IOE 611: Midterm

Yulun Zhuang

October 26, 2024

Problem 1

Obtain the dual cone of the set $K \subset \mathbb{R}^3$ defined as:

$$K = \{(x, y, z) : y > 0, ye^{x/y} \leq z\}$$

Solution. By definition, the dual cone K^* can be written as:

$$\begin{aligned} K^* &= \{(u, v, w) \mid xu + yv + zw \geq 0, \forall (x, y, z) \in K\} \\ &= \left\{(u, v, w) \mid u\frac{x}{y} + v + w\frac{z}{y} > 0, \forall y > 0, e^{x/y} \leq z/y\right\} \end{aligned}$$

Note that if $w < 0$, then for any $u, v \in \mathbb{R}$, consider $(x_0, y_0, z_0) = (0, 1, \max\{1, -\frac{|v|+1}{w}\}) \in K$, we have

$$ux_0 + vy_0 + wz_0 = v + wz_0 \leq v + w \cdot \left(-\frac{|v| + 1}{w}\right) = v - |v| - 1 < 0$$

which implies that $(u, v, w) \notin K^*$.

Thus, $\forall (u, v, w) \in K^*$, we must have $w \geq 0$.

If $w = 0$, consider $y = 1, z = e^x$, then $\forall x \in \mathbb{R}, (x, y, z) \in K$, therefore if $(u, v, 0) \in K^*$, we must have $ux + v \geq 0$ for all $x \in \mathbb{R}$.

Then, we must have $u = 0, v \geq 0$.

Conversely, if $u = w = 0, v \geq 0$, then $\forall (x, y, z) \in K, ux + vy + wz = vy \geq 0$.

If $w > 0$, then for any $(x, y, z) \in K$,

$$u\frac{x}{y} + v + w\frac{z}{y} \geq u\frac{x}{y} + v + we^{x/y}$$

and equality holds if and only if $z/y = e^{x/y}$.

Since $\forall x \in \mathbb{R}, (x, 1, e^x) \in \mathcal{K}$, we know $\forall (u, v, w) \in K^*$,

$$f(x) = ux + v + we^x \geq 0, \forall x \in \mathbb{R}.$$

If $u > 0$, then when $x \to -\infty, ux \to -\infty, w_e^x \to 0 \Rightarrow f(x) \to -\infty$, it leads to a contradiction! Therefore $u \leq 0$.

Note that $f'(x) = we^x + u$.

If $u = 0$, then $f(x)$ increases when x increases, i.e., $\inf_{x\in\mathbb{R}} f(x) = \lim_{x\to\infty} f(x) = v$. Therefore, $(0, v, w) \in K^*$ for $w > 0$ if $v \geq 0$.

If $u < 0$, then $f(x)$ is monotonically increasing for $x \in \left(-\infty, \log \frac{-u}{w}\right)$ and monotonically decreasing for $x \in \left(\log \frac{-u}{w}, +\infty\right)$.

Therefore

$$\min_{x\in\mathbb{R}} f(x) = f\left(\log \frac{-u}{w}\right) = u \log \frac{-u}{w} + v + w \cdot \left(\frac{-u}{w}\right) = u\left(\log \frac{-u}{w} - 1\right) + v$$

Consequently, $(u, v, w) \in K^*$ for $u < 0, w > 0$ if $v \geq u\left(1 - \log \frac{-u}{w}\right)$.

In conclusion,

$$K^* = \{(u, v, w) \mid u = 0, v, w \geq 0\} \cup \{(u, v, w) \mid u < 0, w > 0, v \geq u(1 - \log \frac{-u}{w})\}$$

$\square$

Problem 2

Consider the following optimization problem:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & \mathbf{1}^\top x = 1 \\ & x \succeq 0 \end{aligned} \tag{P1}$$

where $\mathbf{1}$ is a vector of ones, i.e. $\mathbf{1} = [1, \ldots, 1]^\top \in \mathbb{R}^n$. Assume that $f(x)$ is convex and differentiable.

(a) Prove that x^* is optimal for (P1) if and only if x^* is feasible and

$$\min_{1 \le i \le n} \nabla f(x^*)_i \ge \nabla f(x^*)^\top x^*$$

where $\nabla f(x^*)_i$ is the ith element of $\nabla f(x^*)$.

