Homework #3

Due: 2024-11-17 23:59 | 7 Problems, 100 Pts

Name: XXX, ID: XXX

Problem 1 (10'). Suppose that the singular value decomposition of square matrix $\boldsymbol{A}$ is $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\top}$. Find out the maximum value of $\|\boldsymbol{A} - \boldsymbol{W}\|_F$ where $\boldsymbol{W}$ is an orthogonal matrix. ◀

Problem 2 (12'). Suppose $\boldsymbol{A} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\top}$ is the SVD of matrix $\boldsymbol{A}$ where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ and $r \geq 1$. Answer the following problems. You don't need to prove your result.

(1) (3') Determine whether there exists $\alpha \in \mathbb{R}$, such that there exists an absolute constant $c > 0$, for any $r \geq 1$, any matrix $\boldsymbol{A}$ and $k = 1, 2, \cdots, r$,

$$\min_{\mathrm{rank}(\boldsymbol{B}) \leq k} \frac{\|\boldsymbol{A} - \boldsymbol{B}\|_F}{\|\boldsymbol{A}\|_F} \leq ck^{\alpha}.$$

If so, write down the value of minimum α as well.

(2) (3') Determine whether there exists $\alpha \in \mathbb{R}$, such that there exists an absolute constant $c > 0$, for any $r \geq 1$, any matrix $\boldsymbol{A}$ and $k = 1, 2, \cdots, r$,

$$\min_{\mathrm{rank}(\boldsymbol{B}) \leq k} \frac{\|\boldsymbol{A} - \boldsymbol{B}\|_F}{\|\boldsymbol{A}\|_2} \leq ck^{\alpha}.$$

If so, write down the value of minimum α as well.

(3) (3') Determine whether there exists $\alpha \in \mathbb{R}$, such that there exists an absolute constant $c > 0$, for any $r \geq 1$, any matrix $\boldsymbol{A}$ and $k = 1, 2, \cdots, r$,

$$\min_{\mathrm{rank}(\boldsymbol{B}) \leq k} \frac{\|\boldsymbol{A} - \boldsymbol{B}\|_2}{\|\boldsymbol{A}\|_F} \leq ck^{\alpha}.$$

If so, write down the value of minimum α as well.

(4) (3') Determine whether there exists $\alpha \in \mathbb{R}$, such that there exists an absolute constant $c > 0$, for any $r \geq 1$, any matrix $\boldsymbol{A}$ and $k = 1, 2, \cdots, r$,

$$\min_{\mathrm{rank}(\boldsymbol{B}) \leq k} \frac{\|\boldsymbol{A} - \boldsymbol{B}\|_2}{\|\boldsymbol{A}\|_2} \leq ck^{\alpha}.$$

If so, write down the value of minimum α as well.

◀

Problem 3 (26'). Recall the Johnson-Lindenstrauss lemma we learned in class.

Theorem 1 (Johnson-Lindenstrauss Lemma). *Given $\epsilon \in (0, 1)$ and n vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n \in \mathbb{R}^m$. Pick a random matrix $\boldsymbol{\Pi} \in \mathbb{R}^{k \times m}$ as $\boldsymbol{\Pi} = \frac{1}{\sqrt{k}}\boldsymbol{W}$ where each entry $W_{i,j}$ ($1 \leq i \leq k, 1 \leq j \leq m$) is sampled independently from $\mathcal{N}(0, 1)$. Then there exists an absolute constant $c_1 > 0$, such that when $k \geq \frac{c_1 \ln n}{\epsilon^2}$, with probability at least $1 - \frac{3}{2n}$, for all $i, j \in \{1, 2, \cdots, n\}$,*

$$(1 - \epsilon)\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2 \leq \|\boldsymbol{\Pi}\boldsymbol{x}_i - \boldsymbol{\Pi}\boldsymbol{x}_j\|_2 \leq (1 + \epsilon)\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2.$$

In this problem, we are going to use Johnson-Lindenstrauss lemma to prove all big matrices are approximately low-rank.

