Plan

Logistics

Review

Singular Value Decomposition

Low-Rank Approximation

Great job on problem set!

↳ Come talk to me with questions

Goal: Less problem time outside class

1. I'll provide guidance
2. Make sure you have *rough* solution before you leave
3. Calibrate to my solutions
4. Start in assigned groups

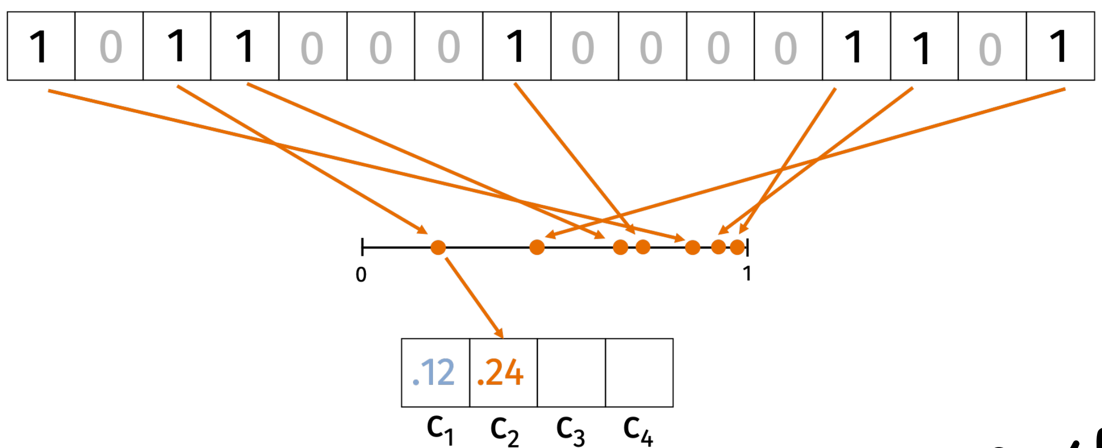
Self-grade

Proposal

Locality Sensitive Hashing

Find "similar" vectors

e.g. song



$$\Pr(c_i(x) = c_i(y)) = \frac{|X \cap Y|}{|X \cup Y|} = J(x,y)$$

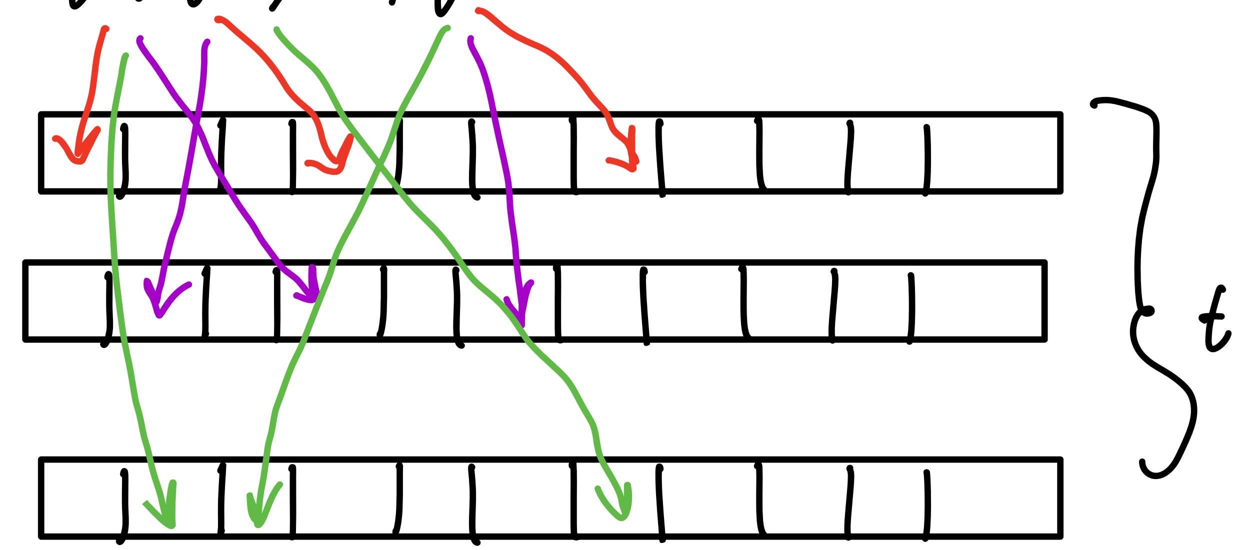
$$c_1, \dots, c_r : \{0,1\}^d \to [0,1]$$

