Optimization Methods **Fall 2021**

Homework 3

Instructor: Lijun Zhang *Name:* Student name, *StudentId:* Student id

Notice

- The submission email is: **opt4grad@163.com**.
- Please use the provided LaTeX file as a template. If you are not familiar with LaTeX, you can also use Word to generate a **PDF** file.

Problem 1: Negative-entropy Regularization
Please show how to compute

$$\underset{x\in\Delta^n}{\operatorname{argmin}}\, b^\top x + c\cdot\sum_{i=1}^{n} x_i \ln x_i$$

where $\Delta^n = \{x | \sum_{i=1}^n x_i = 1, x_i \geq 0, i = 1, \cdots, n\}$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Problem 2: One inequality constraint
(a) With $c \neq 0$, express the dual problem of

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & f(x) \leq 0, \end{aligned}$$

in terms of the conjugate $f^\star$.
(b) Explain why the problem you give is convex. We do not assume f is convex.

Problem 3: KKT conditions
Consider the problem

$$\begin{aligned} & \min_{x\in\mathbb{R}^2} x_1^2 + x_2^2 \\ \text{s.t.} \quad & (x_1-1)^2 + (x_2-1)^2 \leq 2 \\ & (x_1-1)^2 + (x_2+1)^2 \leq 2 \end{aligned}$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in \mathbb{R}^2$.
a) Write the Lagrangian for this problem.
b) Does strong duality hold in this problem?
c) Write the KKT conditions for this optimization problem.

Problem 4: Matrix eigenvalues
We denote by $f(A)$ the sum of the largest r eigenvalues of a symmetric matrix $A \in \mathbb{S}^n$ (with $1 \leq r \leq n$), i.e.,

$$f(A) = \sum_{k=1}^{r} \lambda_k(A),$$

where $\lambda_1(A), \ldots, \lambda_n(A)$ are the eigenvalues of A sorted in decreasing order. Show that the optimal value of the optimization problem

$$\begin{aligned} \max \quad & \boldsymbol{tr}(AX) \\ \text{s.t.} \quad & \boldsymbol{tr}X = r \\ & 0 \preceq X \preceq I, \end{aligned}$$

with variable $X \in \mathbb{S}^n$, is equal to $f(A)$.

Problem 5: Dual Problem

Logistic Regression is a commonly used classification model in machine learning. Given the training data set $\{(a_1, b_1), \ldots, (a_n, b_n)\}$, where $a_i \in \mathbf{R}^d$, $b_i \in \{-1,\ 1\}$, the problem of Logistic Regression can be described as:

$$\min_{w \in \mathbf{R}^d} \sum_{i=1}^{n} \log\left(1 + e^{-b_i a_i^T w}\right) + \frac{\lambda}{2}\|w\|_2^2,$$

where $\lambda > 0$ is a constant. Derive the dual problem of this optimization problem.