The Singular Value Decomposition (SVD)

Objectives

- Define singular value decomposition (SVD)
- Express skinny SVD
- Write SVD as sum of outer products
- Use SVD to find best low-rank approximation
- Interpret matrix as an operator

SVD

- matrix decomposition that leads to good low-rank approximations
- vast range of applications

Definition:

Any $N \times M$ matrix $\underline{A}$ can be written as

$$\underline{A} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T$$

- $\underline{U}$: $N \times N$, orthonormal columns
- $\underline{V}$: $M \times M$, orthonormal columns
- $\underline{\Sigma}$: $N \times M$, diagonal, $\underline{\Sigma}_{ii} \geq 0$

$N > M$

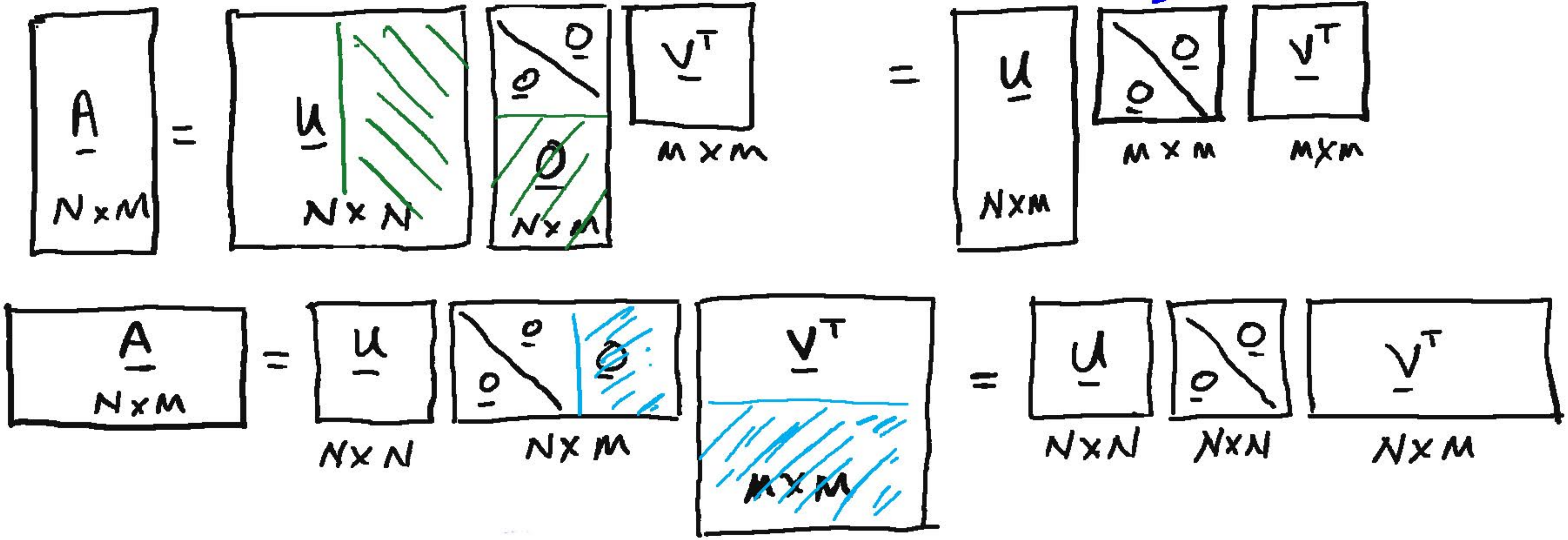
$$\begin{bmatrix} \sigma_1 & & \underline{0} \\ & \sigma_2 \ddots & \\ \underline{0} & & \sigma_M \\ & \underline{0} & \end{bmatrix}$$

$M > N$

$$\begin{bmatrix} \sigma_1 & & \underline{0} & \\ & \sigma_2 & & \underline{0} \\ \underline{0} & & \ddots & \\ & & \sigma_N & \end{bmatrix}$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{N/M} \geq 0$$

SVD Dimensions   Skinny SVD



Sum of Outer Products Form:

"rank 1"

$$\underline{A} = \begin{bmatrix} | & | & & | \\ \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_M \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \underline{0} \\ & \sigma_2 & & \\ & & \ddots & \\ \underline{0} & & & \sigma_M \end{bmatrix} \begin{bmatrix} - & \underline{v}_1^T & - \\ - & \underline{v}_2^T & - \\ & \vdots & \\ - & \underline{v}_M^T & - \end{bmatrix} = \sum_{i=1}^{M} \sigma_i \underline{u}_i \underline{v}_i^T = \sum_{i=1}^{M} \boxed{\begin{matrix} \sigma_i \underline{u}_i \underline{v}_i^T \\ N \times M \end{matrix}}$$

