Optimization Methods for Data Science, Machine Learning and AI
Prof. Yinyu Ye

Homework Assignment 2

Individual Homework (110'):

1. (15') Consider problem 5 of Homework Assignment 1 where the second-order cone is replaced by the p-th order cone for $p \geq 1$:

$$\begin{array}{rl} \min_{\mathbf{x}} & 2x_1 + x_2 + x_3 \\ \text{s.t.} & x_1 + x_2 + x_3 = 1, \\ & x_1 - \|(x_2,\ x_3)\|_p \geq 0. \end{array}$$

   (a) (5') Write out the conic dual problem.

   (b) (5') Compute the dual optimal solution $(y^*, \mathbf{s}^*)$.

   (c) (5') Using the zero duality condition to compute the primal optimal solution $\mathbf{x}^*$.

2. (20') Consider the distributionally robust optimization (DRO) problem

$$\text{minimize}_{\mathbf{x}\in X} \quad \left[\max_{\mathbf{d}\in D} \quad \sum_{k=1}^{N} (\hat{p}_k + d_k) h(\mathbf{x}, \xi_k)\right] \tag{1}$$

   where the distribution set D is now given by

$$D = \{\mathbf{d} : \sum_{k=1}^{N} d_k = 0,\ \|\mathbf{d}\|^2 \leq 1/N,\ \hat{p}_k + d_k \geq 0,\ \forall k.\}$$

   (a) (3') What is the interpretation of D? Answer within 2 sentences.

   (b) (4') Represent D in standard conic form. (Hint: one set of the slack variables are in the second-order cone and the others are in the non-negative orthant cone.)

   (c) (7') Construct the conic dual of the inner max-problem.

   (d) (6') Replace the inner max-problem (1) by its dual, and simplify the DRO problem as much as possible.

3. (10') Consider the SOCP relaxation in problem 8 of Homework Assignment 1:

$$\begin{array}{rc} \min_{\mathbf{x}} & \mathbf{0}^T\mathbf{x} \\ \text{s.t.} & \|\mathbf{x} - \mathbf{a}_i\|^2 \leq d_i^2,\ i = 1, 2, 3, \end{array}$$

   where $\mathbf{x} \in R^2$.

(a) (4') Write down the first-order KKT optimality conditions.

(b) (3') Interpret (with no more than 2 sentences) the three optimal multipliers when the true position of the sensor is inside the convex hull of the three anchors.

(c) (3') Could the true position $\bar{\mathbf{x}} \in R^2$ of the sensor satisfy the optimality conditions if it is outside the convex hull of the three anchors? What would be the multiplier values?

4. (10') Consider the following parametric QCQP problem for a parameter $\kappa > 0$:

$$\begin{aligned} \min \quad & (x_1 - 1)^2 + x_2^2 \\ \text{s.t.} \quad & -x_1 + \tfrac{x_2^2}{\kappa} \geq 0 \end{aligned}$$

(a) (5') Is $\mathbf{x} = \mathbf{0}$ a first-order KKT solution?

(b) (5') Is $\mathbf{x} = \mathbf{0}$ a second-order KKT necessary or sufficient solution for some value of κ?

5. (20') (Central-Path and Potential) Given standard LP problem

$$\begin{aligned} \text{minimize}_{\mathbf{x} \in R^n} \quad & \mathbf{c}^T \mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{LP}$$

The **Analytic Center** of the primal feasible region $\mathcal{F}_p := \{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ is defined as the solution of the following linear-constrained convex optimization problem:

$$\begin{aligned} \text{minimize}_{\mathbf{x} \in R^n} \quad & -\sum_{j=1}^{n} \log x_j, \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} > \mathbf{0}. \end{aligned} \tag{PB}$$

The **Central Path** $\mathbf{x}(\mu)$ of (LP) is defined as the solution of the following Barrier LP problem (where $\mu > 0$ is a parameter):

$$\begin{aligned} \text{minimize}_{\mathbf{x} \in R^n} \quad & \mathbf{c}^T \mathbf{x} - \mu \cdot \sum_{j=1}^{n} \log x_j, \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} > \mathbf{0}. \end{aligned} \tag{BLP}$$

Part I Now consider the following example:

$$\begin{aligned} \text{minimize}_{\mathbf{x} \in R^3} \quad & x_1 + x_2, \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 1, \\ & (x_1, x_2, x_3) \geq \mathbf{0}. \end{aligned} \tag{2}$$

(a) (4') What is the analytic center of the primal feasible region in (2)?

