# DSA3102: 2023/24 (Sem 1): Solutions to Midterm (Total 30 points)

Name: ______________________________

Matriculation Number: ____________________

Score: Q1: Q2: Q3: Total:

You have 1.5 hours (90 mins) for this quiz. There are SIX (6) printed pages. You're allowed 1 sheet of handwritten notes. Please provide *careful explanations* for all your solutions.

1. (a) For $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$, $x_{[k]}$ denotes the $k$-th largest entry of $x$, so, for example, $x_{[1]} = \max_{1\le i\le n} x_i$ and $x_{[n]} = \min_{1\le i\le n} x_i$. For each function, determine whether it is convex, concave, or neither. If you say it is neither convex nor concave, give a counterexample showing it is not convex, and a counterexample showing it is not concave. All functions below have domain $\mathbb{R}^n$.

   i. (3 points) The range of values $f(x) = x_{[1]} - x_{[n]}$;
   **Solution:** We know that the functions $f_{\max}(x) = x_{[1]}$ and $f_{\min}(x) = x_{[n]}$ are convex and concave respectively. Hence, $f_{\max}$ and $-f_{\min}$ are both convex. Since the sum of convex functions is convex, so is $f = f_{\max} - f_{\min}$.

   ii. (3 points) $\text{median}(x) = x_{[(n+1)/2]}$ (you may assume $n$ is odd);
   **Solution:** Consider the points

$$x = [0, 2, 0]^T, \quad x' = [2, 0, 0]^T.$$

Then it is easy to check that

$$\text{median}(x) = 0, \quad \text{median}(x') = 0.$$

However,

$$\text{median}(0.5x + 0.5x') = 1.$$

If median$(\cdot)$ were convex, we would have

$$\text{median}(0.5x + 0.5x') \le 0.5\text{median}(x) + 0.5\text{median}(x')$$

which is clearly not the case here since $1 \not\le 0$.
Again consider

$$x = [0, -2, 0]^T, \quad x' = [-2, 0, 0]^T.$$

Then it is easy to check that

$$\text{median}(x) = 0, \quad \text{median}(x') = 0.$$

However,

$$\text{median}(0.5x + 0.5x') = -1.$$

If median$(\cdot)$ were concave, we would have

$$\text{median}(0.5x + 0.5x') \geq 0.5\text{median}(x) + 0.5\text{median}(x')$$

which is clearly not the case here since $-1 \not\geq 0$.

(b) (4 points) Find the set of $\alpha \in \mathbb{R}$ such that the following function $f : \mathbb{R}^2 \to \mathbb{R}$ is convex

$$f(x_1, x_2) = \alpha x_1^2 + \frac{1}{4}\alpha x_2^2 + (4 - \alpha)x_1 x_2 + 4x_1 + 8x_2 + 5.$$

**Solution:** Differentiating once yields

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2\alpha x_1 + (4 - \alpha)x_2 \\ \frac{1}{2}\alpha x_2 + (4 - \alpha)x_1 \end{bmatrix}$$

Differentiating again yields

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 2\alpha & 4 - \alpha \\ 4 - \alpha & \frac{1}{2}\alpha \end{bmatrix}$$

We need $\alpha \geq 0$ for both diagonal entries to be nonnegative. For the determinant, we can calculate that $2\alpha \cdot \frac{1}{2}\alpha - (4 - \alpha)^2 \geq 0$. This is the same as $\alpha \geq 2$. Since $\alpha \geq 2$ dominates, the desired set of $\alpha$ such that $f$ is convex is $[2, \infty)$.

2. (a) (5 points) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable convex function. Consider the following problem:

$$\text{(P)} \qquad \min_x \ f(x) \quad \text{s.t.} \quad x \succeq 0, \quad \sum_{i=1}^n x_i = 1.$$

Show using the first-order optimality conditions that if $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)^T$ is an optimal solution for problem (P), then

$$\min_{i=1,\ldots,n} (\nabla f(\bar{x}))_i = \sum_{i=1}^n \bar{x}_i (\nabla f(\bar{x}))_i,$$

where $(\nabla f(\bar{x}))_i = \frac{\partial f}{\partial x_i}(\bar{x})$ is the $i$-th component of the gradient vector evaluated at $\bar{x}$.

**Solution:** The first-order optimality condition for $\bar{x}$ to be optimal is that it is feasible, i.e., $\bar{x} \geq 0$ and $\sum_{i=1}^n \bar{x}_i = 1$ and

$$\nabla f(\bar{x})^T (y - x) \geq 0 \qquad \forall\, y \in \mathbb{R}^n \quad \text{s.t.} \quad y \geq 0, \sum_{i=1}^n y_i = 1.$$

This means that

$$\sum_{i=1}^n y_i (\nabla f(\bar{x}))_i \geq \sum_{i=1}^n \bar{x}_i (\nabla f(\bar{x}))_i,$$

for all $y \geq 0, \sum_{i=1}^n y_i = 1$. Let $i^* = \arg\min_{1 \leq i \leq n} (\nabla f(\bar{x}))_i$. Set $y = e_{i^*}$ which satisfies $y \geq 0, \sum_{i=1}^n y_i = 1$. Then the LHS of the above display writes $\min_{1 \leq i \leq n} (\nabla f(\bar{x}))_i$ so that

$$\min_{1 \leq i \leq n} (\nabla f(\bar{x}))_i \geq \sum_{i=1}^n \bar{x}_i (\nabla f(\bar{x}))_i.$$

Since $\bar{x}$ itself forms a probability distribution, the strict inequality above $>$ is not possible. Hence,

$$\min_{1 \leq i \leq n} (\nabla f(\bar{x}))_i = \sum_{i=1}^n \bar{x}_i (\nabla f(\bar{x}))_i,$$

as desired.

