IERG 6120: Midterm Examination

November 9th, 2023

Due: 12:00PM (noon), November 10th, 2023 (at my office SHB 708).

- This is a take-home exam. **You must solve the problems independently. Do not discuss the problems with other students.**
- You can consult any textbooks/papers. However, if you use materials from textbooks other than the ones we use in class, you need to cite them. You also need to cite all papers that you use.
- You will need to turn in the exam paper by 12:00pm (noon), Friday, November 10th, 2023 in my office (SHB 708). If you would like to turn it in earlier, you can slip your exam paper under my office door.
- **Write your name and email at the space provided below.**
- There are **six** problems in the exam. The total points are 100.
- Email the instructor at xjlin@ie.cuhk.edu.hk if there are any questions.

____

Your Name

____

Email Address

(1) **(15 points)** (Yes or No) Is each of the following sets a convex set? **No justification is necessary.**

(a) (3 points) Let $y = [y_1, y_2, y_3]$ be a given point in $\mathbf{R}^3$. The set of points $x = [x_1, x_2, x_3] \in \mathbf{R}^3$ such that

$$d([x_1, x_2, x_3], [y_1, y_2, y_3]) \geq x_3,$$

where $d(x, y)$ is the Euclidean distance between x and y.

(b) (3 points) For any symmetric matrix $\mathbf{X} \in \mathbf{R}^{n \times n}$ and positive number a, we say that $\mathbf{X} \succ a$ if the matrix $\mathbf{X} - a\mathbf{I}$ (where $\mathbf{I}$ is the identity matrix) is positive semi-definite. (Recall that a matrix $\mathbf{X}$ is positive semi-definite if $v^T \mathbf{X} v \geq 0$ for all vector $v \in \mathbf{R}^n$). The set of symmetric matrices $\mathbf{X}$ such that $\mathbf{X} \succ 1$.

(c) (3 points) The set of points $x = [x_1, x_2] \in \mathbf{R}^2$ such that $x_1 x_2^2 \leq 1, x_1 \geq 0$ and $x_2 \geq 0$.

(d) (3 points) Consider two wireless links. Let $P_1, P_2 \geq 0$ be the transmission power of the first and second link, respectively. The SINR (signal-to-interference-and-noise ratio) of the first link is then given by $P_1/(P_2 + N)$, where N denote the background noise power. The set of (P_1, P_2) such that the SINR of the first link is higher than 2.

(e) (3 points) Let $\vec{p} = [p_i, i = 1, ..., n]$ denote a discrete probability distribution on $\{1, ..., n\}$, i.e., $p_i \geq 0$ for all i and $\sum_{i=1}^{n} p_i = 1$. Let $\vec{q} = [q_i, i = 1, ..., n]$ denote another discrete probability distribution. The KL distance between these two distributions is defined as

$$D_{KL}(\vec{p}||\vec{q}) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}.$$

The set of $(\vec{p}, \vec{q})$ such that $D_{KL}(\vec{p}||\vec{q}) \leq 0.5$.

(2) **(10 points)** Let A and B be two sets in $\mathbf{R}^n$. Define $A - B$ as the set of all points $x - y$ such that $x \in A$ and $y \in B$.

Prove the following statement: If A and B are convex, then $A - B$ must also be convex. Show all intermediate steps to get full credits.

(3) **(15 points)** (Yes or No) Is each of the following functions a convex function? **No justification is necessary.**

(a) (3 points) Let $(x_i, y_i), i = 1, ..., N$ be a given sequence of training samples, where each $x_i \in \mathbf{R}^2$ and each y_i is either 1 or -1. Let $w \in \mathbf{R}^2$ be a weight to be determined. The (loss) function $L(w)$ of the weight w, given by:

$$L(w) = \sum_{i:y_i=1} \log\left(\frac{e^{w^T x_i}}{1+e^{w^T x_i}}\right) + \sum_{i:y_i=-1} \log\left(\frac{1}{1+e^{w^T x_i}}\right)$$

(b) (3 points) Let $x, y \in \mathbf{R}$. The function

$$f(x, y) = \log(y^2/x + 1)$$

on the domain $\{(x, y)|x > 0, y \in \mathbf{R}\}$.