$$g : [0,1]^r \to \{1, \dots, m\}$$

$$\Pr(g(x) = g(y)) = \Pr(c_i(x) = c_i(y))^r$$

Preprocess

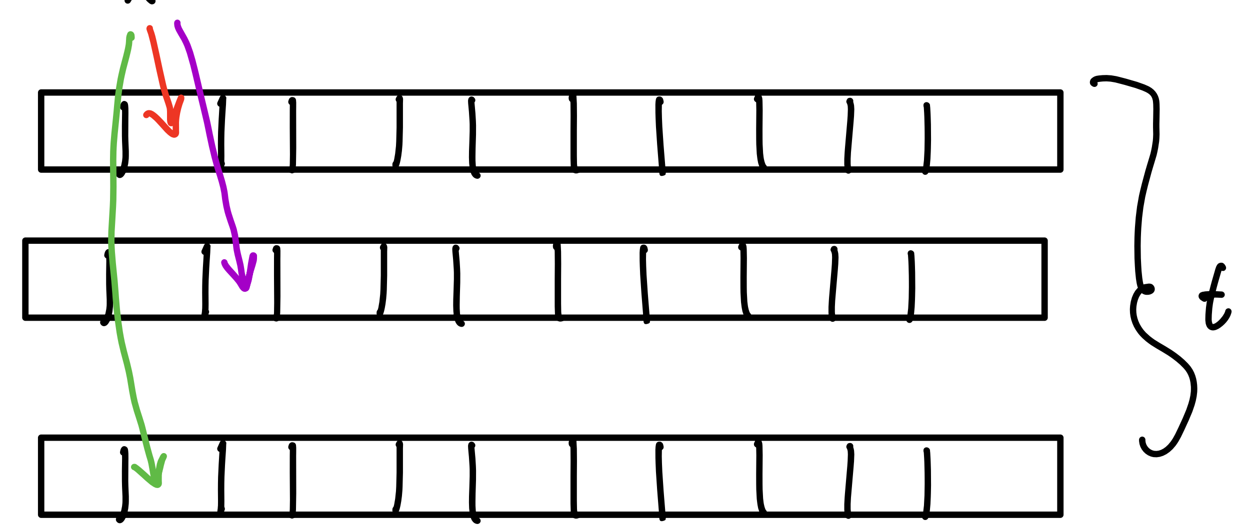
$y_1, y_2, \dots, y_n$



Query

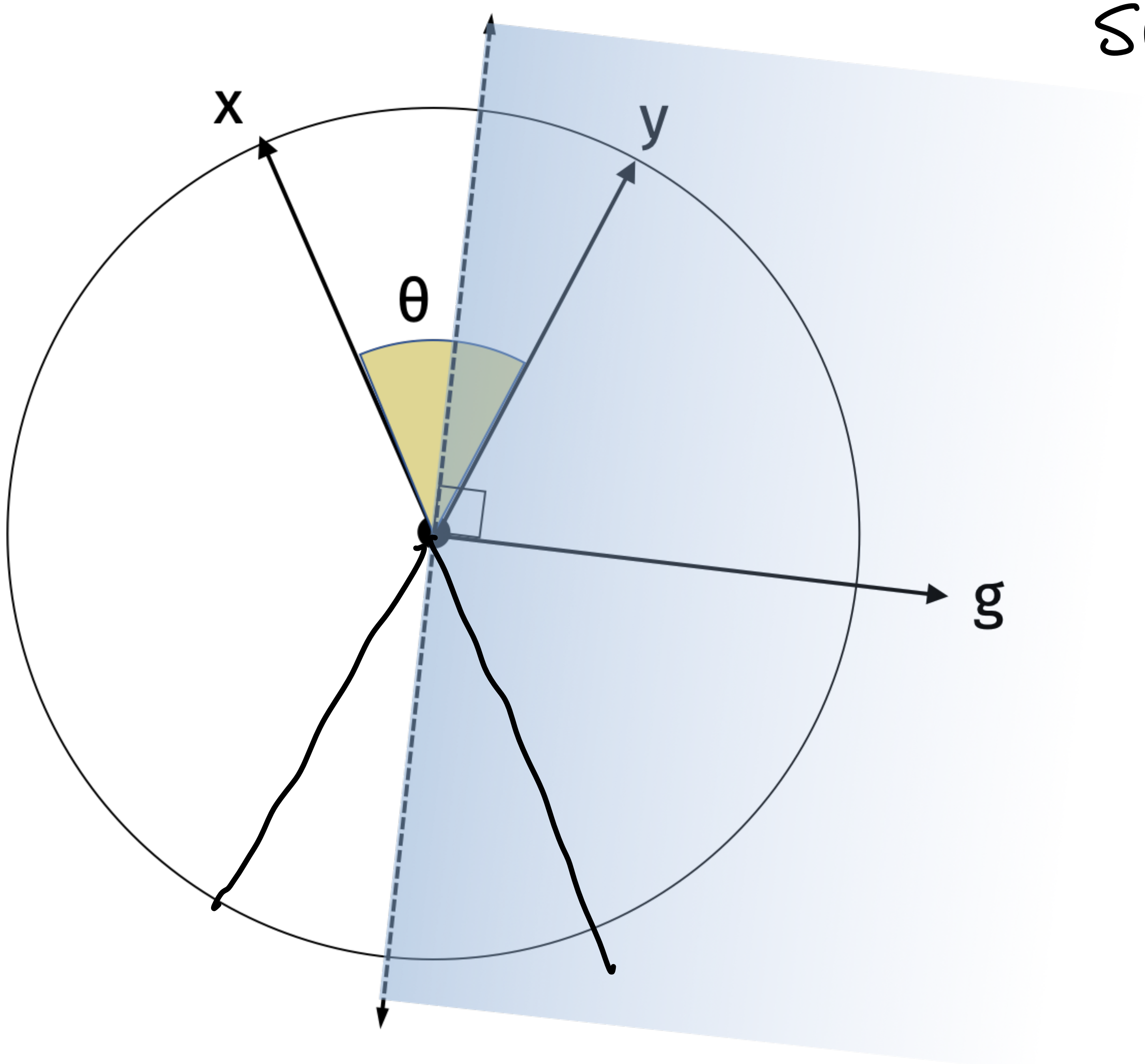
Hits: y.

