MGMT 690 - Pset 2

Spring 2024

Instructions:

- This pset is due on Sunday, April 7 at 11:59pm.
- Completed psets should be submitted to Gradescope.
- **Exercises** are for your own review only. They do not need to be submitted and will not be graded.
- **Complete all problems 1–5.**

1 Exercises

Excercise 1

Proposition 1. *Suppose $f_0, \ldots, f_k : \mathbb{R}^n \to \mathbb{R}$ are SOCR functions and $\mathcal{X}_1, \ldots, \mathcal{X}_m \subseteq \mathbb{R}^n$ are SOCR set. Then the following problem*

$$\inf_{x \in \mathbb{R}^n} \left\{ f_0(x) : \begin{array}{l} f_i(x) \leq 0,\ \forall i \in [k] \\ x \in \mathcal{X}_i,\ \forall i \in [m] \end{array} \right\}$$

has an SOC representation.

Proof. By assumption

$$\text{epi}(f_i) = \Pi_{(x,t)} \left\{ (x, t, u^{(i)}) :\ A^{(i)}(x, t, u^{(i)}) - a^{(i)} \in K^{(i)} \right\}$$

where $K^{(i)}$ is a product of second order cones. Similarly,

$$\mathcal{X}_j = \Pi_x \left\{ (x, v^{(j)}) :\ B^{(j)}(x, v^{(j)}) - b^{(j)} \in W^{(j)} \right\}$$

where $W^{(j)}$ is a product of second order cones.

Then,

$$\begin{aligned} &\inf_{x \in \mathbb{R}^n} \left\{ f_0(x) : \begin{array}{l} f_i(x) \leq 0,\ \forall i \in [k] \\ x \in \mathcal{X}_i,\ \forall i \in [m] \end{array} \right\} \\ &= \inf_{x, u^{(i)}, t^{(i)}, v^{(j)}} \left\{ t^{(0)} : \begin{array}{l} A^{(i)}(x, t^{(i)}, u^{(i)}) - a^{(i)} \in K^{(i)},\ \forall i = 0, \ldots, k \\ t^{(i)} \leq 0,\ \forall i = 1, \ldots, k \\ B^{(j)}(x, v^{(j)}) - b^{(j)} \in W^{(j)},\ \forall j \in [m] \end{array} \right\}. \blacksquare \end{aligned}$$

Exercise 2

Recall that $S_k : \mathbb{S}^n \to \mathbb{R}$ is the function where $S_k(X)$ is the sum of the k largest eigenvalues of X.

Proposition 2. *It holds that*

$$S_k(X) = \max_{Y \in \mathbb{S}^n} \left\{ \langle X, Y \rangle : \begin{array}{l} 0 \preceq Y \preceq I \\ \operatorname{tr}(Y) = k \end{array} \right\}.$$

In particular, $S_k(X)$ is convex.

Proof. Without loss of generality, X is diagonal (apply the spectral theorem). Without loss of generality, Y is diagonal: given any $Y \in \mathbb{S}^n$ that is feasible, the diagonal matrix $\operatorname{Diag}(\operatorname{diag}(Y))$ is also feasible with the same objective function.

We will thus consider the restricted maximization problem where $X = \operatorname{Diag}(\lambda_1, \dots, \lambda_n)$ and $Y = \operatorname{Diag}(y_1, \dots, y_n)$:

$$\max_{y \in \mathbb{R}^n} \left\{ \sum_i x_i y_i : \begin{array}{l} y_i \in [0,1],\ \forall i \in [n] \\ \sum_i y_i = k \end{array} \right\}.$$

The optimal solution to this maximization problem places one unit of weight on each i corresponding to the k largest entries of x_i. Thus, the optimum value of this problem is $S_k(X)$.

Note that for every fixed Y, $\langle X, Y \rangle$ is a convex (in fact linear) function of X. Thus, $S_k(X)$ is the maximum of a family of convex functions (parameterized by Y). We deduce that S_k is convex. ■

We derive the dual to this problem as follows. Let $Z, W \succeq 0$ and let $s \in \mathbb{R}$. Then,

$$\langle I - Y, Z \rangle + \langle Y, W \rangle - s \langle I, Y \rangle + sk \geq 0$$

Rearranging, we have that

$$sk + \operatorname{tr}(Z) \geq \langle Z, Y \rangle - \langle Y, W \rangle + \langle sI, Y \rangle .$$

Thus, the dual to the above problem is

$$\begin{aligned} &\inf_{Z,W,s} \left\{ sk + \operatorname{tr}(Z) : \begin{array}{l} Z - W + sI = X \\ W, Z \succeq 0 \end{array} \right\} \\ &\quad = \inf_{Z,s} \left\{ sk + \operatorname{tr}(Z) : \begin{array}{l} Z + sI \succeq X \\ Z \succeq 0 \end{array} \right\}. \end{aligned}$$