(b) (4') Find the central path $\mathbf{x}(\mu) = (x_1(\mu), x_2(\mu), x_3(\mu))$ for (2).

(c) (4') Show that as μ decreases to 0, $\mathbf{x}(\mu)$ converges to the unique optimal solution of (2).

Part II Consider another example with different objective but the same feasible region:

$$\begin{aligned} \text{minimize}_{\mathbf{x} \in R^3} \quad & x_1 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 1 \\ & (x_1, x_2, x_3) \geq \mathbf{0} \end{aligned} \tag{3}$$

(d) (4') Find the central path $\mathbf{x}(\mu) = (x_1(\mu), x_2(\mu), x_3(\mu))$ for (3).

(e) (4') Which point does the central path converge to now (as $\mu \to 0+$)?

6. (15') Consider the following SVM problem, where $\mu \geq 0$ is a prescribed constant:

$$\begin{aligned} \min \quad & \beta + \mu \|\mathbf{x}\|^2 \\ \text{s.t.} \quad & a_i^T \mathbf{x} + x_0 + \beta \geq 1, \ \forall i, \\ & b_j^T \mathbf{x} + x_0 - \beta \leq -1, \ \forall j, \\ & \beta \geq 0. \end{aligned}$$

(a) (8') Write out the Lagrangian dual problem of the SVM problem. Write it as explicit as possible (at least remove the inner minimization). (Hint: You may want to consider two separate cases: $\mu = 0$ and $\mu > 0$)

(b) (7') Suppose that we have 6 training data in R^2: $a_1 = (0; 0)$, $a_2 = (1; 0)$, $a_3 = (0; 1)$ and $b_1 = (0; 0)$, $b_2 = (-1; 0)$, $b_3 = (0; -1)$. Use the optimality conditions (or any approach you want) to find optimal solutions for $\mu = 0$ and $\mu = 10^{-5}$, respectively. Are the two optimal solutions unique for the given μ? Prove your claim.

7. (20') Consider a generalized Arrow–Debreu equilibrium problem in which the market has n agents and m goods. Agent i, $i = 1, ..., n$, has a bundle amount of $\mathbf{w}_i = (w_{i1}, w_{i2}, \ldots, w_{im}) \in R^m_+$ goods initially and has a linear utility function whose coefficients are $\mathbf{u}_i = (u_{i1}, u_{i2}, \ldots, u_{im}) > 0 \in R^m$. The goal is to price each good so that the market clears. Note that, given the price vector $\mathbf{p} = (p_1, p_2, \ldots, p_m) > 0$, agent i's utility maximization problem is:

$$\begin{aligned} \text{maximize} \quad & \mathbf{u}_i^T \mathbf{x}_i \\ \text{subject to} \quad & \mathbf{p}^T \mathbf{x}_i \leq \mathbf{p}^T \mathbf{w}_i \\ & \mathbf{x}_i \geq 0 \end{aligned}$$

(a) (5') For a given $\mathbf{p} \in R^m$, write down the optimality conditions for agent i's utility maximization problem. Without loss of generality, you may fix $p_m = 1$ since the budget constraints are homogeneous in p.

(b) (5') Suppose that $\mathbf{p} \in R^m$ and $\mathbf{x}_i \in R^m$ satisfy the constraints:

$$\sum_{i=1}^{n} \mathbf{x}_i = \sum_{i=1}^{n} \mathbf{w}_i,$$
$$\frac{\mathbf{u}_i^T \mathbf{x}_i}{\mathbf{p}^T \mathbf{w}_i} p_j \geq u_{ij}, \quad \forall i, j,$$
$$\mathbf{p} \geq \mathbf{0},$$
$$\mathbf{x}_i \geq \mathbf{0}, \quad \forall i.$$

Show that $\mathbf{p}$ is then an equilibrium price vector.

