Plan

Logistics

Review

Power Method

Lanczos Method

Problem Set

Proposal Feedback

Singular Value Decomposition

$X \in \mathbb{R}^{n \times d}$ rank k

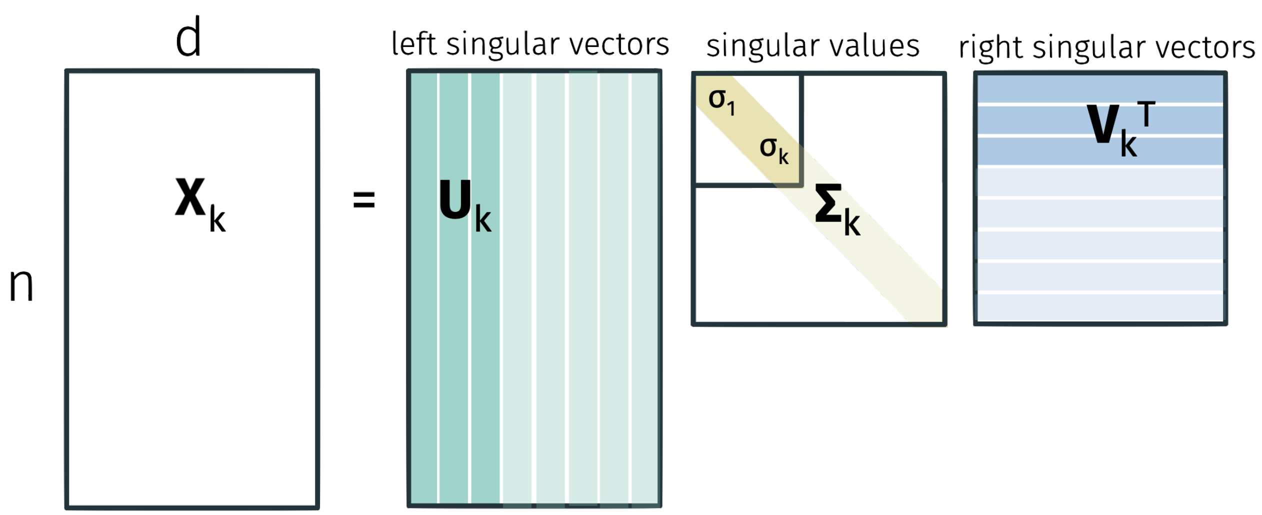
$k \neq d$ in general

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

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

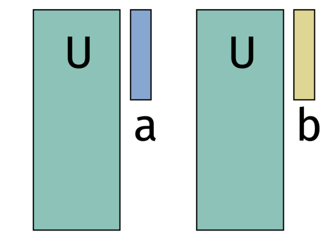
$U^T U = I$

$V^T V = I$

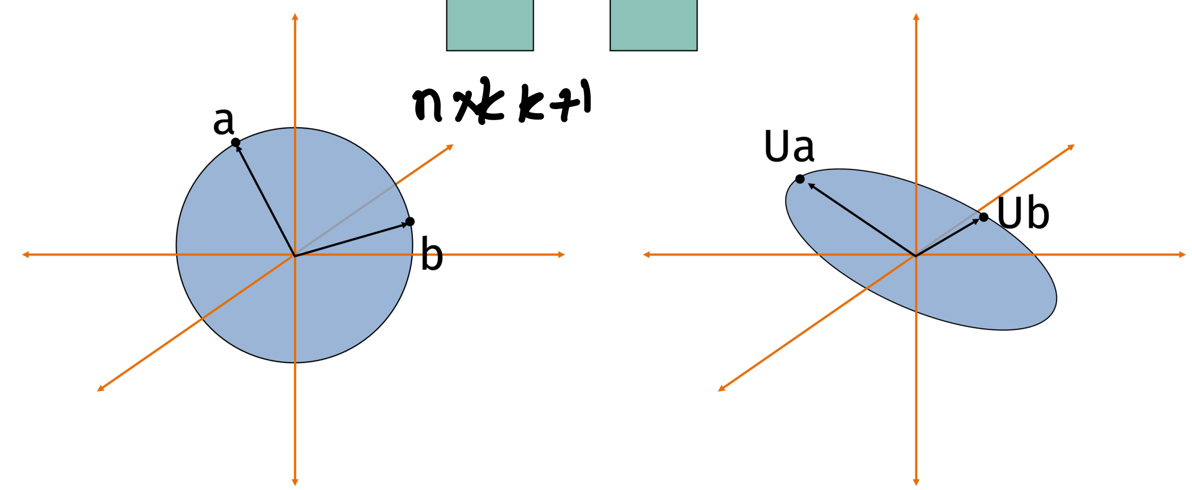


$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k$$

Multiplying by U



$n \times k$ $k \times 1$



$$Ua = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \dots & u_k \\ | & | & & | \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$$

