SI251 - Convex Optimization, Fall 2021
Homework 1 - Solution

Due on Oct. 24, 2021, 23:59

Read all the instructions below carefully before you start working on the assignment, and before you make a submission.

- You are required to write down all the major steps towards making your conclusions; otherwise you may obtain limited points ($\leq 20\%$) of the problem.
- Write your homework in English; otherwise you will get no points of this homework.
- You should submit the electronic version of your answer on the blackboard. LaTeX is encouraged; If you submit the handwriting version, scan it clearly.
- Do your homework by yourself. Any form of plagiarism will lead to 0 point of this homework. If more than one plagiarisms during the semester are identified, we will prosecute all violations to the fullest extent of the university regulations, including but not limited to failing this course, academic probation, or expulsion from the university.
- If you have any doubts regarding the grading, you need to contact the instructor or the TAs within two days since the grade is announced.

I. Convex Set

1. Show that

(a) A set is convex if and only if its intersection with any line is convex. (5 points)

(b) A set is affine if and only if its intersection with any line is affine. (5 points)

Solution:

(a) Proof of "$\Rightarrow$".
The intersection of two convex sets is convex. Therefore if S is a convex set, the intersection of S with a line is convex.
Proof of "$\Leftarrow$".
Conversely, suppose the intersection of S with any line is convex. Take any two distinct points x_1 and $x_2 \in S$. The intersection of S with the line through x_1 and x_2 is convex. Therefore convex combinations of x_1 and x_2 belong to the intersection, hence also to S.

(b) Similar to problem 1(a).

2. *Support function.* The support function of a set $C \subseteq \mathbf{R}^n$ is defined as

$$S_C(y) = \sup \left\{ y^T x \mid x \in C \right\}$$

(We allow $S_C(y)$ to take on the value $+\infty$.) Suppose that C and D are closed convex sets in $\mathbb{R}^n$. Show that $C = D$ if and only if their support functions are equal. (10 points)

Solution:

Obviously, if $C = D$ the support functions are equal. We show that if the support functions are equal, then $C = D$ by showing that $D \subseteq C$ and $C \subseteq D$.

We first show that $D \subseteq C$. Suppose there exists a point $x_0 \in D, x \notin C$. Since C is closed, x_0 can be strictly separated from C, *i.e.*, there exists an $a \neq 0$ with $a^T x_0 > b$ and $a^T x < b$ for all $x \in C$. This means that

$$\sup_{x \in C} a^T x \leq b < a^T x_0 \leq \sup_{x \in D} a^T x$$

which implies that $S_C(a) \neq S_D(a)$. By repeating the argument with the roles of C and D reversed, we can show that $C \subseteq D$.

3. *The monotone nonnegative cone.* We define the *monotone nonnegative cone* as

$$K_{\mathrm{m}+} = \{x \in \mathbf{R}^n \mid x_1 \geq x_2 \geq \cdots \geq x_n \geq 0\}$$

i.e., all nonnegative vectors with components sorted in nonincreasing order.

(a) Show that $K_{\mathrm{m}+}$ is a proper cone.(5 points)

(b) Find the dual cone $K_{\mathrm{m}+}^*$. *Hint.* Use the identity (10 points)

$$\sum_{i=1}^{n} x_i y_i = (x_1 - x_2) y_1 + (x_2 - x_3)(y_1 + y_2) + (x_3 - x_4)(y_1 + y_2 + y_3) + \cdots$$
$$+ (x_{n-1} - x_n)(y_1 + \cdots + y_{n-1}) + x_n (y_1 + \cdots + y_n)$$

Solution:

(a) The set K_{m+} is defined by n homogeneous linear inequalities, hence it is a closed (polyhedral) cone.
The interior of K_{m+} is nonempty, because there are points that satisfy the inequalities with strict inequality, for example, $x = (n, n-1, n-2, \ldots, 1)$.
To show that K_{m+} is pointed, we note that if $x \in K_{m+}$, then $-x \in K_{m+}$ only if $x = 0$. This implies that the cone does not contain an entire line.