(c) (5') For simplicity, assume that all u_{ij} are positive so that all p_j are positive. By introducing new variables $y_j = \log(p_j)$ for $j = 1, ..., m$, the conditions can be written as follows:

$$\begin{array}{ll} \min & 0 \\ \text{s.t.} & \sum_{i=1}^{n} \mathbf{x}_i = \sum_{i=1}^{n} \mathbf{w}_i \\ & \log(\mathbf{u}_i^T \mathbf{x}_i) - \log\left(\sum_{k=1}^{m} w_{ik} e^{y_k}\right) + y_j \geq \log(u_{ij}) \quad \forall i, j \\ & x_{ij} \geq 0, \qquad \forall i, j \end{array}$$

Show that this problem is convex in x_{ij} and y_j. (Hint: Use the fact that $\log\left(\sum_{k=1}^{m} w_{ik} e^{y_k}\right)$ is a convex function in the y_k's.)

(d) (5') Consider the Fisher example on Lecture Note with two agents and two goods, where the utility coefficients are given by

$$\mathbf{u}_1 = (2;\ 1) \quad \text{and} \quad \mathbf{u}_2 = (3;\ 1),$$

while now there are no fixed budgets. Rather, let

$$\mathbf{w}_1 = (1;\ 0) \quad \text{and} \quad \mathbf{w}_2 = (0;\ 1)$$

that is, agent 1 brings in one unit good x and agent brings in one unit of good y. Find the Arrow–Debreu equilibrium prices, where you may assume $p_y = 1$.

8. (Optional:) Consider the dual problem of an SDP,

$$\begin{array}{rl} \max_{\mathbf{y}, S} & by \\ \text{subject to} & Ay + S = C \\ & S \succeq 0, \end{array}$$

where $A, C \in \mathcal{S}^3$ is given. If A is not zero and the above problem is solvable, show that it has a solution $(\mathbf{y}, S)$ satisfies $\text{rank}(S) \leq 2$. (Hint: apply Caratheodory's theorem)

Groupwork (30') (group of 1-4 people):

9. (5') Let $\{(\mathbf{a}_i, c_i)\}_{i=1}^m$ be a given dataset where $\mathbf{a}_i \in R^n$, $c_i \in \{\pm 1\}$. In Logistic Regression (LR), we determine $x_0 \in R$ and $\mathbf{x} \in R^n$ by maximizing

$$\left(\prod_{i, c_i = 1} \frac{1}{1 + \exp(-\mathbf{a}_i^T \mathbf{x} - x_0)} \right) \left(\prod_{i, c_i = -1} \frac{1}{1 + \exp(\mathbf{a}_i^T \mathbf{x} + x_0)} \right).$$

which is equivalent to maximizing the log-likelihood probability

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

In this problem, we consider the quadratic regularized log-logistic-loss function

$$f(\mathbf{x}, x_0) = \sum_{i, c_i = 1} \log\left(1 + \exp(-\mathbf{a}_i^T \mathbf{x} - x_0)\right) + \sum_{i, c_i = -1} \log\left(1 + \exp(\mathbf{a}_i^T \mathbf{x} + x_0)\right) + 0.001 \cdot \|\mathbf{x}\|_2^2.$$

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

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

use the KKT conditions to find a solution of min $f(\mathbf{x}, x_0)$. You can either solve it numerically (e.g., using MATLAB `fsolve`) or analytically (represent the solution by a solution of a simpler (1D) nonlinear equation).