Problems

Problem 1

Proposition 3. *Strong duality fails for the following SDP:*

$$\inf_{X \in \mathbb{S}^2} \left\{ 2X_{1,2} : \begin{array}{l} X_{1,1} = 0 \\ X \succeq 0 \end{array} \right\}$$

Proof. Note that for a 2×2 matrix X, we have that $X \succeq 0$ if and only if both diagonal elements are nonnegative and the determinant is nonnegative. Thus,

$$\begin{aligned} &\inf_{X \in \mathbb{S}^2} \left\{ 2X_{1,2} : \begin{array}{l} X_{1,1} = 0 \\ X \succeq 0 \end{array} \right\} \\ &= \inf_{X \in \mathbb{S}^2} \left\{ 2X_{1,2} : \begin{array}{l} X_{1,1} = 0 \\ X_{2,2} \geq 0 \\ -X_{1,2}^2 \geq 0 \end{array} \right\} \\ &= 0. \end{aligned}$$

We construct the dual. Let $Y \in \mathbb{S}^2_+$ and let $t \in \mathbb{R}$. Then,

$$tX_{1,1} + \langle X, Y \rangle \geq 0$$

Thus, the dual is

$$\begin{aligned} &\sup_{Y \in \mathbb{S}^2,\, t \in \mathbb{R}} \left\{ 0 : \begin{array}{l} \begin{pmatrix} t & \\ & 0 \end{pmatrix} + Y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ Y \succeq 0 \end{array} \right\} \\ &= -\infty. \end{aligned}$$

∎

Problem 2

Fix a discrete metric space $\mathcal{X} = \{x_1, \ldots, x_n\}$. Let $D \in \mathbb{R}^{n \times n}$ denote the matrix where $D_{i,j}$ is the distance between x_i and x_j.

Let P be a probability distribution on $\mathcal{X}$ defined by $P = (p_1, \ldots, p_n)$. Similarly, let $Q = (q_1, \ldots, q_n)$ be a probability distribution on $\mathcal{X}$.

The *Wasserstein distance* between P and Q can be written as the optimal value of the following LP:

$$\inf_{\Gamma \in \mathbb{R}^{n \times n}} \left\{ \langle D, \Gamma \rangle : \begin{array}{l} \Gamma 1_n = p \\ \Gamma^\intercal 1_n = q \\ \Gamma \geq 0 \end{array} \right\}$$

We construct the dual to this LP. Let $y \in \mathbb{R}^n$, $z \in \mathbb{R}^n$ and $\Xi \in \mathbb{R}^{n \times n}_+$. Then,

$$\langle \Gamma, y1_n^\intercal + 1_n z^\intercal + \Xi \rangle \geq \langle p, y \rangle + \langle q, z \rangle .$$

The dual of this LP is thus given by

$$\begin{aligned}
&\sup_{y\in\mathbb{R}^n,\, z\in\mathbb{R}^n,\, \Xi\in\mathbb{R}^{n\times n}} \left\{ \langle p, y\rangle + \langle q, z\rangle : \begin{array}{l} y1_n^\intercal + 1_n z^\intercal + \Xi = D \\ \Xi \geq 0 \end{array} \right\} \\
&= \sup_{y\in\mathbb{R}^n,\, z\in\mathbb{R}^n} \left\{ \langle p, y\rangle + \langle q, z\rangle : \ y_i + z_j \leq D_{i,j},\ \forall i, j \right\}.
\end{aligned}$$

Both primal and dual problems are solvable because they are both feasible: Take $\Gamma = pq^\intercal$ in the primal, and take $y = z = 0$ in the dual. Complementary slackness means that for all i, j, either $\Gamma_{i,j} = 0$ or $y_i + z_j = D_{i,j}$.

One way to think of the dual is: it is a profit maximization problem for a third party. This third party buys dirt at price-per-unit y_i at location x_i and sells it at price-per-unit z_j at location x_j.

Problem 3

Proposition 4. *The set*

$$\left\{ (x, y) \in \mathbb{R}^2 : \begin{array}{l} xy \geq 1 \\ x, y \geq 0 \end{array} \right\}$$

is an SOCR set.

