}$

$V^T = \begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix}_{n\times n}$

$\|V^T\vec{x}\|_2^2 = \|\vec{x}\|_2^2 = \left\| \begin{bmatrix} V_r^T \vec{x} \\ V_{n-r}^T \vec{x} \end{bmatrix} \right\|_2^2 \leftarrow$ fixed.

$\begin{bmatrix} V_r^T \vec{x}_* \\ V_{n-r}^T \vec{x}_* \end{bmatrix} = \begin{bmatrix} \Sigma_r^{-1} U_r^T \vec{y} \\ 0 \end{bmatrix} \Rightarrow V^T \vec{x}_* = \begin{bmatrix} \Sigma_r^{-1} U_r^T \vec{y} \\ 0 \end{bmatrix}$

$$\vec{x}_* = [V_r, V_{n-r}] \begin{bmatrix} \Sigma_r^{-1} U_r^T \vec{y} \\ 0 \end{bmatrix} = \underbrace{V_r \Sigma_r^{-1} U_r^T}_{A^\dagger} \vec{y}$$

$$\vec{x}_* \in \mathrm{col}(V_r) \perp \mathrm{Null}(A)$$

$$\vec{y} = A(\underline{\vec{x}_* + \vec{s}}) \qquad \vec{s} \in \underline{\mathrm{Null}(A)} = \mathrm{col}(V_{n-r})$$

Applications of SVD: Minimum Norm Solution

$$\vec{y} = A(\underbrace{\vec{x}_* + \vec{s}}_{\vec{x}'})$$

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