FoDA L20

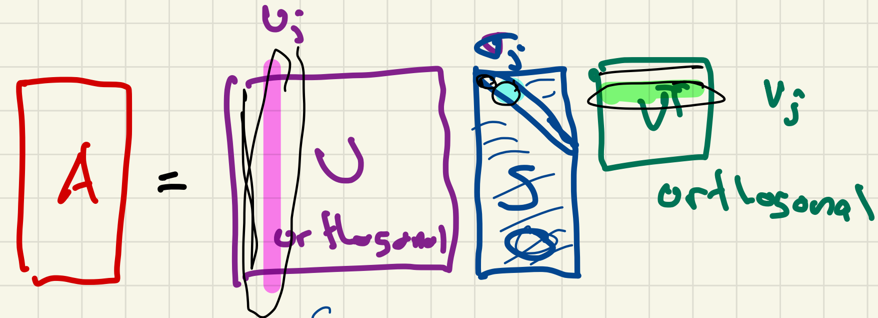
Rank-k Matrix Approximation
Eigen-decomposition & Power Method

Nov 3, 2022

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

$U, S, V^T = \mathrm{svd}(A)$

$A = USV^T$



$S_{11} = \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$

Goal: $A' \in \mathbb{R}^{n \times d}$ $\quad \mathrm{rank}(A') = k \quad k < d < n$

$$A_k = \underset{\mathrm{rank}\,k(A')}{\arg\min} \|A - A'\|_{F \text{ or } 2}$$

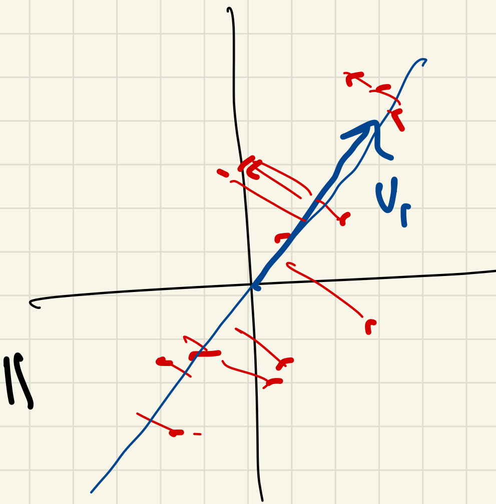
$$A_k = \sum_{j=1}^{k} \sigma_j \boxed{u_j v_j^T} \in \mathbb{R}^{n \times d}$$

$$A = \sum_{j=1}^{d} \sigma_j u_j v_j^T$$

$$A - A_k = \sum_{j=k+1}^{d} \sigma_j u_j v_j^T$$

v_1 = first right sing. vector

$$= \underset{\substack{v \in \mathbb{R}^d \\ \|v\| = 1}}{\arg\max} \underbrace{\|Av\|^2}_{= \sum_{i=1}^{n} \langle a_i, v \rangle^2}$$



$$\boxed{\sigma_1 = \|Av_1\|} = \|A\|_2 = \max_{\substack{v \\ \|v\|=1}} \|Av\|$$

$$\sigma_j^2 = \|Av_j\|^2 = \sum_{i=1}^{n} \langle a_i, v_j \rangle^2$$

v_j maximizes $\|Av_j\|^2$ s.t. $\langle v_j, v_i \rangle = 0 \quad i < j$

v_2 = unit vector that maximizes $\|Av_2\|^2$ subject to $\langle v_2, v_1 \rangle = 0$

$$\|A - A_k\|_2 = \Big\|\sum_{j=k+1}^{d} \sigma_j u_j v_j^T\Big\|_2 = \sigma_{k+1} = \|A v_{k+1}\|$$

$$\|A\|_F^2 = \sum_{j=1}^{d} \sigma_j^2 = \boxed{\sum_{i=1}^{n} \sum_{j=1}^{d} A_{ij}^2} = \sum_{i=1}^{n} \|a_i\|^2$$