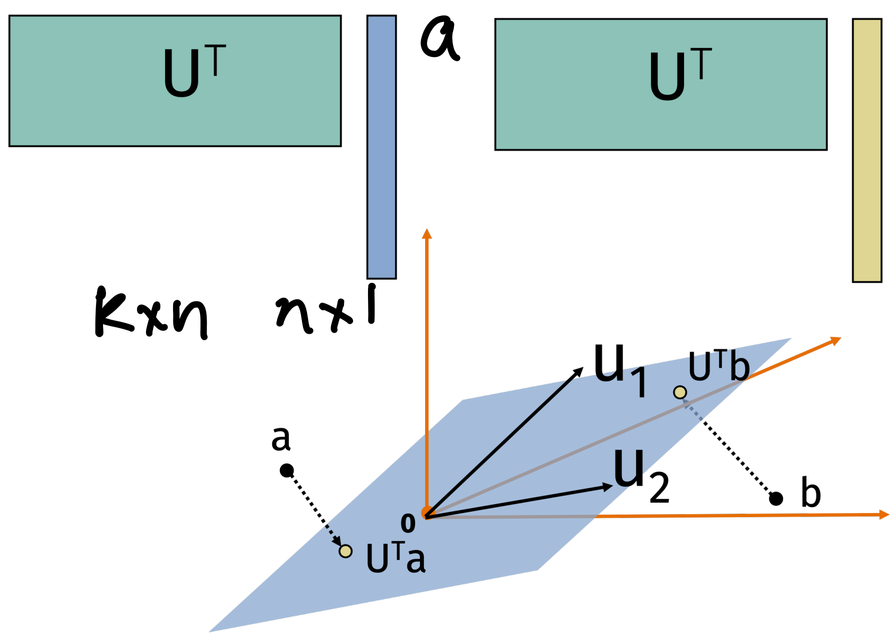
$$= u_1 \cdot a_1 + u_2 \cdot a_2 + \dots + u_k \cdot a_k$$

$$\|Ua\|_2^2 = \langle u_1 \cdot a_1 + \dots + u_k \cdot a_k, u_1 \cdot a_1 + \dots + u_k \cdot a_k \rangle$$

$$= 1 \cdot a_1^2 + 1 \cdot a_2^2 + \dots + 1 \cdot a_k^2$$

$$= \|a_k\|_2^2$$

Multiplying by U^T



$$U^T a = \begin{bmatrix} — & u_1 & — \\ — & u_2 & — \\ & \vdots & \\ — & u & — \end{bmatrix} \begin{bmatrix} | \\ a \\ | \end{bmatrix}$$

$$= \begin{bmatrix} \langle u_1, a \rangle \\ \langle u_2, a \rangle \\ \vdots \\ \langle u_n, a \rangle \end{bmatrix}$$

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

Tools

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

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

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

$$= \|U \Sigma V^T\|_F^2 = \|\Sigma V^T\|_F^2 \quad \text{by } 1$$
$$= \|V \Sigma^T\|_F^2 = \|\Sigma\|_F^2 \quad \text{by } 2, 1$$
$$= \sum_{i=1}^{k} \sigma_i^2$$

Finding SVD

$X = U \Sigma V^T$ $\quad$ $X \in \mathbb{R}^{n \times d}$

Compute $X^T X$ in ______

Decompose $X^T X = V \Lambda V^T$ in ______

Compute $L = XV = U \Sigma V^T V = U \Sigma$ in ______

$\sigma_i = \|L_{:i}\|_2$

$u_{:i} = L_i / \sigma_i$

Total: $O(nd^2)$

Faster?

Power Method

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

$$X^T X = V \Sigma^2 V^T \triangleq A$$

A has eigenvalues $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_d^2$

eigenvectors $v_1, v_2, \dots, v_d$

$$\underset{d\times d}{A} = \underset{d\times 1}{v_1} \underset{1\times d}{v_1^T} \cdot \sigma_1^2 + v_2 v_2^T \cdot \sigma_2^2 + \dots + v_d v_d^T \sigma_d^2$$

$$A v_i = \left(v_1 v_1^T \cdot \sigma_1^2 + v_2 v_2^T \cdot \sigma_2^2 + \dots + v_d v_d^T \sigma_d^2\right) v_i$$

$$= v_i \, v_i^T v_i \cdot \sigma_i^2 = v_i \, \sigma^2$$

$$A^q = \overbrace{(V \Lambda V^T) \cdots (V \Lambda V^T)}^{q}$$

$$= V \Lambda \underbrace{V^T V}_{I} \Lambda V^T \dots V \Lambda V^T$$

$$= V \Lambda^q V^T$$

Algorithm

$z^{(0)} \sim \mathcal{N}(0, I)$

$z^{(0)} = z^{(0)} / \|z\|_2$

For $i = 1, \dots, q$

$z^{(i)} = A z^{(i-1)}$

$n_i = \|z^{(i)}\|_2$

$z^{(i)} = z^{(i)} / n_i$

Return $z^{(q)}$

Time

Produces scaling of

$z^{(0)}, A z^{(0)}, A^2 z^{(0)}, \dots, A^q z^{(0)}$

Because eigenvectors span space $\swarrow$ some constant $c_i^{(0)} = \langle z^{(0)}, v_i \rangle$

