Plan

Logistics

Review

Power Method

Michael's talk at noon in warner 101

↳ Nuclear physics and Brownian motio

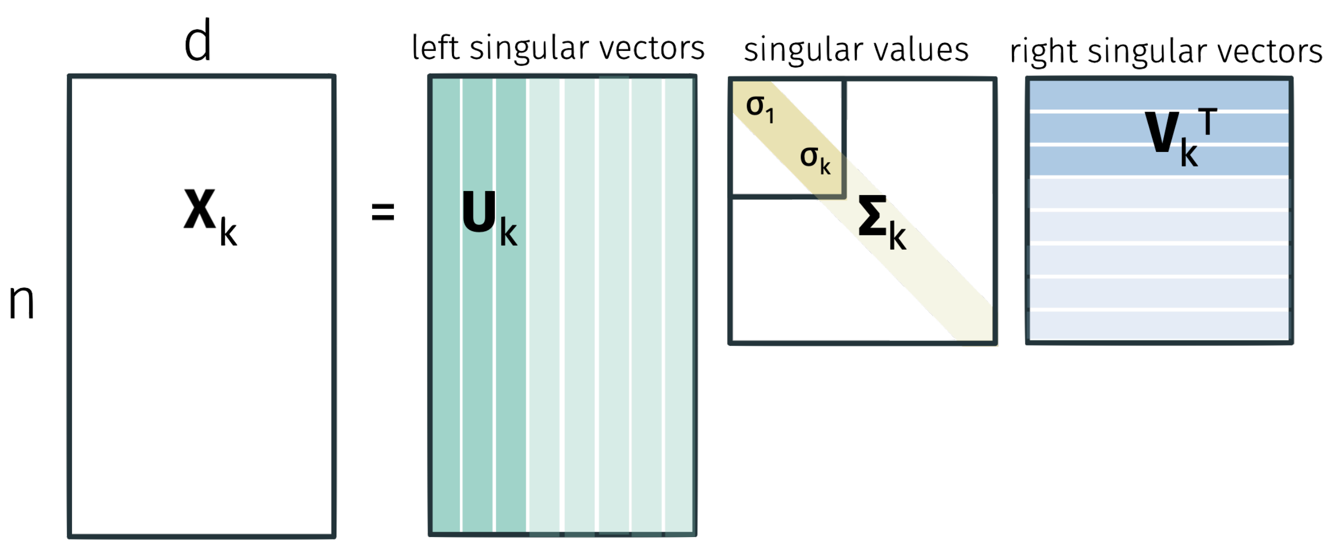
Singular Value Decomposition

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

$k \neq d$

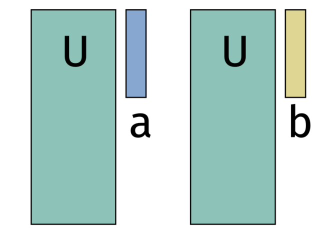
$$X = U \Sigma V^T \qquad U^T U = I$$

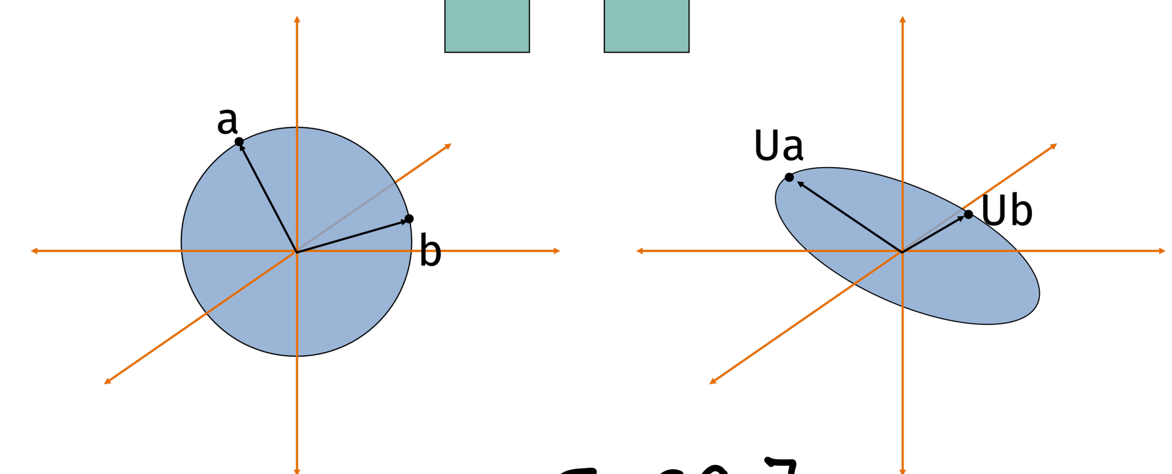
$n \times d \quad n \times k \quad k \times k \quad k \times d \qquad V^T V = I$