x



$$\Pr(\text{find } y) = 1 - (1 - J(x,y)^r)^t$$

or $\left(1 - \frac{\theta}{\pi}\right)$



$\mathrm{sign}(\langle g, x \rangle)$

$$\Pr\left(\mathrm{sign}(\langle g, x \rangle) = \mathrm{sign}(\langle g, y \rangle)\right)$$

$$1 - \frac{2\theta}{2\pi} = 1 - \frac{\theta}{\pi}$$

$$\cos\theta = \frac{\langle x, y \rangle}{\|x\|_2 \, \|y\|_2}$$

Linear Algebra

$X \in \mathbb{R}^{d \times d}$

Eigenvector $v \in \mathbb{R}^d$ $\|v\|_2 = 1$

and eigenvalue $\lambda \in \mathbb{R}$ if

$$Xv = \lambda v$$

Suppose X has d eigenvector/values

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$

$v_1, \; v_2, \; \dots, v_3$

$\langle v_i, v_j \rangle = 0$ if $i \neq j$

$$V = \underset{d \times d}{\begin{bmatrix} \text{---} \; v_1 \; \text{---} \\ \text{---} \; v_2 \; \text{---} \\ \vdots \\ \text{---} \; v_d \; \text{---} \end{bmatrix}} \qquad \Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_d \end{bmatrix}$$

$$X = V \Lambda V^T$$

$$V V^T = I = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots & \\ & & & 1 \end{bmatrix}$$

$$= \underset{V}{\begin{bmatrix} \text{---} \; v_1 \; \text{---} \\ \text{---} \; v_2 \; \text{---} \\ \vdots \\ \text{---} \; v_d \; \text{---} \end{bmatrix}} \underset{V^T}{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_d \\ | & | & & | \end{bmatrix}}$$

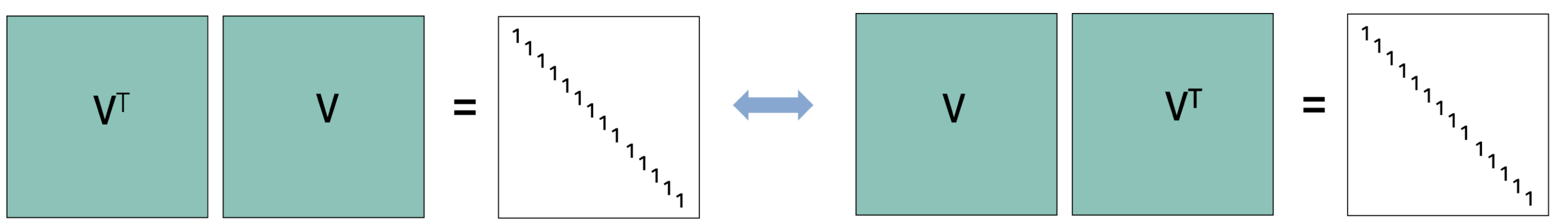
$$V^{-1} = V^T$$

$$V V^{-1} = I$$

$$V^T (V^T)^{-1} = I$$

$$= V^T (V^{-1})^{-1}$$

$$= V^T V$$



By definition, $\|X\|_F^2 = \sum_{i=1}^{d} \sum_{j=1}^{d} (X_{i,j})^2$

(3) $\|VX\|_F^2 = \|X\|_F^2$

(1) Show $\|\underset{d\times d}{V}\,\underset{d\times 1}{x}\|_2^2 = \langle Vx, Vx \rangle = \underset{1\times d}{(Vx)^T}\underset{d\times 1}{(Vx)} = x^T V^T V x = x^T I x$