(c) (3 points) Let X be a given random variable in $\mathbf{R}$. Then, for any $\theta \in \mathbf{R}$, $M_X(\theta) = \mathbf{E}[e^{\theta X}]$ is known as the moment generating function of X. (Note that $M_X(\theta)$ is a function of θ.) Let $x \in \mathbf{R}$ be a real number. The function

$$I(x) = \sup_{\theta}[\theta x - \log M_X(\theta)].$$

(d) (3 points) Let $x, y \in \mathbf{R}$. The function

$$f(x, y) = -\sqrt{x(x + y)}$$

on the domain $\{(x, y)|x > 0, y > 0\}$.

(e) (3 points) Let $x \in \mathbf{R}$. The function $f(x) = e^{x^2}$.

(4) **(15 points)** *(Minimizing electricity cost with flexible load)*

Next generation electrical power systems will need to support both flexible load and inflexible load. Inflexible load is the electrical demand that must be met right away (e.g., when we turn on a lamp, the electricity generation must be immediately increased to power the lamp). On the other hand, flexible load is the demand that can be shifted in time. For example, suppose that we need to charge an electric vehicle (EV) before we leave for work at 8am tomorrow. We have the flexibility in delaying the charging time, as long as the total amount of energy charged is equal to the capacity of the EV battery. In this problem, we will study how to minimize the electricity generation cost by utilizing such flexibility.

In particular, consider the following simple model. Time is slotted with $t = 1, 2, ..., T$ (e.g., $T = 24$ hours). At each time-slot t, let x_t be the inflexible load that must be met right away. Further, there is a flexible load with total capacity C (e.g., the EV battery needs $C = 40$KWh of electricity). Let y_t be electricity consumed at time t that is used to fulfill the flexible load. Thus, we need $\sum_{t=1}^{T} y_t \geq C$. At each time-slot t, the cost of electricity is given by $g(x_t + y_t)$, where we assume that $g(\cdot)$ is a strictly convex and increasing function. Thus, we wish to solve the following optimization problem:

$$\begin{aligned} \min_{[y_t]} \quad & \sum_{t=1}^{T} g(x_t + y_t) && (1) \\ \text{subject to} \quad & y_t \geq 0, \text{ for all } t, \\ & \sum_{t=1}^{T} y_t \geq C. && (2) \end{aligned}$$

(a) (8 points) Associate a multiplier λ for the constraint (2). Write down the KKT condition for the optimal primal and dual solutions for problem (1).

(b) (7 points) Recall that the function $g(\cdot)$ is strictly convex, i.e., $g''(z) > 0$ for all z. Using the KKT condition from part (a), show that the optimal (primal) solution to (1) has the following "equal consumption" property: For any two time-slots t and s such that the flexible load is served with positive electricity consumption, i.e., $y_t > 0$ and $y_s > 0$, we must have

$$x_t + y_t = y_s + y_s.$$

In other words, the optimal fulfillment schedule $[y_t]$ is such that the total electrical consumption amount (of both flexible load and inflexible load) at each time-slot is as common as possible across all times.

Show all intermediate steps to get full credits.

(5) **(20 points)** *(An optimization problem with robust constraints)*

Start with the following optimization problem with respect to $x = [x_1, x_2] \in \mathbf{R}^2$:

$$\begin{array}{ll} \min\limits_{x} & f_0(x) \\ \text{subject to} & a^T x \geq b, \end{array}$$

where $a = [a_1, a_2] \in \mathbf{R}^2$ and $b \in \mathbf{R}$ are given. Assume that $f_0(x)$ is a convex function of x. Then clearly, this is a standard convex optimization problem.

In this question, however, we will consider a "robust" version of the above problem. That is, we will consider the setting where the vector $a \in \mathbf{R}^2$ is not fixed, but instead can be any vector inside a given set $\mathcal{P}$. Since we do not know the exact value of a, we wish to make sure that, with whatever choice of x that we select, the constraint $a^T x \geq b$ holds for **all** $a \in \mathcal{P}$. Assume that b is still fixed. Then, the optimization problem that we are interested in becomes:

$$\min_{x} \quad f_0(x) \tag{3}$$

$$\text{subject to} \quad \min_{a \in \mathcal{P}} a^T x \geq b. \tag{4}$$

As we can see, the problem (3) is "bi-level," because the constraint itself is described by another optimization problem! In general, this type of bi-level optimization problems can be difficult to solve. Fortunately, for this problem, we can use duality to convert it to a single-level optimization problem, which is also convex. You will be asked to perform this transformation below.