11. (15') Consider standard LP problem

$$\begin{aligned} \text{minimize}_{\mathbf{x} \in R^n} \quad & \mathbf{c}^T \mathbf{x}, \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{LP}$$

with its dual

$$\begin{aligned} \text{maximize}_{\mathbf{y} \in R^m, \mathbf{s} \in R^n} \quad & \mathbf{b}^T \mathbf{y}, \\ \text{subject to} \quad & A^T \mathbf{y} + \mathbf{s} = \mathbf{c}, \quad \mathbf{s} \geq \mathbf{0}. \end{aligned} \tag{LD}$$

For any $\mathbf{x} \in \text{int } \mathcal{F}_p := \{\mathbf{x} \in R^n : A\mathbf{x} = \mathbf{b}, \mathbf{x} > 0\}$ and $\mathbf{s} \in \text{int } \mathcal{F}_d := \{\mathbf{s} \in R^n : \mathbf{s} = \mathbf{c} - A^T \mathbf{y}, \mathbf{s} > \mathbf{0}, \mathbf{y} \in R^m\}$, the **Primal-Dual Potential Function** is defined by

$$\psi_{n+\rho}(\mathbf{x}, \mathbf{s}) := (n + \rho) \log(\mathbf{x}^T \mathbf{s}) - \sum_{j=1}^n \log(\mathbf{x}_j \mathbf{s}_j)$$

where $\rho > 0$ is a parameter.

Task: for two LP examples in Problem 5, namely (2) and (3), draw $\mathbf{x}$ part of the primal-dual potential function level sets

$$\psi_6(\mathbf{x}, \mathbf{s}) \leq 0 \quad \text{and} \quad \psi_6(\mathbf{x}, \mathbf{s}) \leq -10,$$

and

$$\psi_{12}(\mathbf{x}, \mathbf{s}) \leq 0 \quad \text{and} \quad \psi_{12}(\mathbf{x}, \mathbf{s}) \leq -10;$$

respectively in int $\mathcal{F}_p$ (on a plane).

Hint: To plot the $\mathbf{x}$ part of the level set of the potential function, say $\psi_6(\mathbf{x}, \mathbf{s}) \leq 0$, you plot

$$\{\mathbf{x} \in \text{int } \mathcal{F}_p : \min_{\mathbf{s} \in \text{int } \mathcal{F}_d} \psi_6(\mathbf{x}, \mathbf{s}) \leq 0\}.$$

This can be approximately done by sampling as follows. You randomly generate N primal points $\{\mathbf{x}^p\}_{p=1}^N$ from int $\mathcal{F}_p$, and N primal points of $\{\mathbf{s}^q\}_{q=1}^N$ from int $\mathcal{F}_d$. For each primal point $\mathbf{x}^p$, you find if it is true that

$$\min_{q=1,\ldots,N} \psi_6(\mathbf{x}^p, \mathbf{s}^q) \leq 0.$$

Then, you plot those $\mathbf{x}^p$ who give an "yes" answer.

10. (10') Recall the Fisher's Equilibrium prices problem (discussed in Lecture Note 6), which we describe here again for reference. Let B be the set of buyers and G be the set of goods. Each buyer $i \in B$ has a budget $w_i > 0$, and utility coefficients $u_{ij} \geq 0$ for each good $j \in G$. Under price $\mathbf{p}$, buyer $i \in B$'s optimal purchase quantity $\mathbf{x}_i^*(\mathbf{p})$ is the solution of the following optimization problem:

$$\begin{aligned} \mathbf{x}_i^*(\mathbf{p}) \in \arg\max \quad & \mathbf{u}_i^T \mathbf{x}_i := \sum_{j \in G} u_{ij} x_{ij} \\ \text{s.t.} \quad & \mathbf{p}^T \mathbf{x}_i := \sum_{j \in G} p_j x_{ij} \leq w_i, \\ & \mathbf{x}_i \geq 0 \end{aligned}$$

Suppose each good $j \in G$ has a supply level $\bar{s}_j$. We call a price vector $\mathbf{p}^*$ an **equilibrium price vector** if the market clears, namely for all $j \in G$,

$$\sum_{i \in B} x^*(\mathbf{p}^*)_{ij} = \bar{s}_j.$$

In the lecture, we discussed how to compute the equilibrium price $\mathbf{p}^*$ and buyers' activities $\{\mathbf{x}_i^*(\mathbf{p}^*)\}_{i \in B}$ under the equilibrium price based on utility coefficients $\{\mathbf{u}_i\}_{i \in B}$, budgets $\{w_i\}_{i \in B}$ and supplies $\bar{\mathbf{s}}$:

$$(\{\mathbf{u}_i\}_{i \in B}, \{w_i\}_{i \in B}, \bar{\mathbf{s}}) \Rightarrow (\mathbf{p}^*, \{\mathbf{x}_i^*(\mathbf{p}^*)\}_{i \in B}) \tag{4}$$