$= \|x\|_2^2$

(2) Show $\|V^T x\|_2^2 = (V^T x)^T (V^T x) = x^T V V^T x = \|x\|_2^2$

Hint: $(Vx)^T = x^T V^T$

In general, $(AB)^T = B^T A^T$

$$\|VX\|_F^2 = \sum_{i=1}^{d} \sum_{j=1}^{d} (VX)_{i,j}^2$$

$$= \sum_{i=1}^{d} \sum_{j=1}^{d} \langle v_i , x_j \rangle^2$$

$$\leftarrow \|v_i X\|_2^2$$

$$= \sum_{j=1}^{d} \sum_{i=1}^{d} \langle v_i , x_j \rangle^2$$

$$= \sum_{j=1}^{d} \| V x_j \|_2^2$$

$$= \sum_{j=1}^{d} \| x_j \|_2^2$$

$$= \| X \|_F^2$$

$$\underset{d \times d}{\overset{V}{\begin{bmatrix} \quad & \quad \\ & \end{bmatrix}}} \quad \underset{d \times d}{\overset{X}{\begin{bmatrix} \quad & \quad \\ & \end{bmatrix}}}$$

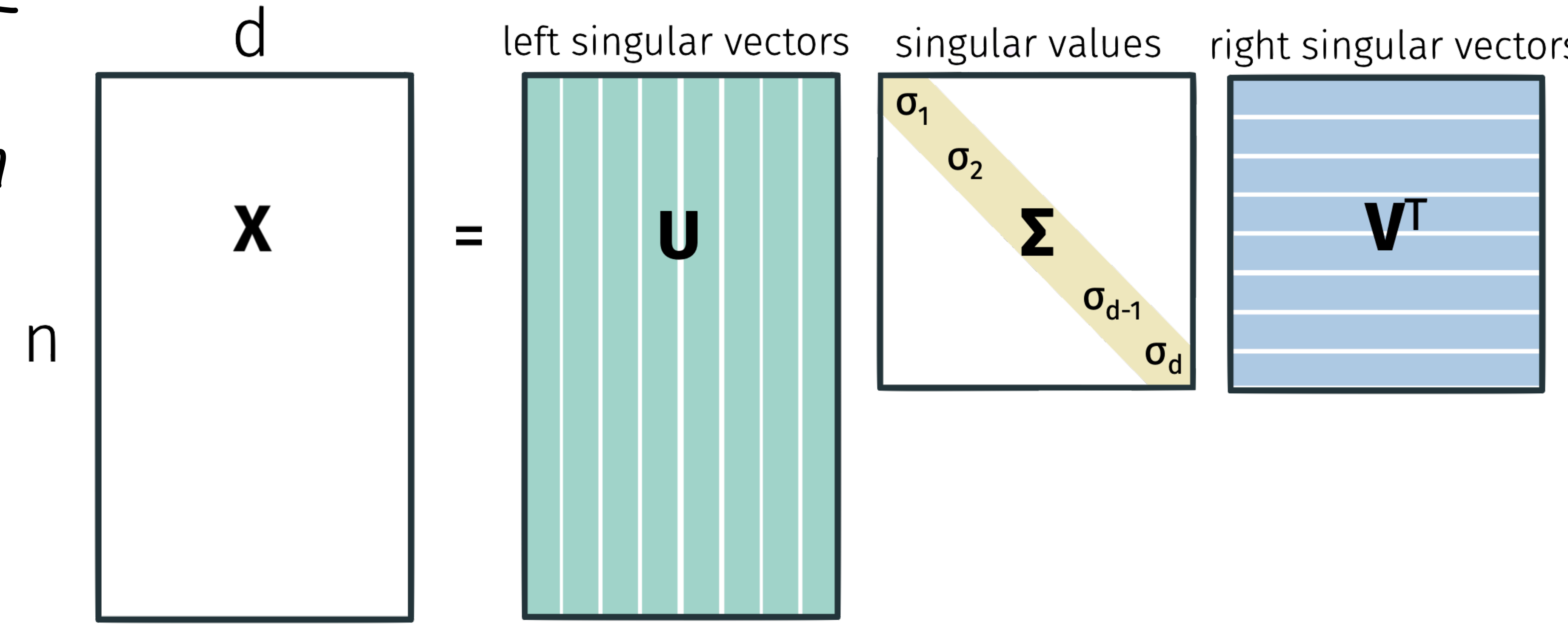
$$\underset{d \times d}{\begin{bmatrix} \langle v_1, x_1 \rangle & \\ & \end{bmatrix}}$$

