CME 307 / MS&E 311 — Winter 2022

Optimization — Jan 5, 2022

Prof. Yinyu Ye — Optional - not graded

Homework Assignment 0

This is a diagnostic homework that covers prerequisite materials that you should be familiar with. This homework will not be graded and will not be counted towards the final grade.

Solve the following problems:

1. Consider the iterative process

$$x_{k+1} = \frac{1}{2}\left(x_k + \frac{a}{x_k}\right),$$

   where $a > 0$. Assuming the process converges, to what does it converge?

2. Let $\{(\mathbf{a}_i, c_i)\}_{i=1}^m$ be Wa given dataset where $\mathbf{a}_i \in R^n$, $c_i \in \{\pm 1\}$.

   (a) Compute the gradient of the following log-logistic-loss function,

$$f(\mathbf{x}, x_0) = \sum_{i:c_i=1} \log\left(1 + \exp(-\mathbf{a}_i^T\mathbf{x} - x_0)\right) + \sum_{i:c_i=-1} \log\left(1 + \exp(\mathbf{a}_i^T\mathbf{x} + x_0)\right),$$

   where $\mathbf{x} \in R^n$ and $x_0 \in R$.

   (b) Consider the following data set

$$\mathbf{a}_1 = (0;0), \quad \mathbf{a}_2 = (1;0), \quad \mathbf{a}_3 = (0;1), \quad \mathbf{a}_4 = (0;0), \quad \mathbf{a}_5 = (-1;0), \quad \mathbf{a}_6 = (0;-1),$$

   with label

$$c_1 = c_2 = c_3 = 1, \quad c_4 = c_5 = c_6 = -1,$$

   show that there is no solution for $\nabla f(\mathbf{x}, x_0) = 0$.

3. Given a symmetric matrix $A \in R^{n\times n}$ s.t. A has eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, show that for every $k = 1, 2, \cdots, n$, we have:

$$\lambda_k = \max_U \left\{ \min_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \middle| \mathbf{x} \in U, \mathbf{x} \neq \mathbf{0} \right\} \middle| U \text{ is a linear subspace of } R^n \text{ of dimension } k \right\} \quad (1)$$

$$= \min_U \left\{ \max_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \middle| \mathbf{x} \in U, \mathbf{x} \neq \mathbf{0} \right\} \middle| U \text{ is a linear subspace of } R^n \text{ of dimension } n-k+1 \right\} \quad (2)$$

4. Given symmetric matrices $A, B, C \in R^{n\times n}$ s.t. A has eigenvalues $a_1 \geq a_2 \geq \cdots \geq a_n$, B has eigenvalues $b_1 \geq b_2 \geq \cdots \geq b_n$ and C has eigenvalues $c_1 \geq c_2 \geq \cdots \geq c_n$, if $A = B + C$, show that for every $k = 1, 2, \cdots, n$, we have:

$$b_k + c_n \leq a_k \leq b_k + c_1. \tag{3}$$

5. Let $A \in R^{n\times n}$ be a positive-semidefinite matrix with Schur decomposition $A = Q\Lambda Q^T$, where $Q = [\mathbf{q}_1|\cdots|\mathbf{q}_n]$ is an orthogonal matrix, $\Lambda = \mathbf{diag}\{\lambda_1, \ldots, \lambda_n\}$ satisfies $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n \geq 0$. Show that for any $k = 1, \ldots, n$,

$$\min_{\mathbf{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \lambda_{k+1}, \tag{4}$$

and

$$\min_{\mathbf{rank}(B)=k} \|A - B\|_F = \|A - A_k\|_F = \sqrt{\sum_{j=k+1}^{n} \lambda_j^2}, \tag{5}$$

where A_k is defined as

$$A_k := \sum_{j=1}^{k} \lambda_j \mathbf{q}_j \mathbf{q}_j^T. \tag{6}$$

Here $\|\cdot\|_2$ stands for the spectrum (L_2) norm and $\|\cdot\|_F$ stands for the Frobenius norm.