(1) (2') Given $\epsilon \in (0,1)$ and n vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n \in \mathbb{R}^m$. Prove that, there exists an absolute constant $c_2 > 0$, such that when $k \geq \frac{c_2 \ln n}{\epsilon^2}$, there exists a matrix $\boldsymbol{\Pi} \in \mathbb{R}^{k \times m}$ satisfying

$$(1 - \epsilon)\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2^2 \leq \|\boldsymbol{\Pi}\boldsymbol{x}_i - \boldsymbol{\Pi}\boldsymbol{x}_j\|_2^2 \leq (1 + \epsilon)\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2^2$$

for all $i, j \in \{1, 2, \cdots, n\}$.

(2) (7') Given $\epsilon \in (0,1)$ and n vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n \in \mathbb{R}^m$. Prove that, there exists an absolute constant $c_3 > 0$, such that for $r \geq \frac{c_3 \ln(n+1)}{\epsilon^2}$, there exists a matrix $\boldsymbol{Q} \in \mathbb{R}^{r \times m}$ satisfying

$$|\boldsymbol{x}_i^\top \boldsymbol{x}_j - \boldsymbol{x}_i^\top \boldsymbol{Q}^\top \boldsymbol{Q}\boldsymbol{x}_j| \leq \epsilon(\|\boldsymbol{x}_i\|_2^2 + \|\boldsymbol{x}_j\|_2^2 - \boldsymbol{x}_i^\top \boldsymbol{x}_j)$$

for all $i, j \in \{1, 2, \cdots, n\}$.

[Hint: Consider $2\boldsymbol{x}_i^\top \boldsymbol{x}_j = \|\boldsymbol{x}_i\|_2^2 + \|\boldsymbol{x}_j\|_2^2 - \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2^2$.]

(3) (7') For any matrix $\boldsymbol{M}$, define

$$\|\boldsymbol{M}\|_{\max} = \max_{i,j} |M_{ij}|.$$

Given matrix $\boldsymbol{X} \in \mathbb{R}^{m \times n}$.

  (a) (2') Prove that, $\|\boldsymbol{X}\|_{\max} \leq \|\boldsymbol{X}\|_2$.

  (b) (5') Prove that, there exists matrix $\boldsymbol{U} = (\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_m) \in \mathbb{R}^{n \times m}, \boldsymbol{V} = (\boldsymbol{v}_1 \ \cdots \ \boldsymbol{v}_n) \in \mathbb{R}^{n \times n}$ such that $\boldsymbol{X} = \boldsymbol{U}^\top \boldsymbol{V}$ and $\|\boldsymbol{u}_i\|_2^2 \leq \|\boldsymbol{X}\|_2, \ \|\boldsymbol{v}_j\|_2^2 \leq \|\boldsymbol{X}\|_2$.

(4) (5') Suppose $\boldsymbol{X} \in \mathbb{R}^{m \times n}$ where $m \geq n$ and $\epsilon \in (0,1)$. Prove that, there exists an absolute constant $c_4 > 0$, such that with $r = \left\lceil \frac{c_4 \ln(m+n+1)}{\epsilon^2} \right\rceil$,

$$\min_{\text{rank}(\boldsymbol{Y}) \leq r} \|\boldsymbol{X} - \boldsymbol{Y}\|_{\max} \leq \epsilon \|\boldsymbol{X}\|_2.$$

(5) (5') A side note is that the result in problem (4) doesn't work for small matrix (small n). Consider $\boldsymbol{X} = \boldsymbol{I}_2$. Find out the value of

$$\min_{\text{rank}(\boldsymbol{Y}) \leq 1} \|\boldsymbol{X} - \boldsymbol{Y}\|_{\max}.$$

Prove your result.

◀

Problem 4 (22').

(1) (8') Given $\boldsymbol{S} = \mathrm{diag}(s_1, s_2, \cdots, s_n)$ where $s_1 \geq s_2 \geq \cdots \geq s_n \geq 0$. Find out the value of

$$\max_{\boldsymbol{W}^\top \boldsymbol{W} = \boldsymbol{I}_r} \|\boldsymbol{W}^\top \boldsymbol{S}\|_F$$

where $\boldsymbol{W} \in \mathbb{R}^{n \times r}$ $(n \geq r)$.