$$\underset{d \times d}{\begin{bmatrix} | & \\ x_j & \\ | & \end{bmatrix}} \quad X$$

Singular Value Decomposition

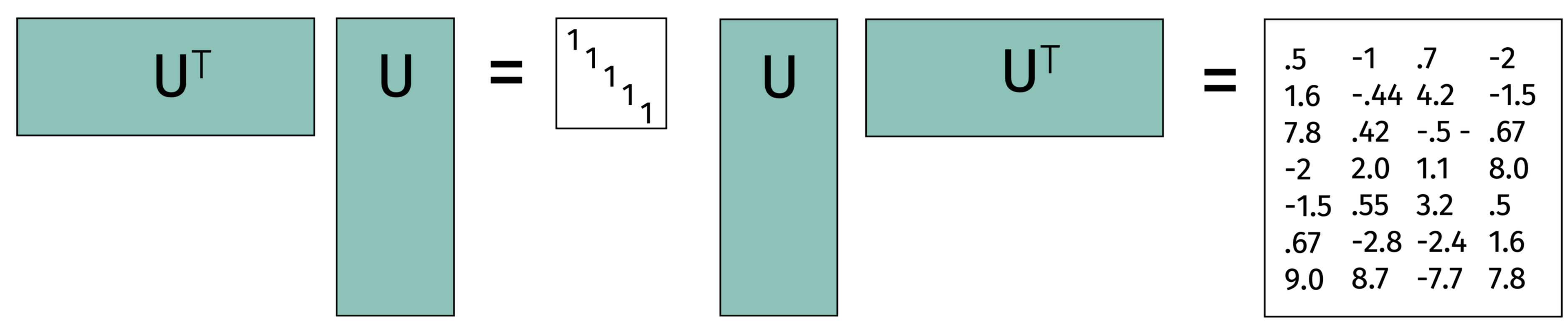
$X \in \mathbb{R}^{n \times d} \quad n \geq d \quad \text{WLOG}$

$$\underset{n\times d}{X} = \underset{n \times d}{U} \, \underset{d\times d}{\Sigma} \, \underset{d \times d}{V^T}$$



