IE 539 Convex Optimization Assignment 4

Fall 2024

Out: 18th November 2024

Due: 1st December 2024 at 11:59pm

Instructions

- Submit a PDF document with your solutions through the assignment portal on KLMS by the due date. Please ensure that your name and student ID are on the front page.
- Late assignments will be subject to a penalty. Special consideration should be applied for in this case.
- It is **required** that you typeset your solutions in LaTeX. Handwritten solutions will not be accepted.
- Spend some time ensuring your arguments are **coherent** and your solutions **clearly** communicate your ideas.

Question:	1	2	3	4	5	6	7	Total
Points:	10	10	10	20	20	20	10	100

1. (10 points) Let $h : \mathbb{R}^d \to \mathbb{R}$ be a closed convex function. Show that for any $x, y \in \mathbb{R}^d$ and $\eta > 0$,

$$\|\mathrm{prox}_{\eta h}(x) - \mathrm{prox}_{\eta h}(y)\|_2 \leq \|x - y\|_2.$$

2. (10 points) Let $h : \mathbb{R}^d \to \mathbb{R}$ be a closed convex function. Show that $x^* \in \arg\min_{x \in \mathbb{R}^d} h(x)$ if and only if $x^* = \mathrm{prox}_h(x^*)$.

3. (10 points) Let $f : \mathbb{R}^d \to \mathbb{R}$ be given by $f = g + h$ where $g : \mathbb{R}^d \to \mathbb{R}$ is a smooth convex function and $h : \mathbb{R}^d$ is closed and convex. Show that for any $\eta > 0$, $x^* \in \arg\min_{x \in \mathbb{R}^d} f(x)$ if and only if

$$x^* = (I + \eta \partial h)^{-1}(I - \eta \nabla g)(x^*).$$

4. In this question, we use Lagrangian duality to derive a solution to the following optimization problem.

$$\min_{x \in \Delta_d} \left\{ v^\top x + \sum_{i=1}^d x_i \log x_i \right\}$$

where $\Delta_d = \{x \in \mathbb{R}^d_+ : \sum_{i=1}^d x_i = 1\}$.

(a) (10 points) The Lagrangian function is defined as

$$\mathcal{L}(x, \lambda, \mu) = v^\top x + \sum_{i=1}^d x_i \log x_i - \lambda^\top x + \mu(1 - \sum_{i=1}^d x_i).$$

Then show that the associated Lagrangian dual function is given by

$$q(\lambda, \mu) = \mu - \sum_{i=1}^d e^{\lambda_i + \mu - v_i - 1}.$$

(b) (5 points) Let (λ^*, μ^*) be an optimal solution to the Lagrangian dual problem. Then show that $\lambda^* = 0$ and μ^* satisfies

$$e^{\mu^* - 1} = \frac{1}{\sum_{i=1}^d e^{-v_i}}.$$

(c) (5 points) Show that the optimal solution x^* satisfies

$$x_j^* = \frac{e^{-v_j}}{\sum_{i=1}^d e^{-v_i}} \quad \text{for } j = 1, \ldots, d.$$

5. Let $a \in \mathbb{R}^d$ be a vector such that $a_1 \geq a_2 \geq \ldots \geq a_d > 0$.

(a) (10 points) Consider the convex optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^d} \quad & -\log\left(\sum_{i \in [d]} a_i x_i\right) - \log\left(\sum_{i \in [d]} x_i / a_i\right) \\ \text{s.t.} \quad & x \geq 0, \ \sum_{i \in [d]} x_i = 1. \end{aligned}$$

Use the KKT conditions to show that the optimal solution occurs when $x_1 = x_d = 1/2$, $x_i = 0$ for $i = 2, \ldots, d-1$.

(b) (10 points) Let $A \in \mathbb{S}^d_{++}$ be a symmetric positive *definite* matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d > 0$. Use part (a) to show that, for any $u \in \mathbb{R}^d$ such that $\|u\|_2 = 1$,

$$2\sqrt{(u^\top A u) \cdot (u^\top A^{-1} u)} \leq \sqrt{\frac{\lambda_1}{\lambda_d}} + \sqrt{\frac{\lambda_d}{\lambda_1}}.$$

6. In this question we prove the convergence of the primal-dual subgradient method for saddle point problems. Let $\phi : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$ be a function such that $\phi(x, y)$ for any fixed $y \in \mathbb{R}^m$ is convex in x and $\phi(x, y)$ for any fixed $x \in \mathbb{R}^d$ is concave in x. Recall that the primal-dual subgradient method proceeds as follows.

- Choose $x_1 \in X$ and $y_1 \in Y$.

- For $t = 1, 2, 3, \ldots, T-1$:
  - Select $g_{x,t} \in \partial_x \phi(x_t, y_t)$, $g_{y,t} \in \partial_y \phi(x_t, y_t)$, and step size $\eta_t > 0$.
  - Compute $x_{t+1} = \mathrm{proj}_X\{x_t - \eta_t g_{x,t}\}$ and $y_{t+1} = \mathrm{proj}_Y\{y_t + \eta_t g_{y,t}\}$.

Assume that X and Y are convex.

(a) (5 points) Show that for any $(\bar{x}, \bar{y}) \in X \times Y$, $g_x \in \partial_x \phi(\bar{x}, \bar{y})$, and $g_y \in \partial_y \phi(\bar{x}, \bar{y})$,

$$\phi(\bar{x}, y) - \phi(x, \bar{y}) \leq -g_x^\top (x - \bar{x}) + g_y^\top (y - \bar{y}) \quad \forall (x, y) \in X \times Y.$$

(b) (15 points) Let $\bar{x}_T$ and $\bar{y}_T$ be defined as

$$\bar{x}_T = \left(\sum_{t=1}^T \eta_t\right)^{-1} \sum_{t=1}^T \eta_t x_t, \quad \bar{y}_T = \left(\sum_{t=1}^T \eta_t\right)^{-1} \sum_{t=1}^T \eta_t y_t.$$

Show that for any $(x, y) \in X \times Y$,

$$\phi(\bar{x}_T, y) - \phi(x, \bar{y}_T) \leq \frac{1}{2\sum_{t=1}^T \eta_t} \left(\|(x_1, y_1) - (x, y)\|_2^2 + \sum_{t=1}^T \eta_t^2 \|(g_{x,t}, g_{y,t})\|_2^2 \right).$$

7. (10 points) Let $f : \mathbb{R}^d \to \mathbb{R}$ be a closed convex function. Using the fact that

$$x = \mathrm{prox}_f(x) + \mathrm{prox}_{f^*}(x),$$

show that for any $\lambda > 0$,

$$x = \mathrm{prox}_{\lambda f}(x) + \lambda \mathrm{prox}_{(1/\lambda) f^*}(x/\lambda).$$