SVD gives the "best" low-rank approximation

Frobenius norm $\|\underline{A}\|_F^2 = \sum_{i=1}^{N} \sum_{j=1}^{M} \left([\underline{A}]_{i,j}\right)^2 = \|\mathrm{vec}(\underline{A})\|_2^2$

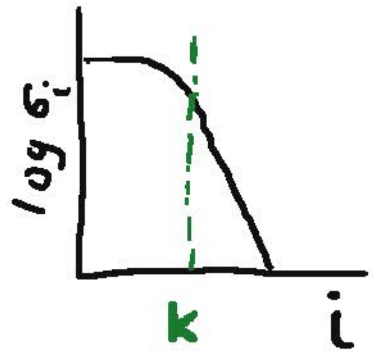
Eckart-Young Theorem (1936) Let $\mathrm{rank}(\underline{A}) = r$ and $k < r$: $\displaystyle\min_{\mathrm{rank}(\underline{B}) \le k} \|\underline{A} - \underline{B}\|_F^2 = \sum_{i=k+1}^{r} \sigma_i^2$ for $\underline{B} = \sum_{i=1}^{k} \sigma_i \underline{u}_i \underline{v}_i^T$

where $\underline{A} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T$ is the SVD.

$$\boxed{\underline{A}} \approx \sigma_1 \boxed{\underline{u}_1 \underline{v}_1^T} + \sigma_2 \boxed{\underline{u}_2 \underline{v}_2^T} + \cdots + \sigma_k \boxed{\underline{u}_k \underline{v}_k^T}$$

cols: scaled $\underline{u}_i$
rows: scaled $\underline{v}_i^T$

patterns: most important 2^{nd} most k^{th} most



σ_i provide ordered ranking of components

SVD describes matrix as an operator

$\underline{A}: N \times M$, $\underline{x}: M \times 1$, $\underline{y}: N \times 1$

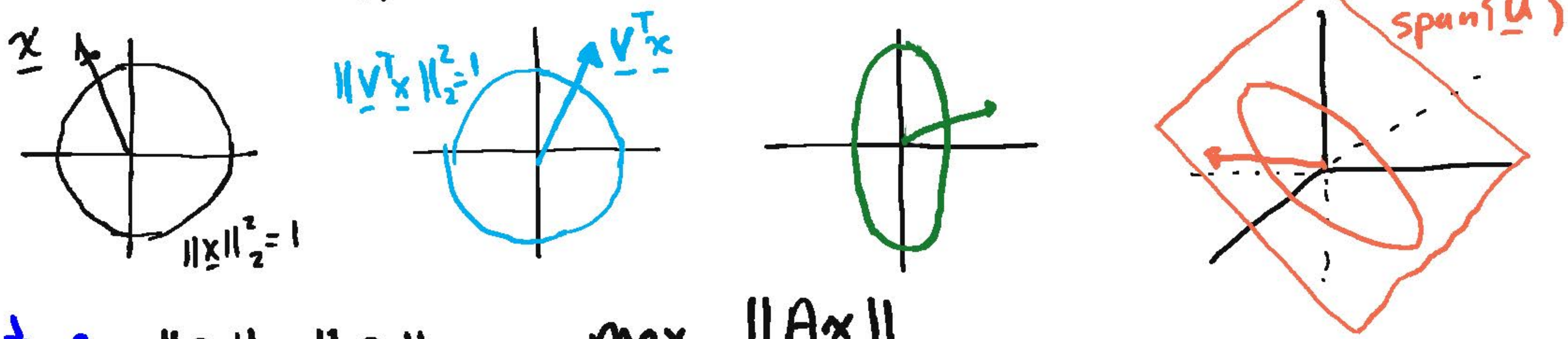
$$\underline{y} = \underline{A}\,\underline{x} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T \underline{x} = \underline{U}\left[\underline{\Sigma}\left(\underline{V}^T \underline{x}\right)\right]$$

$$x: \mathbb{R}_M \xrightarrow[\text{rotate}]{\underline{V}^T} \underline{V}^T\underline{x}: \mathbb{R}_M \xrightarrow[\Sigma]{\text{scale}} \underline{\Sigma}\,\underline{V}^T\underline{x}: \mathbb{R}_M \xrightarrow[\text{rotate (up dim)}]{\underline{U}} \underline{y}: \mathbb{R}_N$$

Note: $\|\underline{V}^T\underline{x}\|_2^2 = \underline{x}^T \underbrace{\underline{V}\,\underline{V}^T}_{\text{o/n}} \underline{x} = \underline{x}^T\underline{x} = \|\underline{x}\|_2^2$

$N = 3$

$M = 2$



Operator Norm $\|\underline{A}\|_2 = \|\underline{A}\|_{op} := \max_{x \neq 0} \dfrac{\|\underline{A}\underline{x}\|_2}{\|\underline{x}\|_2} = \sigma_1$

(proof: notes)