$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k \geq 0$$

Multiplying U





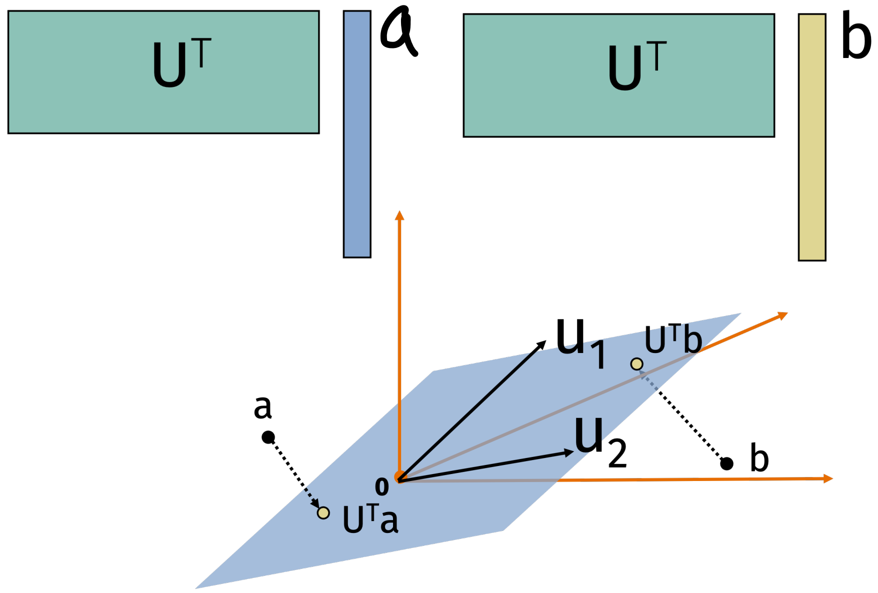
$$Ua = \begin{bmatrix} | & | & & | \\ \bar{u}_1 & \bar{u}_2 & & \bar{u}_k \\ | & | & & | \end{bmatrix} \underbrace{\begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}}_{a}$$

$$= \bar{u}_1 a_1 + \bar{u}_2 \cdot a_2 + \ldots + \bar{u}_k a_k$$

$$\|Ua\|_2^2 = \langle \bar{u}_1 a_1 + \ldots + \bar{u}_k a_k, \bar{u}_1 a_1 + \ldots + \bar{u}_k a_k \rangle$$

$$= a_1^2 + a_2^2 + \ldots + a_k^2 = \|a\|_2^2$$

Multiplying by U^T, projecting



$$U^T a = \begin{bmatrix} — u_1 — \\ — u_2 — \\ \vdots \\ — u_k — \end{bmatrix} \underset{a}{\begin{bmatrix} \\ \\ \\ \end{bmatrix}}$$

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

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

Tools

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

② $\|X^T\|_F^2 = \|X\|_F^2$

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

$= \|U\Sigma V^T\|_F^2 = \|\Sigma V^T\|_F^2$ by ①

$= \|V\Sigma^T\|_F^2 = \|\Sigma\|_F^2$ by ① ②

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

Finding SVD

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

Time

Compute $\underset{d \times d}{X^T X}$ $\quad O(d^2 n)$

Decompose $X^T X = V \Sigma^2 V^T = V \Lambda V^T$ $\quad O(d^3)$

Compute $L = \underset{n \times d}{X}\,\underset{d \times d}{V} = U \Sigma V^T V = U \Sigma$ $\quad O(nd \cdot d)$