Proof. Suppose $x, y \geq 0$. Then

$$\begin{aligned}
& xy \geq 1 \\
\iff \quad & (x+y)^2 - (x-y)^2 \geq 4 \\
\iff \quad & (x+y)^2 \geq 4 + (x-y)^2 \\
\iff \quad & x + y \geq \sqrt{2^2 + (x-y)^2} \\
\iff \quad & \begin{pmatrix} x+y \\ x-y \\ 2 \end{pmatrix} \in \mathcal{L}^{1+2}.
\end{aligned}$$

Thus, the set in question is equal to

$$\left\{ (x, y) \in \mathbb{R}^2 : \begin{array}{c} x, y \geq 0 \\ \begin{pmatrix} x+y \\ x-y \\ 2 \end{pmatrix} \in \mathcal{L}^{1+2} \end{array} \right\}. \qquad \blacksquare$$

Problem 4

Consider an optimization problem of the form

$$\inf_{x\in\mathbb{R}^n} \left\{ f(x) : \ g_i(x) \leq 0,\ \forall i \in [m] \right\}.$$

We make no assumptions on whether f or $g_1, \ldots, g_m$ is convex. Define

$$\mathcal{I} := \left\{ \begin{pmatrix} f(x) \\ g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} : x \in \mathbb{R}^n \right\} + \mathbb{R}^{1+m}_+.$$

Proposition 5. *If f and $g_1, \ldots, g_m$ are convex functions, then $\mathcal{I}$ is a convex set.*

Proof. Suppose $(\phi, \gamma_1, \ldots, \gamma_m) \in \mathcal{I}$ and $(\hat{\phi}, \hat{\gamma}_1, \ldots, \hat{\gamma}_m) \in \mathcal{I}$. That is, there exists x and $\hat{x}$ for which

$$\begin{pmatrix} f(x) \\ g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} \leq \begin{pmatrix} \phi \\ \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}$$

and similarly for the "hat" versions.

Now, suppose $\mu \in [0, 1]$ and define $x_\mu = (1-\mu)x + \mu\hat{x}$. As f and each of the g_is are convex, we have that

$$\begin{pmatrix} f(x_\mu) \\ g_1(x_\mu) \\ \vdots \\ g_m(x_\mu) \end{pmatrix} \leq (1-\mu) \begin{pmatrix} f(x) \\ g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} + \mu \begin{pmatrix} f(\hat{x}) \\ g_1(\hat{x}) \\ \vdots \\ g_m(\hat{x}) \end{pmatrix} \leq \begin{pmatrix} (1-\mu)\phi + \mu\hat{\phi} \\ (1-\mu)\gamma_1 + \mu\hat{\gamma}_1 \\ \vdots \\ (1-\mu)\gamma_m + \mu\hat{\gamma}_m \end{pmatrix}. \quad \blacksquare$$

Proposition 6. *Suppose $\mathcal{I}$ is convex and there exists $\bar{x}$ so that $g_i(\bar{x}) < 0$ for all $i \in [m]$, then*

$$\begin{aligned} &\inf_{x \in \mathbb{R}^n} \{f(x) : g_i(x) \leq 0, \forall i \in [m]\} \\ &= \sup_{u \in \mathbb{R},\, \lambda \in \mathbb{R}^m} \left\{ u : \begin{array}{l} \lambda \geq 0 \\ f(x) + \sum_{i=1}^m \lambda_i g_i(x) \geq u, \forall x \in \mathbb{R}^n \end{array} \right\} \end{aligned}$$

and the dual problem is solvable.

Proof. The statement is vacuously true if $\mathrm{Opt(Primal)} = -\infty$. Thus, assume that $\mathrm{Opt(Primal)}$ is finite.

Define the set

$$\mathcal{S} := \left\{ \begin{pmatrix} \phi \\ \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} : \begin{array}{l} \phi < \mathrm{Opt(Primal)} \\ \gamma_i \leq 0, \forall i \in [m] \end{array} \right\}.$$

By assumption, $\mathcal{I}$ and $\mathcal{S}$ are disjoint nonempty convex sets. Thus, by the hyperplane separation theorem, there exists $(u, \lambda) \in \mathbb{R}^{1+m}$ not all zero so that

$$\sup_{(\phi,\gamma)\in\mathcal{S}} u\phi + \langle \lambda, \gamma \rangle \leq \inf_{(\phi,\gamma)\in\mathcal{I}} u\phi + \langle \lambda, \gamma \rangle .$$

We claim that $u, \lambda \geq 0$: Otherwise, take the relevant entry in $\mathcal{S}$ increasingly negative to contradict the separation statement.

We claim that $u \neq 0$: Otherwise, the LHS evaluates to zero, whereas our strictly feasible $\bar{x}$ shows that the RHS is negative (recall λ cannot also be identically zero).