(b) Prove that x^* is optimal for (P1) if and only if x^* is feasible and, for each k,

$$x_k^* > 0 \Rightarrow \nabla f(x^*)_k = \min_{1 \le i \le n} \nabla f(x^*)_i$$

Proof. Since $f(x)$ is convex, by the first order optimality condition, x is optimal if and only if x^* is feasible and for any feasible $y \in \mathbf{dom}\, f$,

$$\nabla f(x^*)^\top (y - x^*) \ge 0 \Rightarrow \nabla f(x^*)^\top y \ge \nabla f(x^*)^\top x^*$$

(a)

"$\Rightarrow$": If x^* is optimal, for any $i = 1, \ldots, n$, let e_i be the unit vector such that all elements except the i-th element are 0 , and the i-th element is 1.

Therefore, $\mathbf{1}^\top e_i = 1$ and $e_i \ge 0$, which indicates that e_i is feasible for any $i = 1, \ldots, n$.

Then $\forall i = 1, \ldots, n$, we have

$$\begin{aligned} & \nabla f\left(x^*\right)^\top \left(e_i - x^*\right) \ge 0 \\ \Leftrightarrow \quad & \nabla f\left(x^*\right)^\top e_i \ge \nabla f\left(x^*\right)^\top x^* \\ \Leftrightarrow \quad & \nabla f\left(x^*\right)_i \ge \nabla f\left(x^*\right)^\top x^* \\ \Leftrightarrow \quad & \min_{1 \le i \le n} \nabla f\left(x^*\right)_i \ge \nabla f\left(x^*\right)^\top x^* \end{aligned}$$

"$\Leftarrow$": If $\min_{1 \le i \le n} \nabla f\left(x^*\right)_i \ge \nabla f\left(x^*\right)^\top x^*$, then for any $y \succeq 0, \mathbf{1}^\top y = 1$,

$$\begin{aligned} \nabla f(x^*)^\top y &= \sum_{i=1}^n y_i \nabla f(x^*)_i \\ &\ge \min_{1 \le i \le n} \nabla f(x^*)_i \sum_{i=1}^n y_i \\ &= \min_{1 \le i \le n} \nabla f(x^*)_i \mathbf{1}^\top y \\ &= \min_{1 \le i \le n} \nabla f(x^*)_i \\ &\ge \nabla f(x^*)^\top x^* \end{aligned}$$

which implies $\nabla f(x^*)^\top y \geq \nabla f(x^*)^\top x^*$ for any feasible y. Thus, x^* is optimal.

(b) From the result of part (a), we only need to prove that

$$\min_{1\leq i\leq n} \nabla f(x^*)_i \geq \nabla f(x^*)^\top x^*$$

if and only if x^* is feasible, and for each $k = 1, \ldots, n$,

$$x_k^* > 0 \Rightarrow \nabla f(x^*)_k = \min_{1\leq i\leq n} \nabla f(x^*)_i$$

"$\Rightarrow$": If $\min_{1\leq i\leq n} \nabla f(x^*)_i \geq \nabla f(x^*)^\top x^*, x^*$ is feasible and $x_k^* > 0$ for each $k = 1, \ldots, n$, then we have

$$\begin{aligned}
\nabla f(x^*)^\top x^* &= \nabla f(x^*)_k x_k^* + \sum_{i\neq k} \nabla f(x^*)_i x_i^* \\
&\geq \nabla f(x^*)_k x_k^* + \sum_{i\neq k} x_i^* \min_{1\leq i\leq n} \nabla f(x^*)_i \\
&= \nabla f(x^*)_k x_k^* + (1 - x_k^*) \min_{1\leq i\leq n} \nabla f(x^*)_i
\end{aligned}$$

Given that $\min_{1\leq i\leq n} \nabla f(x^*)_i \geq \nabla f(x^*)^\top x^*$, and $x_k^* > 0$, we have

$$\begin{aligned}
& \min_{1\leq i\leq n} \nabla f(x^*)_i - (1 - x_k^*) \min_{1\leq i\leq n} \nabla f(x^*)_i \geq x_k^* \nabla f(x^*)_k \\
\Leftrightarrow \quad & \min_{1\leq i\leq n} \nabla f(x^*)_i \geq \nabla f(x^*)_k .
\end{aligned}$$

Since we also have $\nabla f(x^*)_k \geq \min_{1\leq i\leq n} \nabla f(x^*)_i$, we must have $\nabla f(x^*)_k = \min_{i\leq i\leq n} \nabla f(x^*)_i$.