(b) (5 points) Find the convex conjugate of the function

$$f(x) = \sqrt{1+x^2} \quad \text{with} \quad \mathbf{dom}\ f = \mathbb{R}.$$

As usual, denote the convex conjugate as $f^*(y)$.

*Hint: First, argue that the domain of $f^*$ is $[-a, a]$ for some $a > 0$. You need to find $a$.*

**Solution:** Consider

$$f^*(y) = \sup_{x \in \mathbb{R}} \left\{ xy - \sqrt{1+x^2} \right\}$$

Note that if $|y| > 1$, the objective is unbounded because $\sqrt{1+x^2}$ behaves as $|x|$ for $x$ large. So the domain of $f^*$ is $[-1, 1]$.

Hence, we restrict ourselves to $y \in [-1, 1]$. Let $g(x) = xy - \sqrt{1+x^2}$. Then $g'(x) = y - x/\sqrt{1+x^2}$. Setting this to zero yields that $x^* = \text{sign}(y)\sqrt{\frac{y^2}{1-y^2}}$. Substituting this into the objective yields

$$f^*(y) = -\sqrt{1-y^2} \qquad \text{for} \quad |y| \leq 1.$$

3. (a) (5 points) Consider the following optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \sum_{i=1}^{n} c_i |x_i - d_i| \\ \text{s.t.} \quad & Ax = b, \quad x \geq 0. \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and $d \in \mathbb{R}^n$ are given. Assume that $c_i \geq 0$ for all $i = 1, \ldots, n$. As such this is not a linear program since the objective function involves absolute values. Show how this problem can be formulated equivalently as a linear program. Write your answer as a linear optimization problem where the decision variables are in a length-$2n$ vector $y \in \mathbb{R}^{2n}$ of the form

$$\min_{y \in \mathbb{R}^{2n}} \tilde{c}^T y \quad \text{subject to} \quad \tilde{A} y = \tilde{b}, \quad \tilde{G} y \leq \tilde{h}.$$

Identify the vectors $\tilde{c}, \tilde{b}$ and $\tilde{h}$ and the matrices $\tilde{A}$ and $\tilde{G}$.

**Solution:** This may be written as

$$\begin{aligned} \min_{x \in \mathbb{R}^n, t \in \mathbb{R}^n} \quad & c^T t \\ \text{s.t.} \quad & Ax = b, \quad x \geq 0, \quad t \geq 0, \quad |x_i - d_i| \leq t_i, \quad \forall\, i \in [n]. \end{aligned}$$

Let $y = [x^T, t^T]^T$. Then the objective can be written as

$$\begin{bmatrix} 0 \\ c \end{bmatrix}^T \begin{bmatrix} x \\ t \end{bmatrix} = c^T t.$$

The equality constraints can be written as

$$\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}.$$

The inequality constraints can be written as

$$\begin{bmatrix} -I & 0 \\ 0 & -I \\ I & -I \\ -I & -I \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ d \\ -d \end{bmatrix}.$$

Hence,

$$\tilde{c} = \begin{bmatrix} 0 \\ c \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} b \\ 0 \end{bmatrix}, \quad \tilde{G} = \begin{bmatrix} -I & 0 \\ 0 & -I \\ I & -I \\ -I & -I \end{bmatrix}, \quad \tilde{h} = \begin{bmatrix} 0 \\ 0 \\ d \\ -d \end{bmatrix}.$$

(b) (5 points) In this problem, the decision variable $x \in \mathbb{R}^n$. The matrix $A \in \mathbf{S}_+^n$ is fixed. Show that the following optimization problem can be written as a semidefinite program (SDP):

$$\begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & x^T(A - bb^T)x \leq 0 \\ & b^T x \geq 0 \\ & Dx = z \end{array}$$

by writing the first constraint as

$$\begin{bmatrix} f(x,b,A) & g(x,b,A) \\ g(x,b,A)^T & h(x,b,A) \end{bmatrix} \succeq 0$$

for suitable functions $f(x,b,A), g(x,b,A)$, and $h(x,b,A)$.

*Hint: Since $A \in \mathbf{S}_+^n$, it may be written as $A = V^T V$ for some $V \in \mathbb{R}^{r \times n}$. You might want to use the Schur complement lemma.*

**Solution:** Since $A$ is PSD, we may write $A$ as $V^T V$ for some $V \in \mathbb{R}^{r \times n}$. Hence, the first constraint is

$$x^T(V^T V - bb^T)x \leq 0 \qquad \Longleftrightarrow \qquad x^T V^T V x \leq x^T bb^T x.$$

This is the same as

$$\|Vx\|_2^2 \leq (b^T x)^2.$$

Using the fact that $b^T x \geq 0$, we have that

$$\|Vx\|_2 \leq b^T x.$$

This is a second-order cone constraint. We can use the Schur complement lemma to write this as the following semidefinite constraint:

$$\begin{bmatrix} (b^T x)I & Vx \\ (Vx)^T & b^T x \end{bmatrix} \succeq 0.$$

Hence $f(x,b,A) = (b^T x)I$, $g(x,b,A) = Vx$ and $h(x,b,A) = b^T x$.