ENGG 5501: Foundations of Optimization	**2023–24 First Term**
Homework Set 2 Solution	
Instructor: Anthony Man–Cho So	October 28, 2023

Problem 1 (20pts).

(a) **(10pts).** Let $A = U\Sigma U^T$ be a spectral decomposition of A. Then, we have

$$A \bullet B = \mathrm{tr}(AB) = \mathrm{tr}\left(\Sigma(U^T BU)\right) = \sum_{i=1}^{n} \Sigma_{ii}(U^T BU)_{ii}. \tag{1}$$

Since $A \in \mathcal{S}^n_+$, we have $\Sigma_{ii} \geq 0$ for $i = 1, \ldots, n$. Moreover, since $B \in \mathcal{S}^n_+$, we have $U^T BU \in \mathcal{S}^n_+$, which implies that $(U^T BU)_{ii} \geq 0$ for $i = 1, \ldots, n$. It follows from (1) that $A \bullet B \geq 0$, as desired.

(b) **(10pts).** Let $X \in \mathcal{S}^n$ be such that $A \bullet X \geq 0$ for any $A \in \mathcal{S}^n_+$. We claim that $X \in \mathcal{S}^n_+$. Suppose that this is not the case. Let $X = U\Sigma U^T$ be a spectral decomposition of X. Then, there exists an $i \in \{1, \ldots, n\}$ such that $\Sigma_{ii} < 0$. Consider the matrix $A = Ue_ie_i^T U^T \in \mathcal{S}^n_+$, where $e_i \in \mathbb{R}^n$ is the i-th basis vector. A simple calculation shows that $A \bullet X = \Sigma_{ii} < 0$, which is a contradiction.

Problem 2 (30pts).

(a) **(10pts).** Observe that for any $y \in \mathbb{R}$, we have

$$f^*(y) = \sup_{x\in\mathbb{R}}\{yx - f(x)\} = \sup_{x\geq 0}\{yx - x\ln x\}.$$

It is easy to verify (using, e.g., calculus) that the function $x \mapsto yx - x\ln x$ is maximized at $x^* = \exp(y-1) > 0$. It follows that

$$f^*(y) = y\exp(y-1) - (y-1)\exp(y-1) = \exp(y-1).$$

(b) **(10pts).** Observe that for any $y \in \mathbb{R}$, we have

$$f^*(y) = \sup_{x\in\mathbb{R}}\{yx - |x|\} = \sup_{x\in\mathbb{R}}(|y|-1)|x|.$$

It follows that

$$f^*(y) = \begin{cases} 0 & \text{if } |y| \leq 1, \\ +\infty & \text{otherwise.} \end{cases}$$

(c) **(10pts).** Observe that for any $y \in \mathbb{R}^n$, we have

$$i_C^*(y) = \sup_{x\in\mathbb{R}^n}\left\{y^T x - i_C(x)\right\} = \sup_{x\in C} y^T x.$$

Now, let $x \in C$ and $y \in \mathbb{R}^n$ be arbitrary. We compute

$$y^T x = \sum_{i:y_i\geq 0} x_iy_i + \sum_{i:y_i<0} x_iy_i \leq \sum_{i:y_i\geq 0} x_iy_i = y^T x' = y_+^T x,$$

where $x' \in C$ is given by

$$x'_j = \begin{cases} x_j & \text{if } y_j \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

and $y_+ \in \mathbb{R}^n$ is given by $(y_+)_j = \max\{y_j, 0\}$, for $j = 1, \ldots, n$. It follows from the Cauchy–Schwarz inequality that $i_C^*(y) = \|y_+\|_2$.

Problem 3 (25pts).

(a) **(15pts).** Let $x, y \in \mathbb{R}^n$ be arbitrary. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by $g(t) = f(x + t(y - x))$. By the Fundamental Theorem of Calculus and the Chain Rule, we have

$$f(y) - f(x) = g(1) - g(0) = \int_0^1 g'(t)\, dt = \int_0^1 \nabla f(x + t(y - x))^T (y - x)\, dt.$$

Upon writing

$$\int_0^1 \nabla f(x+t(y-x))^T(y-x)\, dt = \int_0^1 \left[\nabla f(x + t(y-x)) - \nabla f(x)\right]^T (y-x)\, dt + \nabla f(x)^T (y-x)$$

and using the Cauchy–Schwarz inequality together with the Lipschitz continuity of ∇f, we obtain

$$\begin{aligned} \left| f(y) - f(x) - \nabla f(x)^T (y - x) \right| &\leq \int_0^1 \left| \left[\nabla f(x + t(y-x)) - \nabla f(x)\right]^T (y - x) \right| dt \\ &\leq L\|y - x\|_2^2 \int_0^1 t\, dt \\ &= \frac{L}{2}\|y - x\|_2^2, \end{aligned}$$

as desired.