"$\Leftarrow$": If x^* is feasible and for each k, $x_k^* > 0 \Rightarrow \nabla f(x^*)_k = \min_{1\leq i\leq n} \nabla f(x^*)_i$, we can separate x as two parts. First part we have x_k for some k such that $x_k > 0$, and the second part we have x_l for l includes the rest of x such that $x_l = 0$.

$$\begin{aligned}
\nabla f(x^*)^\top x^* &= \sum_{i=1}^{n} x_i^* \nabla f(x^*)_i \\
&= \sum_{k} x_k^* \nabla f(x^*)_k + \sum_{l} x_l^* \nabla f(x^*)_l \\
&= \min_{1\leq i\leq n} \nabla f(x^*)_i \sum_{k} x_k^* \\
&= \min_{l\leq i\leq n} \nabla f(x^*)_i (\mathbf{1}^\top x^* - \sum_{l} x_l^*) \\
&= \min_{1\leq i\leq n} \nabla f(x^*)_i
\end{aligned}$$

Therefore, $\nabla f(x^*)^\top x^* \leq \min_{1\leq i\leq n} \nabla f(x^*)_i$. $\square$

Problem 3

Consider the function $f : \mathbb{R}^{n\times n} \times \mathbb{R}^n \to \mathbb{R}$ defined by $f(X, y) = y^\top X^{-1} y$, where $X \in \mathbb{R}^{n\times n}$ and $y \in \mathbb{R}^n$.

(a) Suppose that $\mathbf{dom}\, f = \{(X, y) : X \in \mathbb{S}^n_{++}\}$. Is $f(X, y)$ convex? If so, prove it. If not, give a counter-example.

(b) Suppose that $\mathbf{dom}\, f = \{(X, y) : X + X^\top \in \mathbb{S}^n_{++}\}$. Is $f(X, y)$ convex? If so, prove it. If not, give a counter-example.

Solution. (a) To determine if $f(X, y)$ is convex, we can check if the epigraph of f, defined as

$$\mathbf{epi}\, f = \{(X, y, t) \mid y^\top X^{-1} y \le t,\ X \succ 0\},$$

is convex. Using the Schur complement, we have

$$t - y^\top X^{-1} y \ge 0 \quad \text{and} \quad X \succ 0 \quad \iff \quad \begin{bmatrix} X & y \\ y^\top & t \end{bmatrix} \succ 0, X \succ 0$$

Thus, the epigraph can be expressed as

$$\mathbf{epi}\, f = \{(X, y, t) \mid \begin{bmatrix} X & y \\ y^\top & t \end{bmatrix} \succ 0,\ X \succ 0\}.$$

This set is the intersection of two convex sets (a positive semi-definite matrix constraint and $X \succ 0$), which is itself convex.

Therefore, $f(X, y)$ is convex when $\mathbf{dom}\, f = \{(X, y) \mid X \in \mathbb{S}^n_{++}\}$.

(b) When the domain is changed to $\mathbf{dom}\, f = \{(X, y) \mid X + X^\top \in \mathbb{S}^n_{++}\}$, the function $f(X, y)$ is no longer convex. We can find a counterexample which violates the convexity.

Consider the matrices and vectors:

$$X_1 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \quad y_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad y_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Evaluate it on the Jensen's inequality.

$$f(\frac{1}{2}(X_1 + X_2), \frac{1}{2}(y_1 + y_2)) = \frac{1}{4}$$

and

$$\frac{1}{2} f(X_1, y_1) + \frac{1}{2} f(X_2, y_2) = \frac{2}{3}$$

which does not satisfy convexity in this domain. $\square$

Problem 4

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function with $f(0) \leq 0$.

(a) Show that the perspective function $tf(x/t)$ with domain $\{(x,t) \mid t > 0, x/t \in \textbf{dom}\, f\}$ is nonincreasing as a function of t for a fixed x.