Thus, we may assume WLOG that $u = 1$. That is to say,

$$\text{Opt(Primal)} \leq f(x) + \sum_{i=1}^{m} \lambda_i g_i(x),\ \forall x. \qquad \blacksquare$$

Problem 5

Proposition 7. *Suppose K is a proper cone and consider the primal and dual conic problems:*

$$\inf_{x\in\mathbb{R}^n} \left\{ c^\intercal x : \begin{array}{l} Ax - a \in K \\ Bx - b = 0 \end{array} \right\} \geq \sup_{y\in\mathbb{R}^m,\, z\in\mathbb{R}^k} \left\{ \langle a, y\rangle + \langle b, z\rangle : \begin{array}{l} A^\intercal y + B^\intercal z = c \\ y \in K_* \end{array} \right\}.$$

Furthermore, assume that the primal problem is feasible and that:

$$\ker\left(\begin{pmatrix} c^\intercal \\ A \\ B \end{pmatrix}\right) = \{0\} .$$

Then, the primal problem has bounded sublevel sets, i.e.,

$$\forall t \in \mathbb{R},\ \text{ the set } \mathcal{S}_t := \left\{ x \in \mathbb{R}^n : \begin{array}{l} c^\intercal x \leq t \\ Ax - a \in K \\ Bx - b = 0 \end{array} \right\} \text{ is bounded}$$

if and only if the dual problem is strictly feasible.

Proof. First, suppose the dual is strictly feasible. That is, there exists $\bar{y} \in \mathbb{R}^m$ and $\bar{z} \in \mathbb{R}^k$ so that

$$A^\intercal \bar{y} + B^\intercal \bar{z} = c \quad \text{and} \quad \bar{y} \in \operatorname{int}(K_*).$$

For the sake of contradiction, suppose $\mathcal{S}_t$ is unbounded so that $x_0 + \mathbb{R}_+ \hat{x} \subseteq \mathcal{S}_t$ for some $\hat{x}$ nonzero. Thus, it must hold that

$$c^\intercal \hat{x} \leq 0, \qquad A\hat{x} \in K, \quad \text{and} \quad B\hat{x} = 0.$$

The second deduction here requires that K is closed, which holds by assumption. Thus, $-c^\intercal \hat{x}$, $\langle \bar{y}, A\hat{x}\rangle$, and $\langle B\hat{x}, \bar{z}\rangle$ are nonnegative quantities summing to

$$\langle A\hat{x}, \bar{y}\rangle + \langle B\hat{x}, \bar{z}\rangle - c^\intercal \hat{x} = \langle A^\intercal \bar{y} + B^\intercal \bar{z} - c, \hat{x}\rangle = 0.$$

We deduce that $\langle A\hat{x}, \bar{y}\rangle = 0$. As $\bar{y} \in \operatorname{int}(K_*)$, we deduce that $A\hat{x} = 0$. This contradicts the nondegeneracy assumption.

Now, suppose that the primal problem has bounded sublevel sets. Let t be such that $\mathcal{S}_t$ is nonempty. Thus,

$$\left\{ x \in \mathbb{R}^n : \begin{array}{l} Ax \in K \\ Bx = 0 \\ \langle c, x\rangle \le 0 \end{array} \right\} = \{0\}.$$

We write this set in terms of duals:

$$\{0\} = \left\{ x \in \mathbb{R}^n : \begin{array}{l} \langle A^\intercal y, x\rangle \ge 0,\ \forall y \in K_* \\ \langle B^\intercal z, x\rangle \ge 0,\ \forall z \in \mathbb{R}^n \\ \langle -c, x\rangle \ge 0 \end{array} \right\}.$$

Here, we have used that $(K_*)_* = K$. Now, taking the dual of either side, we get

$$\mathbb{R}^n = \left\{ A^\intercal y + B^\intercal z - \lambda c : \begin{array}{l} y \in K_* \\ z \in \mathbb{R}^n \\ \lambda \ge 0 \end{array} \right\} = \begin{pmatrix} A^\intercal & B^\intercal & -c \end{pmatrix} (K_* \times \mathbb{R}^n \times \mathbb{R}_+).$$

We may now take the relative interior of either side:

$$\mathbb{R}^n = \operatorname{int}(\mathbb{R}^n) = \begin{pmatrix} A^\intercal & B^\intercal & -c \end{pmatrix} (\operatorname{int}(K_*) \times \mathbb{R}^n \times \mathbb{R}_{++}).$$

We deduce that there exists $\bar{y} \in \operatorname{int}(K_*)$, $\bar{z} \in \mathbb{R}^n$, and $\lambda > 0$ such that

$$0 = A^\intercal \bar{y} + B^\intercal \bar{z} - \lambda c.$$

Dividing by λ completes the proof. ∎