THE SINGULAR VALUE DECOMPOSITION

- Orthogonal subspaces
- The Singular Value Decomposition
- Properties of the SVD. Relations to eigenvalue problems

The Singular Value Decomposition (SVD)

Theorem For any matrix $A \in \mathbb{R}^{m\times n}$ there exist unitary matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ such that

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

where Σ is a diagonal matrix with entries $\sigma_{ii} \geq 0$.

$$\sigma_{11} \geq \sigma_{22} \geq \cdots \sigma_{pp} \geq 0 \text{ with } p = \min(n, m)$$

➤ The σ_{ii}'s are the singular values. Notation change $\sigma_{ii} \longrightarrow \sigma_i$

Proof: Let $\sigma_1 = \|A\|_2 = \max_{x, \|x\|_2=1} \|Ax\|_2$. There exists a pair of unit vectors v_1, u_1 such that

$$Av_1 = \sigma_1 u_1$$

➤ Complete v_1 into an orthonormal basis of $\mathbb{R}^n$

$$V \equiv [v_1, V_2] = n \times n \text{ unitary}$$

➤ Complete u_1 into an orthonormal basis of $\mathbb{R}^m$

$$U \equiv [u_1, U_2] = m \times m \text{ unitary}$$

✍ 1 Define U, V as single Householder reflectors.

➤ Then, it is easy to show that

$$AV = U \times \begin{pmatrix} \sigma_1 & w^T \\ 0 & B \end{pmatrix} \rightarrow U^T AV = \begin{pmatrix} \sigma_1 & w^T \\ 0 & B \end{pmatrix} \equiv A_1$$

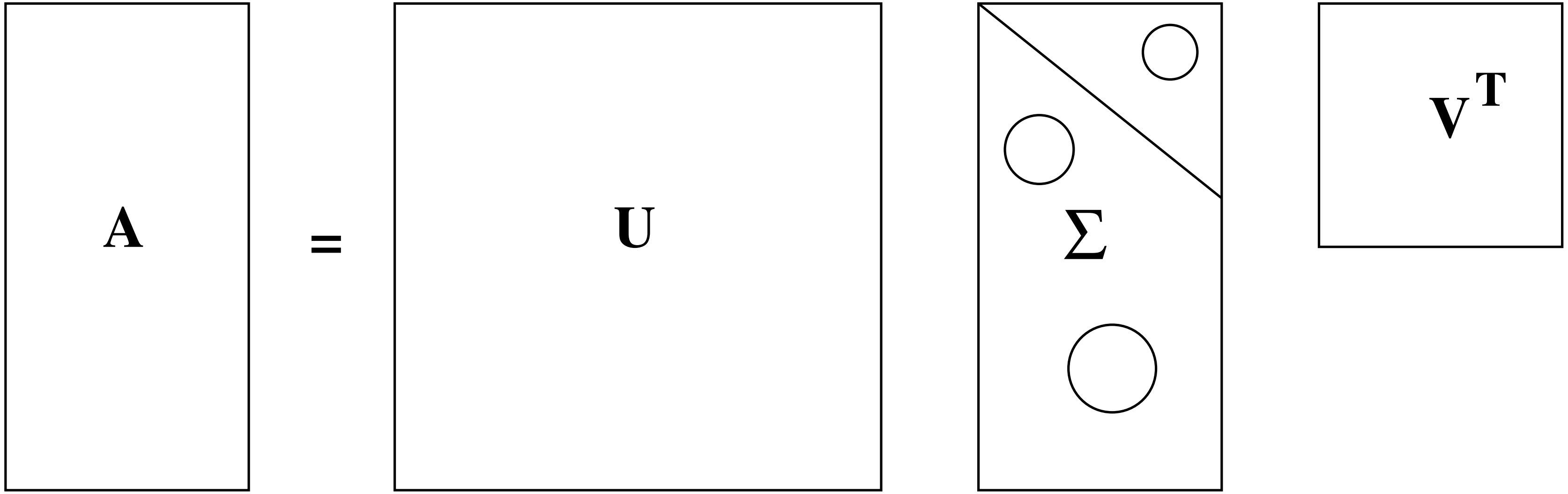
- Observe that

$$\left\| A_1 \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2 \ge \sigma_1^2 + \|w\|^2 = \sqrt{\sigma_1^2 + \|w\|^2} \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2$$

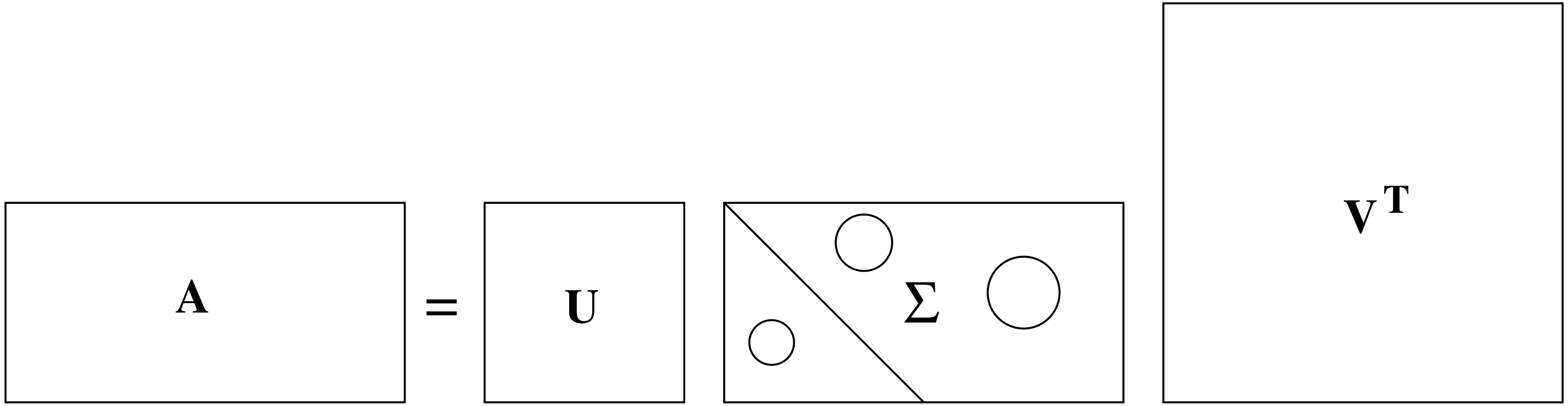
- This shows that w must be zero [why?]