(b) Using the hint, we see that $y^T x \geq 0$ for all $x \in K_{m+}$ if and only if

$$y_1 \geq 0, \quad y_1 + y_2 \geq 0, \quad \ldots, y_1 + y_2 + \cdots + y_n \geq 0.$$

Therefore

$$K_{\mathrm{m}+}^* = \left\{ y \mid \sum_{i=1}^{k} y_i \geq 0, k = 1, \ldots, n \right\}.$$

II. **Convex Function**

1. Determine the convexity (i.e., convex, concave, or neither) of the following functions.

(a) $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}^2_{++}$. (5 points)

(b) $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$, where $0 \le \alpha \le 1$, on $\mathbb{R}^2_{++}$. (5 points)

(c) $f(X) = \mathrm{tr}(X^{-1})$ on $\mathbf{dom} f = \mathbb{S}^n_{++}$. (5 points)

(d) $f(X) = (\det X)^{1/n}$ on $\mathbf{dom} f = \mathbb{S}^n_{++}$. (5 points)

Solution:

(a) Neither. The Hessian of f is

$$\nabla^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

which is neither positive semi-definite nor negative semi-definite. Therefore, f is neither convex nor concave. It is quasi-concave, since its superlevel sets

$$\left\{(x_1, x_2) \in \mathbf{R}^2_{++} \mid x_1 x_2 \ge \alpha\right\}$$

are convex. Is is not quasi-convex.

(b) Concave. The Hessian is

$$\begin{aligned}
\nabla^2 f(x) &= \begin{bmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)(-\alpha)x_1^{\alpha}x_2^{-\alpha-1} \end{bmatrix} \\
&= \alpha(1-\alpha)x_1^{\alpha}x_2^{1-\alpha} \begin{bmatrix} -1/x_1^2 & 1/x_1x_2 \\ 1/x_1x_2 & -1/x_2^2 \end{bmatrix} \\
&= -\alpha(1-\alpha)x_1^{\alpha}x_2^{1-\alpha} \begin{bmatrix} 1/x_1 \\ -1/x_2 \end{bmatrix} \begin{bmatrix} 1/x_1 \\ -1/x_2 \end{bmatrix}^T \\
&\preceq 0.
\end{aligned}$$

f is not convex.

(c) Convex. Define $g(t) = f(Z + tV)$, where $Z \succ 0$ and $V \in S^n$.

$$\begin{aligned}
g(t) &= \mathrm{tr}\left((Z+tV)^{-1}\right) \\
&= \mathrm{tr}\left(Z^{-1}\left(I + tZ^{-1/2}VZ^{-1/2}\right)^{-1}\right) \\
&= \mathrm{tr}\left(Z^{-1}Q(I+t\Lambda)^{-1}Q^T\right) \\
&= \mathrm{tr}\left(Q^TZ^{-1}Q(I+t\Lambda)^{-1}\right) \\
&= \sum_{i=1}^{n} \left(Q^TZ^{-1}Q\right)_{ii}(1+t\lambda_i)^{-1}
\end{aligned}$$

where we used the eigenvalue decomposition $Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^T$. In the last equality we express g as a positive weighted sum of convex functions $1/(1+t\lambda_i)$, hence it is convex.

(d) Concave. Define $g(t) = f(Z + tV)$, where $Z \succ 0$ and $V \in S^n$.

$$\begin{aligned}
g(t) &= (\det(Z+tV))^{1/n} \\
&= \left(\det Z^{1/2} \det\left(I + tZ^{-1/2}VZ^{-1/2}\right)\det Z^{1/2}\right)^{1/n} \\
&= (\det Z)^{1/n}\left(\prod_{i=1}^{n}(1+t\lambda_i)\right)^{1/n}
\end{aligned}$$

where $\lambda_i, i = 1, \dots, n$, are the eigenvalues of $Z^{-1/2}VZ^{-1/2}$. From the last equality we see that g is a concave function of t on $\{t | Z + tV \succ 0\}$, since $\det Z > 0$ and the geometric mean $(\prod_{i=1}^n x_i)^{1/n}$ is concave on $\mathbb{R}^n_{++}$.