(2) (14') Suppose the singular values of a $m \times n$ $(m \geq n)$ matrix $\boldsymbol{A}$ are $\sigma_i(\boldsymbol{A})$, and $\sigma_1(\boldsymbol{A}) \geq \cdots \geq \sigma_n(\boldsymbol{A})$.

    (a) (12') Prove that, for any $k \in [n]$,

$$\sum_{i=1}^{k} \sigma_i(\boldsymbol{A}) = \max_{\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_k,\, \boldsymbol{V}^\top \boldsymbol{V} = \boldsymbol{I}_k} |\mathrm{Tr}(\boldsymbol{U}^\top \boldsymbol{A} \boldsymbol{V})|,$$

       where $\boldsymbol{I}_k$ is the rank-k identity matrix, $\boldsymbol{U}$ is a $m \times k$ matrix, and $\boldsymbol{V}$ is a $n \times k$ matrix.

    (b) (2') Prove that, for any $k \in [n]$ and $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m \times n}$ $(m \geq n)$, $\sum_{i=1}^{k} \sigma_i(\boldsymbol{A}+\boldsymbol{B}) \leq \sum_{i=1}^{k} \sigma_i(\boldsymbol{A}) + \sum_{i=1}^{k} \sigma_i(\boldsymbol{B})$.

◀

Problem 5 (5'). Use power method presented in section 3.7.1 of textbook to compute the largest 10 singular values of the random matrix $\boldsymbol{A}$ generated using NumPy as follows.

Algorithm 1: Generate the matrix
np.random.seed(20241025)   // Set random seed.
A = np.random.randn(2000, 1000)

Write down the singular values and submit the code as an attachment. You could use any functions provided by the NumPy package. If you use other programming languages or other packages to solve this problem, please generate the random matrix in a similar fashion. ◀

Problem 6 (5'). Consider a labeling $f : \{0,1\}^d \to \{-1, +1\}$: $f(x_1, x_2, x_3, \cdots, x_d) := (-1)^{\sum_{i=1}^{d} x_i}$. Is $\{0,1\}^d$ linearly separable if it is labeled by f? Prove your result. ◀

Problem 7 (20'). Consider function $K(\boldsymbol{x}, \boldsymbol{y}) = (1 + a\boldsymbol{x}^\top \boldsymbol{y})^2$ where $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $a \in \mathbb{R}$ is a constant.

(1) (6') Find out all possible $a \in \mathbb{R}$, such that $K(\boldsymbol{x}, \boldsymbol{y})$ is a kernel function. Prove your result.

(2) (14') Suppose $a \neq 0$ and $K(\boldsymbol{x}, \boldsymbol{y})$ is a kernel function, i.e., there exists $\boldsymbol{\varphi} : \mathbb{R}^n \to \mathbb{R}^m$ such that $K(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{\varphi}(\boldsymbol{x}), \boldsymbol{\varphi}(\boldsymbol{y}) \rangle$. Prove that, for any $\boldsymbol{\varphi}$ satisfying the conditions above, $m \geq \frac{n(n+3)}{2} + 1$.

[Hint: Consider matrix $M_{ij} = K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ $(1 \leq i, j \leq r)$, then $\boldsymbol{M}$ is a semi-definite matrix and $\boldsymbol{M} = \boldsymbol{G}^\top \boldsymbol{G}$ where $\boldsymbol{G} = (\boldsymbol{\varphi}(\boldsymbol{x}_1) \mid \cdots \mid \boldsymbol{\varphi}(\boldsymbol{x}_r))$. Take specific $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r \in \mathbb{R}^n$ and consider the rank of $\boldsymbol{M}$. You may find the basic results in linear algebra that $\mathrm{rk}(\boldsymbol{G}^\top \boldsymbol{G}) = \mathrm{rk}(\boldsymbol{G})$ useful.]

◀