(b) Let $g : \mathbb{R}^n \to \mathbb{R}$ be concave and positive on its domain. Show that the following function is convex:

$$h(x) = g(x)f(x/g(x)), \quad \textbf{dom}\, h = \{x \in \textbf{dom}\, g \mid x/g(x) \in \textbf{dom}\, f\}$$

(c) Using the result of part (b), show that the following function is convex:

$$h(x) = \frac{\|x\|_2^2}{(\prod_{i=1}^n x_i)^{1/n}}, \quad \textbf{dom}\, h = \mathbb{R}^n_{++}$$

Proof. (a) Take t and s, such that $0 < t \leq s$ and $x/t, x/s \in \textbf{dom}\, f$,

$$\begin{aligned} f(\frac{x}{s}) &= f(\frac{x}{t} \cdot \frac{t}{s} + 0 \cdot (1 - \frac{t}{s})) \\ &\leq \frac{t}{s} f(\frac{x}{t}) + (1 - \frac{t}{s}) f(0) \\ &\leq \frac{t}{s} f(\frac{x}{t}) \\ \Leftrightarrow \quad s f(\frac{x}{s}) &\leq t f(\frac{x}{t}) \end{aligned}$$

which implies $tf(x/t)$ is nonincreasing for a fixed x.

(b) To show $h(x)$ is convex, we first show $\textbf{dom}\, h$ is convex, i.e. $\forall x, y \in \textbf{dom}\, h, \alpha \in [0,1] \Rightarrow \alpha x + (1-\alpha)y \in \textbf{dom}\, h$, which is equivalent to show that

$$\alpha x + (1-\alpha)y \in \textbf{dom}\, g \ \text{ and } \ \frac{\alpha x + (1-\alpha)y}{g(\alpha x + (1-\alpha)y)} \in \textbf{dom}\, f$$

Take $\forall x, y \in \textbf{dom}\, h, \alpha \in [0,1]$.

Given $g(x)$ is concave, we have $\textbf{dom}\, g$ is convex.

Therefore, $\alpha x + (1-\alpha)y \in \textbf{dom}\, g$.

$$\begin{aligned} \frac{\alpha x + (1-\alpha)y}{g(\alpha x + (1-\alpha)y)} =& \frac{\alpha g(x)}{g(\alpha x + (1-\alpha)y)} \cdot \frac{x}{g(x)} \\ &+ \frac{(1-\alpha)g(y)}{g(\alpha x + (1-\alpha)y)} \cdot \frac{y}{g(y)} \\ &+ (1 - \frac{\alpha g(x) + (1-\alpha)g(y)}{g(\alpha x + (1-\alpha)y)}) \cdot 0 \end{aligned}$$

where $x/g(x), y/g(y), 0 \in \textbf{dom}\, f$ and their corresponding coefficients sum up to 1.

This implies it is a convex combination of elements within a convex set $\textbf{dom}\, f$, which also lies in the same set, i.e.

$$\frac{\alpha x + (1-\alpha)y}{g(\alpha x + (1-\alpha)y)} \in \textbf{dom}\, f$$

Then, we need to show that $h(x)$ satisfies the Jensen's inequality, i.e.

$$h(\alpha x + (1-\alpha)y) \leq \alpha h(x) + (1-\alpha)h(y)$$

$$\begin{aligned}
h(\alpha x + (1-\alpha)y) &= g(\alpha x + (1-\alpha)y) f(\frac{\alpha x + (1-\alpha)y}{g(\alpha x + (1-\alpha)y)}) \\
&\leq (\alpha g(x) + (1-\alpha)g(y)) f(\frac{\alpha x + (1-\alpha)y}{\alpha g(x) + (1-\alpha)g(y)}) \Leftarrow \text{from part (a)} \\
&= (\alpha g(x) + (1-\alpha)g(y)) f(\frac{\alpha g(x)}{\alpha g(x) + (1-\alpha)g(y)} \cdot \frac{x}{g(x)} + \frac{(1-\alpha)g(y)}{\alpha g(x) + (1-\alpha)g(y)} \cdot \frac{y}{g(y)}) \\
&\leq \alpha g(x) f(\frac{x}{g(x)}) + (1-\alpha)g(y) f(\frac{y}{g(y)}) \Leftarrow f \text{ is convex} \\
&= \alpha h(x) + (1-\alpha)h(y)
\end{aligned}$$

Thus, $h(x)$ is convex.