2. Show that the following functions $f : \mathbb{R}^n \to \mathbb{R}$ are convex.

(a) The 'exponential barrier' of a set of inequalities:

$$f(\mathbf{x}) = \sum_{i=1}^{m} e^{-1/f_i(\mathbf{x})}, \ \mathbf{dom}\ f = \{\mathbf{x} \mid f_i(\mathbf{x}) < 0, i = 1, \ldots, m\}.$$

The functions $f_i(\mathbf{x})$ are convex. (10 points)

(b) The function

$$f(x) = \frac{\|Ax+b\|_2^2}{c^T x + d}$$

dom $f = \{x | c^T x + d > 0\}$, where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$ and $d \in \mathbb{R}$. (10 points)

Solution:

(a) $h(u) = \exp(1/u)$ is convex and decreasing on $\mathbb{R}_{++}$:

$$h^{'}(u) = -\frac{1}{u^2} e^{1/u}, \quad h^{''}(u) = \frac{2}{u^3} e^{1/u} + \frac{1}{u^4} e^{1/u}. \tag{1}$$

Therefore the composition $h(-f_i(x)) = \exp(-1/f_i(x))$ is convex if f_i is convex.

(b) This function is the composition of the function $g(y,t) = y^T y/t$ with an affine transformation $(y,t) = (Ax+b, c^T x + d)$. Therefore convexity of f follows from the fact that g is convex on $\{(y,t) | t > 0\}$. For convexity of g one can note that it is the perspective of $x^T x$, or directly verify that the Hessian

$$\nabla^2 g(y,t) = \begin{bmatrix} I/t & -y/t^2 \\ -y^T/t & y^T y/t^3 \end{bmatrix}$$

is positive semi-definite, since

$$\begin{bmatrix} v \\ w \end{bmatrix}^T \begin{bmatrix} I/t & -y/t^2 \\ -y^T/t & y^T y/t^3 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \|tv - yw\|_2^2/t^3 \geq 0$$

for all v and w.

III. Convex Problem

1. (*Linear Programming*)

(a) Formulate the following problem as LP, and explain the relation in detail. (5 points)

$$\text{minimize} \quad \|Ax - b\|_1 + \|x\|_\infty$$

(b) Consider the LP by

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

with A square and nonsingular. Show that the optimal value is given (5 points)

$$p^\star = \begin{cases} c^T A^{-1} b & A^{-T} c \preceq 0 \\ -\infty & \text{otherwise} \end{cases}$$

Solution:

(a) For minimize $\|Ax - b\|_1$, it is equivalent to the LP

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T \mathbf{s} \\ \text{subject to} & Ax - b \preceq \mathbf{s} \\ & Ax - b \geq -\mathbf{s} \end{array} \tag{1}$$

Assume x is fixed in this problem, and we optimize only over s. The constraints say that

$$-s_k \leq a_k^T x - b_k \leq s_k$$

for each k, *i.e.*, $s_k \geq |a_k^T - b_k|$. The objective function of the LP is separable, so we achieve the optimum over s by choosing

$$s_k = |a_k^T - b_k|,$$

and obtain the optimal value $p^*(x) = \|Ax - b\|_1$. Therefore optimizing over x and s simultaneously is equivalent to problem (1).

For minimize $\|x\|_\infty$, it is equivalent to the LP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & x \preceq t\mathbf{1} \\ & x \succeq -t\mathbf{1} \end{array} \tag{2}$$

in the variables x, t. To see the equivalence, assume x is fixed in this problem, and we optimize only over t. The constraints say that

$$-t \le x_k \le t$$

for each k, *i.e.*, $t > |x_k|$, *i.e.*,

$$t \ge \max_k |x_k| = \|x\|_\infty$$

Clearly, if x is fixed, the optimal value of the LP is $\|x\|_\infty$. Therefore optimizing over t and x simultaneously is equivalent to problem (2).