$U^T U = I \qquad V^T V = I$

$U^T U = I \quad \text{but} \quad U U^T \neq I$

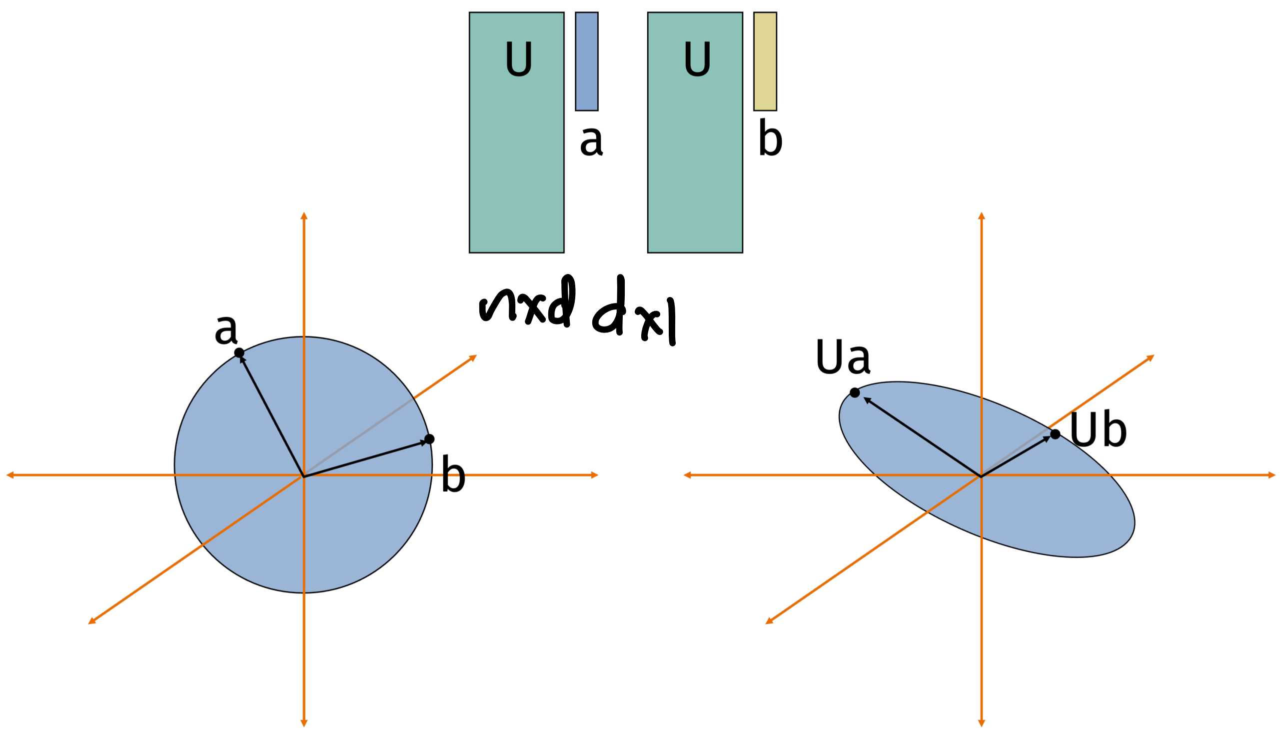


$$\Sigma = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_d \end{bmatrix}$$

$\sigma_1 \geq \sigma_2 \geq \dots \sigma_d \geq 0$

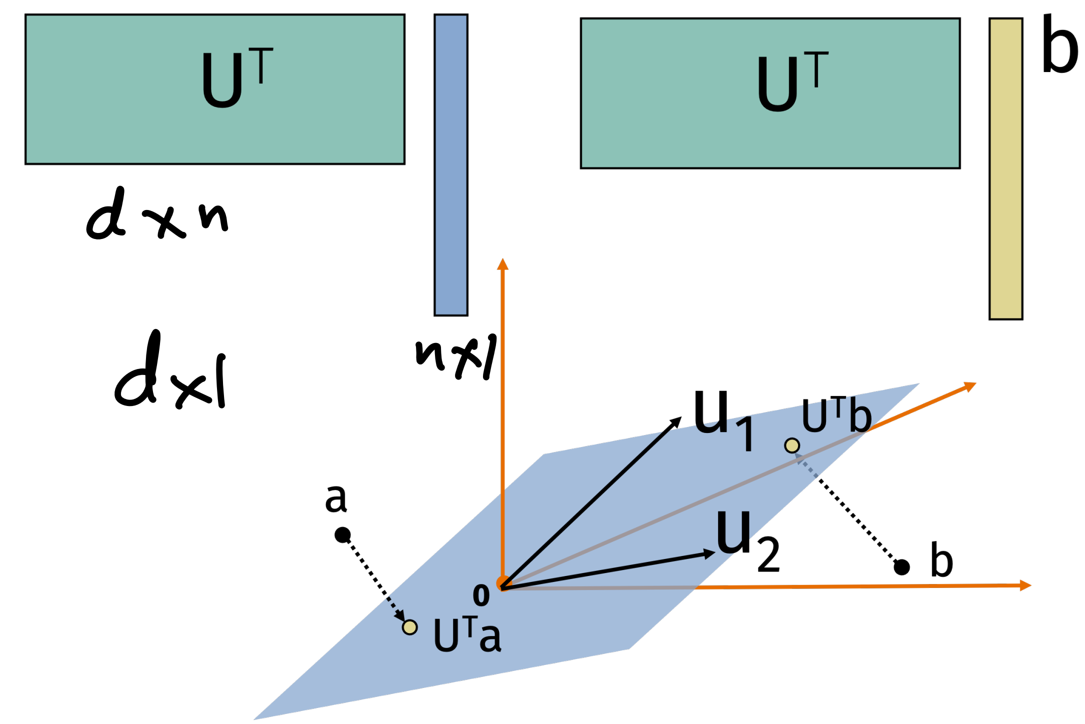
why?