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

$$\|A - A_k\|_F^2 = \sum_{j=k+1}^{d} \sigma_j^2$$

$$\|A - A_k\|_F^2 = \Big\|\sum_{j=k+1}^{d} \sigma_j u_j v_j^T\Big\|_F^2 = \sum_{j=k+1}^{d} \sigma_j^2 \underbrace{\|u_j v_j^T\|_F}_{1}$$

$$A_k = U_k \, S_k \, V_k^T$$

$U_k \in \mathbb{R}^{n \times k}$ first k L S Vec

$S_k \in \mathbb{R}^{k \times k}$ first k S Val

$V_k^T \in \mathbb{R}^{k \times d}$ first k R S Vec

$$A - A_k = \sum_{j=k+1}^{d} \sigma_j u_j v_j^T$$

square matrix $M \in \mathbb{R}^{d \times d}$

p.d.

eigenvalues $\lambda_j \in \mathbb{R}$

$M v_j = \lambda_j v_j$ $\quad v_j$ unit vector

↳ eigen vector

eigen decomposition $M = V L V^T$

↳ set of $\{(v_1, \lambda_1) \dots (v_d, \lambda_d)\}$

$\lambda_j \geq \lambda_{j+1}$ $\quad \langle v_j, v_{j'} \rangle = 0$

$V = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_d \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{d \times d}$ $\quad L = \mathrm{diag}(\lambda_1 \dots \lambda_d)$

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

$V V^T = I$

$M_R = A^T A \in \mathbb{R}^{d\times d}$ $A \in \mathbb{R}^{n\times d}$ $n > d$

↳ p.d. if A is full rank

$A = USV^T \leftarrow svd(A)$

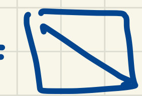
S

$I = V^{-1}V$

$M_R V = A^T A V = (V S^T U^T)(U S V^T) V$

I

$= V S^T S = V S^2$

$S^2 =$ 

$S^2 = diag(\sigma_1^2, \sigma_2^2, \ldots \sigma_d^2)$

for rsvec v_j of A

$M_R v_j = \sigma_j^2 v_j \leftarrow$ eigenvector

λ_j = eigenvalue

$$M_L = AA^T \in \mathbb{R}^{n \times n}$$

$$u_j M_L = \sigma_j^2 u_j$$

left sing vect u_j of A

↳ eigenvector u_j of AA^T

How to compute inverse M_R p.d $\in \mathbb{R}^{d \times d}$

$x^* = (A^T A)^{-1} A^T y$

$$M_R^{-1} = (V L V^{-1})^{-1} = V^{-1} L^{-1} V$$

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

$$L = \text{diag}(\lambda_1 \ldots \lambda_d)$$

$$L^{-1} = \text{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_d}\right)$$

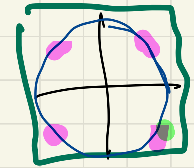
$$M_R^{-1} = V^T \ \text{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_d}\right) V$$

Power Method (to compute top eigenvector)

Input $M \in \mathbb{R}^{d \times d}$ # iterations $q \approx 10 - 20$

0. $u^{(0)} \in \mathbb{R}^d$, $\|u\| = 1$, at random

   Random guess of v_1



1. for $i = 1$ to q

   $u^{(i)} = M u^{(i-1)}$

2. Return $v = \dfrac{u^{(q)}}{\|u^{(q)}\|}$

for $i = 1$ to q

$u = M u$

q larger if $\dfrac{\lambda_1}{\lambda_2}$ is small.

$$M = \sum_{j=1}^{d} \lambda_j v_j v_j^T \qquad M' = L V^T$$

$$M' v = \sum_{j=1}^{d} \lambda_j v_j \langle v_j, v \rangle$$

$$v = M(M(M(\cdots(Mv))\cdots))$$

$$v = M^8 v$$

$$M^2 = (V L \cancel{V^T})(\cancel{V} L V^T)$$
$$= V L^2 V^T$$

$\lambda_1 = 10 \qquad \lambda_1^3 = 1000$

$\lambda_2 = 2 \qquad \lambda_2^3 = 8$

$$M^8 = V L^8 V^T$$

eigenvalues of M^8

$\lambda_1^8, \lambda_2^8 \ \dots \ \lambda_d^8$

Power Method A

0. $v \sim$ random unit

1. for $i = 1$ to g

$$v = A^T (A v)$$

2. $v_1 = v / \|v\|$