To be more specific, below you can assume that the constraint set $\mathcal{P}$ on $a = [a_1, a_2]$ is given by

$$\begin{aligned} 2a_1 + a_2 &\leq 4, \\ a_1 + 2a_2 &\leq 5, \end{aligned}$$

and b is given by $b = 1$.

(a) (10 points) For a given x that is already chosen, focus on the optimization problem in the constraint (4), i.e.:

$$\min_{a} \quad a^T x = a_1 x_1 + a_2 x_2 \tag{5}$$

$$\text{subject to} \quad 2a_1 + a_2 \leq 4, \tag{6}$$

$$a_1 + 2a_2 \leq 5. \tag{7}$$

Associate multipliers λ_1 and λ_2 for the two constraints (6) and (7), respectively. Derive the **dual** problem for this optimization problem (5). (*Hint:* Note that since x is assumed to be already chosen in (5), x_1 and x_2 can be treated as constants.)

(b) (10 points) Show that the bi-level optimization problem (3), with the value of $b = 1$ and $\mathcal{P}$ given above, is equivalent to:

$$\begin{aligned} \min_{x_1, x_2, \lambda_1, \lambda 2} \quad & f_0(x) \\ \text{subject to} \quad & -4\lambda_1 - 5\lambda_2 \geq 1 \\ & x_1 + 2\lambda_1 + \lambda_2 = 0 \\ & x_2 + \lambda_1 + 2\lambda_2 = 0 \\ & \lambda_1 \geq 0, \lambda_2 \geq 0, \end{aligned} \tag{8}$$

which is a single-level convex optimization problem. (*Hint:* Apply strong duality to the result of part (a). Then, use the fact that the following two constraints are equivalent: (i) $\max_z g(x, z) \geq b$; and (ii) there exists z such that $g(x, z) \geq b$.)

Show all intermediate steps for full credit. (*Note:* you do NOT need to develop the optimization algorithm for solving (8).)

(6) **(25 points)** *(Joint congestion control and power control)*

In this question, we will study a network model that jointly optimizes the decisions for end-user's packet injection rate (i.e., a congestion control problem) and the transmission power of the links (i.e., a power control problem).

Consider a network with L links, which is used to serve N flows. For each flow $s = 1, ..., N$, let x_s be its packet injection rate, and let $U_s(x_s)$ be its (concave and increasing) utility function. Each flow s will traverse a single path, given by a (fixed) subset of links. Specifically, let $H_s^l = 1$ if the packets of flow s traverses link l, and $H_s^l = 0$, otherwise. Thus, the total packet injection rate on link l is given by $\sum_{s=1}^N H_s^l x_s$.

Unlike the congestion control model that we studied in the class, however, here the capacity r_l of link l depends on its transmission power p_l. Specifically, the relationship between r_l and p_l is given by:

$$r_l = \log\left(1 + \frac{g_l p_l}{w_0}\right), \tag{9}$$

where w_0 is the (fixed) background noise, and g_l is the (fixed) channel propagation loss on link l. Note that we assume there is no interference among the links. Hence, the capacity r_l of a link l only depends on its own transmission power p_l, and is independent of the transmission power of other links.

Our goal is to maximize the total system utility, minus the total power consumption in the system. This problem can be formulated as follows:

$$\max_{[x_s],[p_l]} \quad \sum_{s=1}^N U_s(x_s) - \sum_{l=1}^L p_l \tag{10}$$

$$\text{subject to} \quad \sum_{s=1}^N H_s^l x_s \le r_l \text{ for all } l, \tag{11}$$

$$r_l = \log\left(1 + \frac{g_l p_l}{w_0}\right) \text{ for all } l, \tag{12}$$

$$p_l \ge 0 \text{ for all } l.$$

(a) (20 points) Problem (10) is not a convex problem yet. Convert (10) to an equivalent convex optimization problem. Then, associate a Lagrange multiplier λ_l to each constraint in (11). As in class, λ_l can be interpreted as the unit-price of using link l. Derive a **dual** gradient algorithm to solve problem (10).

(b) (5 points) Explain how the algorithm that you derived in part (a) is "distributed" and "price-driven," i.e., based on the prices, each flow s can determine its rate x_s, and each link l can determine its transmission power p_l, both as functions of the prices, independently of others. Further, each link can update its price in a way without knowledge of the utility function $U_s(\cdot)$ of each flow, or the rate-power function (9) of each link.

Show all intermediate steps for full credit. (*Note:* you do NOT need to prove the convergence of your algorithm.)