Multiply by U : rotation



$$\|U x\|_2^2 = \|x\|_2^2$$

Multiply by U^T : projection



$$Xa = U\Sigma V^T a = U\left(\Sigma\left(V^T a\right)\right)$$

1. Rotates a
2. Scale coordinates
3. Rotating

Eigendecomposition vs SVD

Eigendecomposition	SVD
$X \in \mathbb{R}^{d \times d}$	$X \in \mathbb{R}^{n \times d}$
λ_i	$\sigma_i \geq 0$
V orthonormal cols	U, V orthonormal columns but not necessarily orthonormal rows

$$\Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots & \\ & & & \sigma_d \end{bmatrix}$$

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

$$\begin{aligned}
X^T X &= (U \Sigma V^T)^T U \Sigma V^T \\
&= V \Sigma^T U^T U \Sigma V^T \\
&= V \Sigma \Sigma V^T \\
&= V \Sigma^2 V^T \\
&\triangleq V \Lambda V^T
\end{aligned}$$

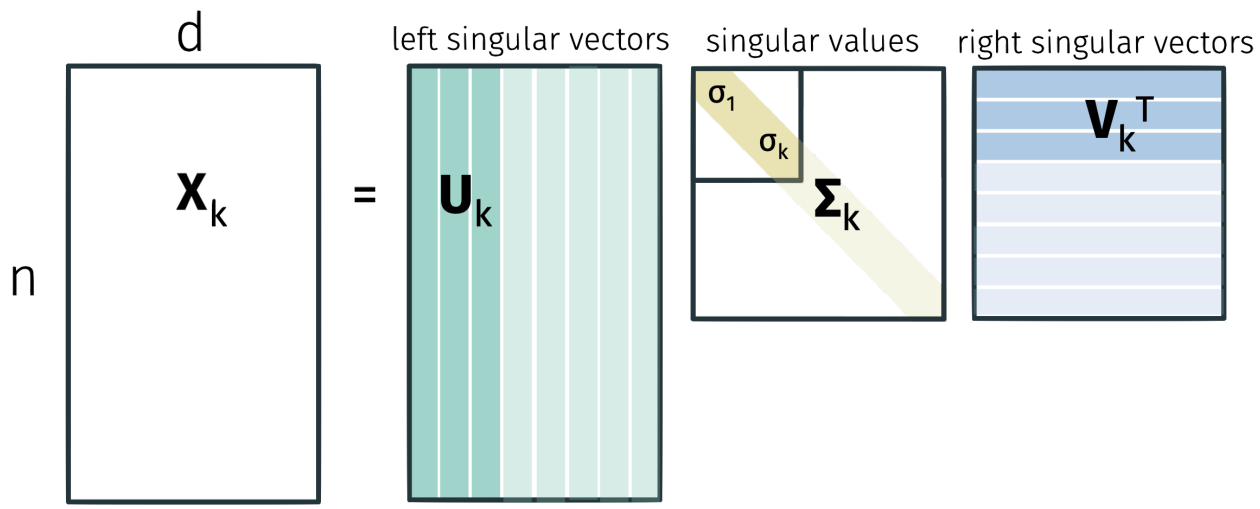
$$\lambda_i = \sigma_i^2$$

Low Rank Approximation

$X \in \mathbb{R}^{n \times d}$ $\quad n \geq d \quad O(nd)$

$X_k = U_k \Sigma_k V_k^T \quad O(nk)$

$n \times k \quad k \times k \quad k \times d$



$X_k \approx X$

Tools

① $\|VX\|_F^2 = \|X\|_F^2$ if $V^TV = I$

② $\|X\|_F^2 = \|X^T\|_F^2$

③ $\|X\|_F^2 = \sum_{i=1}^{d} \sigma_i^2$

Best rank k approximation

$$\underset{\text{rank } k\ B}{\operatorname{argmin}} \|X - B\|_F^2 = \operatorname{argmin} \|U \Sigma V^T - B\|_F^2$$

$$= \operatorname{argmin} \|U^T U \Sigma V^T - U^T B\|_F^2 \quad \text{by } ① \qquad U^T U = I$$

$$= \operatorname{argmin} \|\Sigma V^T - U^T B\|_F^2$$

$$= \operatorname{argmin} \|V \Sigma - B^T U\|_F^2 \quad \text{by } ②$$

$$V^T V = I$$

$$= \operatorname{argmin} \|\Sigma - V^T B^T U\|_F^2$$

$$B = \Sigma_k = \begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_k & \\ 0 & & & 0 \end{bmatrix}$$