In this question, we consider the inverse problem of (4): suppose the market does not know the "private information" of each buyer, namely the utility $\{\mathbf{u}_i\}_{i\in B}$ and the budgets $\{w_i\}_{i\in B}$, but instead you observe the equilibrium prices $\{\mathbf{p}^{*(k)}\}_{k=1}^K$ and their corresponding realized activities $\{\mathbf{x}_i^{*(k)}\}_{k=1}^K$ under K different supply levels $\bar{\mathbf{s}}^{(1)}, \ldots, \bar{\mathbf{s}}^{(K)}$. The query is to infer buyers' utility coefficients $\{\mathbf{u}_i\}_{i\in B}$ and their budgets $\{w_i\}_{i\in B}$. We assume that the utility function is ℓ_1-normalized, namely $\|\mathbf{u}_i\|_1 = 1$ for $i \in B$.

Hint: Mathematically, the query is to find $\{\mathbf{u}_i\}_{i\in B}$ (s.t. $\mathbf{u}_i \geq \mathbf{0}$ and $\|\mathbf{u}_i\|_1 = 1$) and $\{w_i\}_{i\in B}$ (s.t. $w_i > 0$) such that for all $i \in B$, and $k = 1, \ldots, K$,

$$\begin{aligned} \mathbf{x}_i^{*(k)} = \arg\max_{\mathbf{x}_i} \quad & \mathbf{u}_i^T \mathbf{x}_i \\ \text{s.t.} \quad & (\mathbf{p}^{*(k)})^T \mathbf{x}_i \leq w_i \\ & \mathbf{x}_i \geq \mathbf{0} \end{aligned}$$

given $\{\mathbf{x}_i^{*(k)}\}_{i\in B, k\in\{1,\ldots,K\}}$ and $\{\mathbf{p}^{*(k)}\}_{k\in\{1,\ldots,K\}}$.

Question: Now consider the following 2-buyer 2-good example and solve this inverse problem. Let $B = \{1, 2\}$ and $G = \{1, 2\}$. Suppose we observe the following 5 scenarios:

- $\mathbf{p}^{*(1)} = (\frac{9}{5}; \frac{3}{5})$, $\mathbf{x}_1^{*(1)} = (1; \frac{1}{3})$, $\mathbf{x}_2^{*(1)} = (0; \frac{5}{3})$;
- $\mathbf{p}^{*(2)} = (2; 1)$, $\mathbf{x}_1^{*(2)} = (1; 0)$, $\mathbf{x}_2^{*(2)} = (0; 1)$;
- $\mathbf{p}^{*(3)} = (1; 1)$, $\mathbf{x}_1^{*(3)} = (2; 0)$, $\mathbf{x}_2^{*(3)} = (0; 1)$;
- $\mathbf{p}^{*(4)} = (\frac{1}{2}; 1)$, $\mathbf{x}_1^{*(4)} = (4; 0)$, $\mathbf{x}_2^{*(4)} = (0; 1)$;
- $\mathbf{p}^{*(5)} = (\frac{3}{7}; \frac{6}{7})$, $\mathbf{x}_1^{*(5)} = (\frac{14}{3}; 0)$, $\mathbf{x}_2^{*(5)} = (\frac{1}{3}; 1)$.

Use any approach to find $\{\mathbf{u}_i\}_{i\in B}$ (s.t. $\mathbf{u}_i \geq \mathbf{0}$ and $\|\mathbf{u}_i\|_1 = 1$) and $\{w_i\}_{i\in B}$ (s.t. $w_i > 0$). Describe your approach and report the result.