For minimize $\|Ax - b\|_1 + \|x\|_\infty$, it is equivalent to the LP

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T\mathbf{s} + t \\ \text{subject to} & -\mathbf{s} \preceq Ax - b \preceq \mathbf{s} \\ & -t\mathbf{1} \preceq x \preceq t\mathbf{1}, \end{array} \tag{3}$$

(b) Make a change of variables $y = Ax$. The problem is equivalent to

$$\begin{array}{ll} \text{minimize} & c^T A^{-1} y \\ \text{subject to} & y \preceq b \end{array} \tag{4}$$

If $A^{-T}c \preceq 0$, the optimal solution is $y = b$, with $p^* = c^T A^{-1} b$. Otherwise, the LP is unbounded below.

2. (*Semidefinite Programming*) Consider $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}_0 + x_1\boldsymbol{A}_1 + \cdots + x_n\boldsymbol{A}_n$, where the vector $\boldsymbol{x} \in \mathbb{R}^n$ and the matrix $\boldsymbol{A}_i \in \mathbb{S}^m$, for $i = 0, 1, \cdots, n$. Let $\lambda_1(\boldsymbol{x}) \ge \cdots \ge \lambda_m(\boldsymbol{x})$ denotes the eigenvalues of $\boldsymbol{A}(\boldsymbol{x})$. Equivalently reformulate the following problems as SDPs.

(a) $\min_{\boldsymbol{x}} \quad \lambda_1(\boldsymbol{x})$. (5 points)

(b) $\min_{\boldsymbol{x}} \quad \lambda_1(\boldsymbol{x}) - \lambda_m(\boldsymbol{x})$. (5 points)

(c) $\min_{\boldsymbol{x}} \quad \sum_{i=1}^m |\lambda_i(\boldsymbol{x})|$. (5 points)

Solution:

(a) $\lambda_1(x) \le t$ if and only if $A(x) \preceq tI$, so we have

$$\begin{array}{ll} \underset{\boldsymbol{x},t}{\text{minimize}} & t \\ \text{subject to} & \boldsymbol{A}(\boldsymbol{x}) \preceq t\boldsymbol{I} \end{array}$$

(b) $\lambda_1(x) \le t_1$ if and only if $A(x) \preceq t_1 I$ and $\lambda_m(x) \ge t_2$ if and only if $A(x) \succeq t_2 I$, so we have

$$\begin{array}{ll} \underset{x,t_1,t_2}{\text{minimize}} & t_1 - t_2 \\ \text{subject to} & t_2 I \preceq \boldsymbol{A}(x) \preceq t_1 I \end{array}$$

(c) Method 1:

Suppose $A(x)$ has eigenvalue decomposition $Q\Lambda Q^T$. Let $A(x) = A^+ - A^- = Q\Lambda^+ Q^T - Q\Lambda^- Q^T$. Λ is divided into two parts: Λ^+ and Λ^-. $\lambda_i(x) \ge 0$ are in the Λ^-. Thus $A^+ \succeq 0$ and $A^- \succeq 0$. $\sum_{i=1}^m |\lambda_i(x)|$ is equivalent to $\text{trace}(A^+) + \text{trace}(A^-)$.

$$\begin{array}{ll} \underset{\boldsymbol{x},\boldsymbol{A}^+,\boldsymbol{A}^-}{\text{minimize}} & \text{trace}\left(\boldsymbol{A}^+\right) + \text{trace}\left(\boldsymbol{A}^-\right) \\ \text{subject to} & \boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}^+ - \boldsymbol{A}^- \\ & \boldsymbol{A}^+ \succeq 0 \\ & \boldsymbol{A}^- \succeq 0 \end{array}$$

Method 2:

Similarly to ℓ_1 norm of vector, we have

$$\begin{array}{ll} \underset{\boldsymbol{x},\boldsymbol{Y}}{\text{minimize}} & \text{trace}(\boldsymbol{Y}) \\ \text{subject to} & \boldsymbol{Y} + \boldsymbol{A}(\boldsymbol{x}) \succeq 0 \\ & \boldsymbol{Y} - \boldsymbol{A}(\boldsymbol{x}) \succeq 0 \end{array}$$