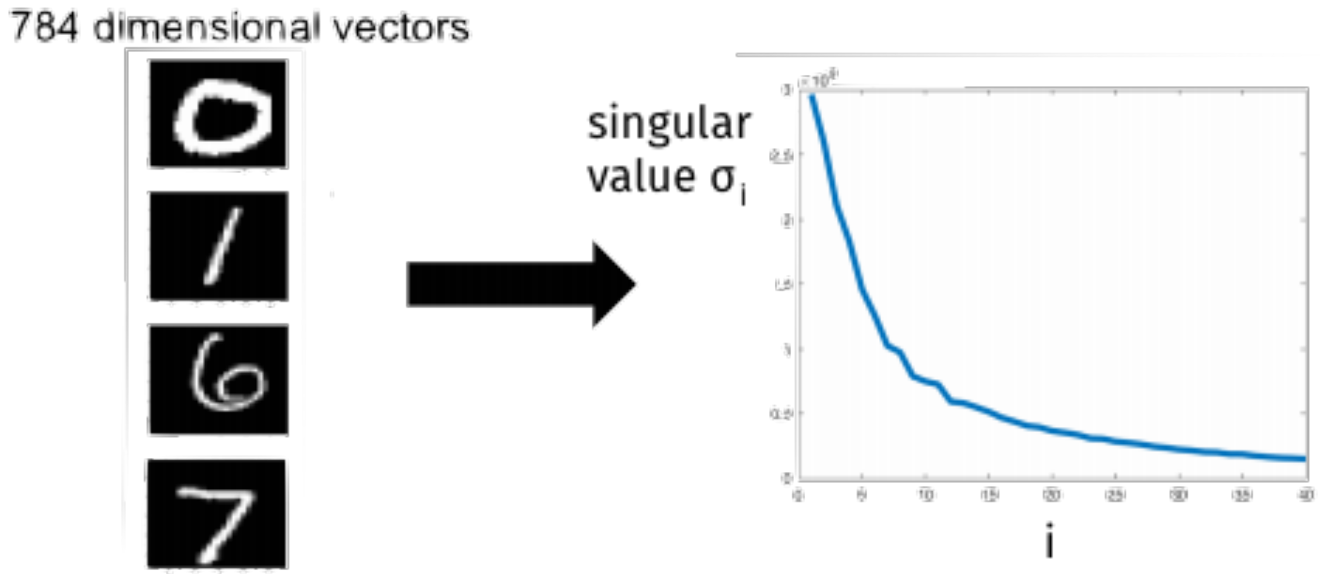
$$\| X - X_k \|_F^2 = \| U \Sigma V^T - U \Sigma_k V^T \|_F^2$$

$$= \| U (\Sigma - \Sigma_k) V^T \|_F^2 \qquad U^T U = I \quad V^T V = I$$

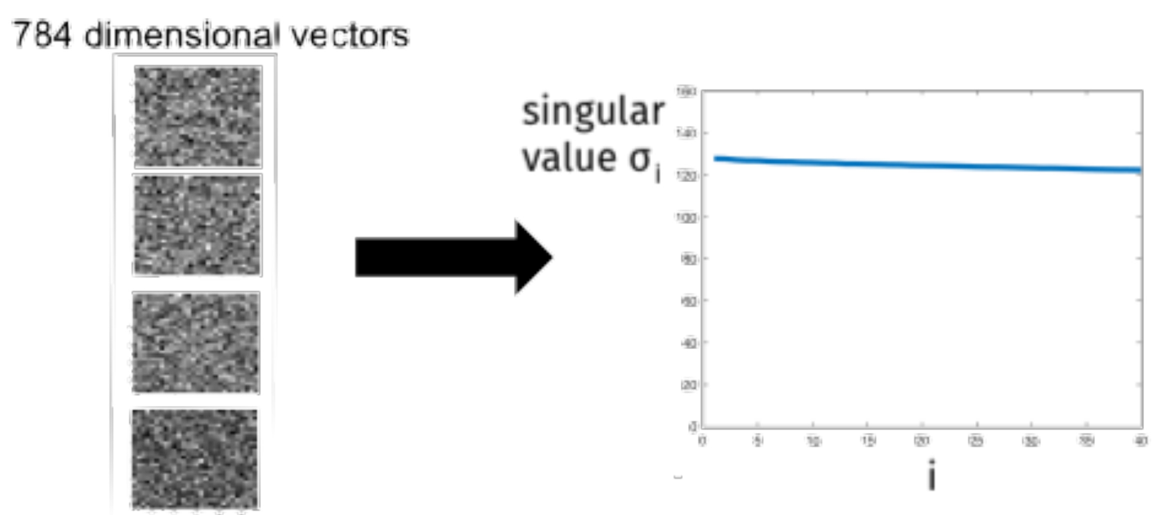
$$= \| \Sigma - \Sigma_k \|_F^2$$

$$= \left\| \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_d \end{bmatrix} - \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_k & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix} \right\|_F^2$$

$$= \sum_{i=k+1}^{d} \sigma_i^2$$



Structured



unstructured