(b) **(10pts).** Using the result in (a) with $y = x - \alpha\nabla f(x)$, we have

$$\begin{aligned} f(y) - f(x) &\leq \nabla f(x)^T(-\alpha\nabla f(x)) + \frac{L}{2}\|x - y\|_2^2 = -\alpha\|\nabla f(x)\|_2^2 + \frac{L}{2}\|x - y\|_2^2 \\ &= \left(-\frac{1}{\alpha} + \frac{L}{2}\right)\|x - y\|_2^2, \end{aligned}$$

as desired.

Problem 4 (10pts).

(a) **(10pts).** The systems (I) and (II) cannot be simultaneously solvable. Indeed, suppose that $\bar{x} \in \mathbb{R}^n$ solves (I) and $\bar{y} \in \mathbb{R}^m$ solves (II). Then, since $\bar{y} > \mathbf{0}$, $A\bar{x} \leq \mathbf{0}$ and $A\bar{x} \neq \mathbf{0}$, we have $\bar{y}^T A\bar{x} < 0$. On the other hand, since $\bar{x} \geq \mathbf{0}$ and $A^T\bar{y} \geq \mathbf{0}$, we have $\bar{y}^T A\bar{x} \geq 0$. This results in a contradiction.

Suppose that (I) is not solvable. Then, by a simple scaling argument, we see that

$$(\text{I}') \quad Ax \leq \mathbf{0},\ e^T Ax = -1,\ x \geq \mathbf{0}$$

is not solvable either. By introducing slack variables, we see that (I$'$) is equivalent to

$$\begin{bmatrix} A & I \\ e^T A & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -1 \end{bmatrix}, \ (x,s) \geq \mathbf{0}.$$

Hence, by Farkas' lemma, there exists a $\bar{z} = (\bar{u}, \bar{t}) \in \mathbb{R}^{m+1}$ such that

$$\begin{bmatrix} A^T & A^T e \\ I & \mathbf{0} \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{t} \end{bmatrix} \geq \mathbf{0}, \ \bar{t} > 0,$$

or equivalently,

$$A^T(\bar{u} + \bar{t}e) \geq \mathbf{0}, \ \bar{u} \geq \mathbf{0}, \ \bar{t} > 0.$$

Now, let $\bar{y} = \bar{u} + \bar{t}e \in \mathbb{R}^m$. Clearly, we have $A^T \bar{y} \geq \mathbf{0}$. Moreover, since $\bar{u} \geq \mathbf{0}$ and $\bar{t} > 0$, we have $\bar{y} \geq \bar{t}e > \mathbf{0}$. It follows that (II) is solvable, as desired.

(b) **(15pts).** The systems (I) and (II) cannot be simultaneously solvable. Indeed, if $\tilde{x} \in \mathbb{R}^n$ solves (I) and $\tilde{y} \in \mathbb{R}^m$ solves (II), then

$$d < c^T \tilde{x} = \tilde{y}^T A \tilde{x} \leq b^T \tilde{y} \leq d,$$

which is a contradiction.

We claim that (I) is solvable iff

$$(\text{I}') \quad Ax - bt \leq \mathbf{0}, \ \ c^T x - dt > 0, \ \ t \geq 0$$

is solvable. Indeed, if x' solves (I), then $(x', 1)$ solves (I$'$). Conversely, suppose that (x', t') solves (I$'$). If $t' > 0$, then it is easy to verify that x'/t' solves (I). If $t' = 0$, then we have $Ax' \leq \mathbf{0}$ and $c^T x' > 0$. Since $A\bar{x} \leq b$ by assumption, upon letting $x'' = \bar{x} + \theta x'$ with $\theta > 0$ sufficiently large, we have $c^T x'' = c^T \bar{x} + \theta c^T x' > d$ and $Ax'' = A(\bar{x} + \theta x') \leq b$. It follows that x'' solves (I).

Now, note that (I$'$) takes the form

$$(\text{I}') \quad \begin{bmatrix} A & -b \\ \mathbf{0} & -1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq \mathbf{0}, \ \ \begin{bmatrix} c^T & -d \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} > 0.$$

Suppose that (I$'$) is not solvable. By Farkas' lemma, we see that

$$(\text{II}') \quad \begin{bmatrix} A^T & \mathbf{0} \\ -b^T & -1 \end{bmatrix} \begin{bmatrix} y \\ s \end{bmatrix} = \begin{bmatrix} c \\ -d \end{bmatrix}, \ \ y \geq \mathbf{0}, \ \ s \geq 0$$

is solvable. This implies that (II) is solvable, as can be easily verified.