- Complete the proof by an induction argument. $\blacksquare$

Case 1:



Case 2:



The "thin" SVD

➤ Consider Case-1. It can be rewritten as

$$\boldsymbol{A} = [\boldsymbol{U}_1 \boldsymbol{U}_2] \begin{pmatrix} \boldsymbol{\Sigma}_1 \\ \boldsymbol{0} \end{pmatrix} \boldsymbol{V}^T$$

Which gives:

$$\boxed{\boldsymbol{A} = \boldsymbol{U}_1 \boldsymbol{\Sigma}_1 \boldsymbol{V}^T}$$

where $\boldsymbol{U}_1$ is $\boldsymbol{m} \times \boldsymbol{n}$ (same shape as $\boldsymbol{A}$), and $\boldsymbol{\Sigma}_1$ and $\boldsymbol{V}$ are $\boldsymbol{n} \times \boldsymbol{n}$

➤ Referred to as the "thin" SVD. Important in practice.

✍2 How can you obtain the thin SVD from the QR factorization of $\boldsymbol{A}$ and the SVD of an $\boldsymbol{n} \times \boldsymbol{n}$ matrix?

A few properties. Assume that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \text{ and } \sigma_{r+1} = \cdots = \sigma_p = 0$$

Then:

- $\mathbf{rank}(A) = r$ = number of nonzero singular values.
- $\mathbf{Ran}(A) = \mathbf{span}\{u_1, u_2, \ldots, u_r\}$
- $\mathbf{Null}(A^T) = \mathbf{span}\{u_{r+1}, u_{r+2}, \ldots, u_m\}$
- $\mathbf{Ran}(A^T) = \mathbf{span}\{v_1, v_2, \ldots, v_r\}$
- $\mathbf{Null}(A) = \mathbf{span}\{v_{r+1}, v_{r+2}, \ldots, v_n\}$

Properties of the SVD (continued)

- The matrix A admits the SVD expansion:

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

- $\|A\|_2 = \sigma_1$ = largest singular value
- $\|A\|_F = \left(\sum_{i=1}^{r} \sigma_i^2\right)^{1/2}$
- When A is an $n \times n$ nonsingular matrix then $\|A^{-1}\|_2 = 1/\sigma_n$

Theorem [Eckart-Young-Mirsky] Let $k \le r$ and $A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$ then

$$\min_{rank(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

Proof: First: $\|A - B\|_2 \geq \sigma_{k+1}$, for any rank-$k$ matrix B.

Consider $\mathcal{X} = \mathbf{span}\{v_1, v_2, \cdots, v_{k+1}\}$. Note:

$$dim(Null(B)) = n - k \rightarrow Null(B) \cap \mathcal{X} \neq \{0\}$$

[Why?]

Let $x_0 \in Null(B) \cap \mathcal{X},\ x_0 \neq 0$. Write $x_0 = Vy$. Then

$$\|(A - B)x_0\|_2 = \|Ax_0\|_2 = \|U\Sigma V^T Vy\|_2 = \|\Sigma y\|_2$$

But $\|\Sigma y\|_2 \geq \sigma_{k+1}\|x_0\|_2$ (Show this). $\rightarrow \|A - B\|_2 \geq \sigma_{k+1}$

Second: take $B = A_k$. Achieves the min. $\square$

Right and Left Singular vectors:

$$\begin{aligned} Av_i &= \sigma_i u_i \\ A^T u_j &= \sigma_j v_j \end{aligned}$$

- v_i's = right singular vectors;
- u_i's = left singular vectors.

- Consequence $\boxed{A^T A v_i = \sigma_i^2 v_i}$ and $\boxed{A A^T u_i = \sigma_i^2 u_i}$
- Right singular vectors (v_i's) are eigenvectors of $A^T A$
- Left singular vectors (u_i's) are eigenvectors of $A A^T$
- Possible to get the SVD from eigenvectors of AA^T and $A^T A$ – but: difficulties due to non-uniqueness of the SVD

Define the $r \times r$ matrix

$$\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$$

➤ Let $A \in \mathbb{R}^{m\times n}$ and consider A^TA $(\in \mathbb{R}^{n\times n})$:

$$A^TA = V\Sigma^T\Sigma V^T \;\rightarrow\; A^TA = V \underbrace{\begin{pmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{pmatrix}}_{n\times n} V^T$$

➤ This gives the spectral decomposition of A^TA.

➤ Similarly, U gives the eigenvectors of AA^T.

$$AA^T = U \underbrace{\begin{pmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{pmatrix}}_{m \times m} U^T$$

Important:

$A^TA = VD_1V^T$ and $AA^T = UD_2U^T$ give the SVD factors U, V up to signs!