$\sigma_i = \|L_i\|_2$

$u_i = L_i / \sigma_i$ $\quad O(nd^2)$

Power Method

$X = U \Sigma V^T$

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

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

eigenvectors $v_1, v_2, \ldots, v_d$

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

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

$$= \sigma_1^2 v_1 v_1^T \cdot v_i + \sigma_2^2 v_2 v_2^T \cdot v_i + \ldots + \sigma_d^2 v_d v_d^T v_i$$

$$= \sigma_i^2 v_i$$

$$A^q = \underbrace{A A A \cdots A}_{q}$$

$$= V \Sigma^2 \underbrace{V^T V}_{I} \Sigma^2 V^T \cdots V \Sigma^2 V^T$$

$$= V \underbrace{\Sigma^2 \Sigma^2 \Sigma^2}_{q} V^T$$

$$= V \Sigma^{2q} V^T$$

Algorithm

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

Time

$z^{(0)}, \; Az^{(0)}, \; A^2 z^{(0)}, \ldots$

$A^q z^{(0)}$

$z^{(0)} \sim N(0, I)$ $\quad z^{(0)} \in \mathbb{R}^d$ $\quad O(d)$

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

For $i = 1, \ldots, q$ $\qquad A' = A - v_1 v_1^T \sigma_1^2$

$\quad z^{(i)} = A z^{(i-1)}$ $\quad O(d^2)$

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

$\quad z^{(i)} = z^{(i)} / n_i$ $\quad O(d)$

Return $z^{(q)}$

$O(d^2 \cdot q)$

Because $v_1, \dots, v_d$

$c_i^{(0)} = \langle z^{(0)}, v_i \rangle$

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

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

$$\vdots$$

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

Goal: show this is large

$$z^{(q)} \approx v_1 \qquad \|v_1 - z^{(q)}\|_2^2 \leq \epsilon$$

$$\text{or} \quad \|-v_1 - z^{(q)}\|_2^2 \leq \epsilon$$

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

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

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

$$z^{(q)} = \left(v_1 \sigma_1^{2q} c_1^{(0)} + v_2 \sigma_2^{2q} c_2^{(0)} + \ldots + v_d \sigma_d^{2q} c_d^{(0)}\right) \frac{1}{\prod_{l=1}^{q} n_l}$$

$$c_i^{(t)} = \frac{\sigma_i^2}{n_t} c_i^{(t-1)} = \frac{\sigma_i^2}{n_t} \cdot \frac{\sigma_i^2}{n_{t-1}} c_i^{(t-2)} = \frac{\sigma_i^{2t}}{\prod_{l=1}^{t} n_l} c_i^{(0)}$$

$$\left|\frac{c_i^{(q)}}{c_1^{(q)}}\right| = \left|\frac{\sigma_i^{2q} c_i^{(0)}}{\prod_{l=1}^{q} n_l} \cdot \frac{\prod_{l=1}^{q} n_l}{\sigma_1^{2q} c_1^{(0)}}\right| = \left|\frac{\sigma_i^{2q}}{\sigma_1^{2q}}\right| \left|\frac{c_1^{(0)}}{c_1^{(0)}}\right|$$

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

$$\gamma = \frac{\sigma_1 - \sigma_2}{\sigma_1} \quad \text{"spectral gap"}$$

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

$$\left|\frac{c_i^{(q)}}{c_1^{(q)}}\right| = \left|\left(\frac{\sigma_i}{\sigma_1}\right)^{2q}\left(\frac{c_i^{(0)}}{c_1^{(0)}}\right)\right| \leq \frac{\sqrt{\epsilon/d}}{d^3}\, d^3 = \sqrt{\epsilon/d}$$

$$|c_1^{(q)}| \leq 1 \quad \Longrightarrow \quad |c_i^{(q)}| \leq |c_i^{(q)}|\sqrt{\epsilon/d} \leq \sqrt{\epsilon/d}$$

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

$$(c_1^{(q)})^2 = 1 - \sum_{k=2}^{d} (c_k^{(q)})^2 \geq 1 - \epsilon$$

$$c_1^{(q)} = \langle v_1, z^{(q)} \rangle$$

constant

$$\Rightarrow \quad |c_1^{(q)}| \geq (1-\epsilon)^2 \approx (1-\epsilon)$$

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

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

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

Using Lanczos, we run for $q = O\left(\frac{\log(d/\epsilon)}{\sqrt{\gamma}}\right)$ steps