(c) Take $f(x) = \|x\|_2^2$, such that $f(x)$ is convex and $f(0) = 0$. Take $g(x) = (\prod_{i=1}^n)^{1/n}, \mathbf{dom}\, g = \mathbb{R}_{++}^n$, such that $g(x)$ is concave and positive on its domain.

From the results of part (a) and (b), the function $h(x)$

$$\begin{aligned}
h(x) &= g(x) f(x/g(x)) \\
&= (\prod_{i=1}^n)^{1/n} \left\| \frac{x}{(\prod_{i=1}^n)^{1/n}} \right\|_2^2 \\
&= \frac{\|x\|_2^2}{(\prod_{i=1}^n x_i)^{1/n}}
\end{aligned}$$

with domain

$$\mathbf{dom}\, h = \{x \in \mathbf{dom}\, g, x/g(x) \in \mathbf{dom}\, f\} = \mathbb{R}_{++}^n$$

is a convex function.

$\square$

Problem 5

(a) Let $X = \{x_1, x_2, \ldots, x_m\}$ be a set of points in $\mathbb{R}^n$, where $m \geq n+2$. Show that X can be partitioned in two sets Y and $Z = X \backslash Y$ such that $\mathbf{conv}\, Y \cap \mathbf{conv}\, Z$ is non-empty. Here, $\mathbf{conv}\, Y$ and $\mathbf{conv}\, Z$ refer to the convex hulls of Y and Z, respectively.

(b) Let $S_1, \ldots, S_m$ be a collection of convex sets in $\mathbb{R}^n$, where $m \geq n+2$. Suppose that the intersection of every $m-1$ sets from this collection is non-empty. Then the intersection of all sets $S_1 \cap \cdots \cap S_m$ is non-empty.

Proof. (a) Since $[x_i, 1]^\top, 1 \leq i \leq m$ are $\mathbb{R}^{n+1}$ vectors and $m > n+1$, we know that they are linearly dependent, i.e. $\exists y \in \mathbb{R}^m \backslash \{0\}$,

$$\begin{bmatrix} x_1 & \cdots & x_m \\ 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = 0 \quad \Rightarrow \quad \begin{cases} \sum_{i=1}^m y_i x_i = 0 \\ \sum_{i=1}^m y_i = 0. \end{cases}$$

Define $S = \{i \mid i = 1, \ldots, m, y_i \geq 0\}$, $T = \{i \mid i = 1, \ldots, m\} \backslash S = \{i \mid i = 1, \ldots, m, y_i < 0\}$. Since $y \neq 0$ we must have $S, T \neq \varnothing$.

Therefore, we have

$$\begin{cases} \sum_{i=1}^m y_i x_i = 0 & \Leftrightarrow \quad \sum_{i \in S} y_i x_i = \sum_{j \in T} (-y_j)\, x_j \\ \sum_{i=1}^m y_i = 0 & \Leftrightarrow \quad \sum_{i \in S} y_i = \sum_{j \in T} (-y_j) > 0 \\ \forall i \in S, y_i \geq 0; \forall j \in T, -y_j > 0 \end{cases} \tag{1}$$

Define $c = \sum_{i \in S} y_i > 0$ and

$$z_{i|j} = \begin{cases} y_i / c & \text{if } i \in S \\ -y_j / c & \text{if } j \in T \end{cases}$$

Then, divide (1) by c, we have

$$\sum_{i \in S} z_i x_i = \sum_{j \in T} z_j x_j := v$$
$$\sum_{i \epsilon s} z_i = \sum_{j \in T} z_j = 1$$
$$\forall i \in \{1, \cdots, m\}, z_i \geq 0$$

which implies that v can be expressed as two convex combinations of $z_{i|j}, x_{i|j}$ for $i \in S, j \in T$.

Consequently, let $Y = \{x_i \mid i \in S\}, Z = \{x_j \mid j \in T\}$, we have $\exists v \in \mathbb{R}^n$, $v \in \mathbf{conv}\, Y$ and $v \in \mathbf{conv}\, Z$, i.e. $\mathbf{conv}\, Y \cap \mathbf{conv}\, Z$ is non-empty. $\square$