$$z^{(0)} = c_1^{(0)} v_1 + c_2^{(0)} v_2 + \dots + c_d^{(0)} v_d$$

$$z^{(1)} = c_1^{(1)} v_1 + c_2^{(1)} v_2 + \dots + c_d^{(1)} v_d$$

$$\vdots$$

$$z^{(q)} = c_1^{(q)} v_1 + c_2^{(q)} v_2 + \dots + c_d^{(q)} v_d$$

$\uparrow$ want to show this is big so $z^{(q)} \approx v_1$

$$A z^{(0)} = \left(v_1 v_1^T \sigma_1^2 + v_2 v_2^T \sigma_2^2 + \ldots + v_d v_d^T \sigma_d^2 \right) \left(c_1^{(0)} v_1 + c_2^{(0)} v_2 + \ldots + c_d^{(0)} v_d \right)$$

$$= v_1 c_1^{(0)} \sigma_1^2 + v_2 c_2^{(0)} \sigma_2^2 + \ldots + v_d c_d^{(0)} \sigma_2^2$$

$$z^{(1)} = \frac{1}{n_1} A z^{(0)} = v_1 \frac{c_1^{(0)}}{n_1} \sigma_1^2 + v_2 \frac{c_2^{(0)} \sigma_2^2}{n_1} + \ldots + v_d \frac{c_d^{(0)} \sigma_2^2}{n_1}$$

$$c_i^{(j)} = \frac{1}{n_j} \sigma_i^2 \cdot c_i^{(j-1)}$$

$$c_i^{(q)} = \frac{1}{\prod_{l=1}^{q} n_l} \cdot \sigma_i^2 \, c_i^{(0)}$$

Analyze

$$\left| \frac{c_i^{(q)}}{c_1^{(q)}} \right| = \left(\frac{\sigma_i^2}{\sigma_1^2} \right)^q \left(\frac{c_i^{(0)}}{c_1^{(0)}} \right)$$

$\frac{c_i^{(0)}}{c_1^{(0)}} \leq d^3$ whp by Gaussian

$$\left(\frac{\sigma_i^2}{\sigma_1^2}\right)^q = \left(\frac{\sigma_i}{\sigma_1}\right)^{2q} = \left(\frac{\sigma_1 + \sigma_i - \sigma_1}{\sigma_1}\right)^{2q} = \left(1 - \frac{\sigma_1 - \sigma_i}{\sigma_1}\right)^{2q}$$

$$\frac{\sigma_1 - \sigma_2}{\sigma_1} = \gamma = \text{“spectral gap”}$$

$$q = \frac{\log\left(d^3 \sqrt{d/\epsilon}\right)}{2\gamma} \quad \text{then} \quad \left(1 - \frac{\sigma_1 - \sigma_i}{\sigma_1}\right)^{2q} \leq \frac{\sqrt{\epsilon/d}}{d^3}$$

$$\left|\frac{c_i^{(q)}}{c_1^{(q)}}\right| = \left(\frac{\sigma_i^2}{\sigma_1^2}\right)^q \left(\frac{c_i^{(0)}}{c_1^{(0)}}\right) \leq \sqrt{\epsilon/d}$$

$$\Rightarrow \left|c_i^{(q)}\right| \leq \sqrt{\epsilon/d} \quad \text{because } \left|c_1^{(q)}\right| \leq 1$$

$$\sum_{k=1}^{d} \left(c_k^{(q)}\right)^2 = 1, \quad \sum_{i \neq 1} \left(c_k^{(q)}\right)^2 \leq \sum_{i \neq 1} \epsilon/d < \epsilon$$

$$|c_1^{(q)}| \geq 1-\epsilon \qquad c_1^{(q)} = \langle v_1, z^{(q)} \rangle$$

$$\|v_1 - z^{(q)}\|_2^2 = 2 - 2\langle v_1, z^{(q)} \rangle$$

$$\langle v_1, z^{(q)} \rangle \geq 1-\epsilon \Rightarrow \|v_1 - z^{(q)}\|_2^2 \leq 2 - 2(1-\epsilon) = 2\epsilon$$

OR

$$\|v_1 + z^{(q)}\|_2^2 = 2 - 2\langle -v_1, z^{(q)} \rangle$$

$$\langle -v_1, z^{(q)} \rangle \geq 1-\epsilon \Rightarrow \|-v_1 - z^{(q)}\|_2^2 \leq 2 - 2(1-\epsilon) = 2\epsilon$$

We ran for $q = O\left(\frac{\log(d/\epsilon)}{\gamma}\right)$ steps

Keeping all iterates, we can improve to $O\left(\frac{\log(d/\epsilon)}{\sqrt{\gamma}}\right)$ steps

Lanczos implemented in python, MATLAB