ENGG 5501: Foundations of Optimization	2021–22 First Term

Homework Set 2 Solution

Instructor: Anthony Man–Cho So — October 11, 2021

Problem 1 (20pts).

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

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

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

(b) **(10pts).** Let $X \in \mathcal{S}^n$ be such that $A \bullet X \ge 0$ 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 $i \in \{1, \dots, n\}$ such that $\Sigma_{ii} < 0$. Consider the matrix $A = Ue_i e_i^T U^T \in \mathcal{S}_+^n$, 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 (25pts).

(a) **(15pts).** Suppose that a continuous function f possesses Property C. Let $t \in \mathbb{R}$ be arbitrary. Since $L_t = f^{-1}((-\infty, t])$ and $(-\infty, t]$ is closed in $\mathbb{R}$, the continuity of f implies that L_t is closed. Now, suppose that L_t is unbounded. Then, we can find a sequence $\{x^k\}_{k\ge 0}$ in L_t such that $\|x^k\|_2 \to +\infty$. However, since $x^k \in L_t$ for all $k \ge 0$, we have $f(x^k) \le t$ for all $k \ge 0$. This contradicts the fact that f possesses Property C. Hence, L_t is bounded, which together with its closedness implies that it is compact.

Conversely, suppose that L_t is compact for any $t \in \mathbb{R}$. If a sequence $\{x^k\}_{k\ge 0}$ in $\mathbb{R}^n$ satisfies $f(x^k) \le T < +\infty$ for all $k \ge 0$, then it belongs to the compact set L_T. In particular, the sequence is bounded. Thus, if $\|x^k\|_2 \to +\infty$, then we necessarily have $f(x^k) \to +\infty$; i.e., f possesses Property C.

Lastly, observe that for any $\bar{x} \in \mathbb{R}^n$, we have

$$\inf_{x\in\mathbb{R}^n} f(x) = \inf_{x\in L_{f(\bar{x})}} f(x).$$

If f is continuous and possesses Property C, then $L_{f(\bar{x})}$ is compact. The desired result then follows from Weierstrass' theorem.

(b) **(10pts).** No. Take, for example, $f(x) = \|x\|_2$ and $A = \mathbf{0}$.

Problem 3 (30pts).

(a) **(5pts).** First, the closedness of K° follows from the fact that it is the (possibly infinite) intersection of halfspaces, which are closed.

Next, let us establish the convexity of K°. Let $w, w' \in K^\circ$ and $\alpha \in [0,1]$ be arbitrary. Then, by definition of K°, for all $x \in K$, we have

$$(\alpha w + (1-\alpha)w')^T x = \alpha w^T x + (1-\alpha)(w')^T x \le 0,$$

which shows that K° is convex. Note that the above argument does not depend on the convexity of K. In other words, K° is convex regardless of whether K is convex.

Lastly, let us show that K° is a cone. Let $w \in K^\circ$ and $\alpha > 0$. Then, by definition of K°, for all $x \in K$, we have $(\alpha w)^T x = \alpha w^T x \le 0$. This completes the proof.

(b) **(15pts).** By Theorem 3 of Handout 2, we have $z^* = \Pi_K(x)$ if and only if $z^* \in K$ and

$$(x - z^*)^T (z - z^*) \le 0 \quad \text{for all } z \in K. \tag{2}$$

Suppose that $z^* = \Pi_K(x)$. Since $z^* \in K$ and K is a cone, we have $\alpha z^* \in K$ for all $\alpha > 0$. In particular, by taking $z = \alpha z^*$ in (2), we have

$$(\alpha - 1)(x - z^*)^T z^* \le 0 \quad \text{for all } \alpha > 0.$$

This implies that $(x - z^*)^T z^* = 0$. Upon substituting this into (2), we have $(x - z^*)^T z \le 0$ for all $z \in K$, which, by definition of K°, means that $x - z^* \in K^\circ$.

Conversely, if $z^* \in K$, $x - z^* \in K^\circ$, and $(x - z^*)^T z^* = 0$, then (2) holds, which implies that $z^* = \Pi_K(x)$.

(c) **(15pts).** Let $z^* = \Pi_K(x)$. By the result in (b), we know that $x - z^* \in K^\circ$. By Theorem 3 of Handout 2, it remains to show that

$$(x - (x - z^*))^T (w - (x - z^*)) \le 0 \quad \text{for all } w \in K^\circ.$$

We compute

$$\begin{aligned} (x - (x - z^*))^T (w - (x - z^*)) &= (z^*)^T (w - (x - z^*)) \\ &= (z^*)^T w - (z^*)^T (x - z^*) \\ &= (z^*)^T w \\ &\le 0, \end{aligned}$$

where the third line follows from the result in (b) that $(x - z^*)^T z^* = 0$, and the last line follows from the fact that $z^* \in K$ and $w \in K^\circ$. It follows that $z^* = \Pi_{K^\circ}(x)$.

Problem 4 (20pts). Let $g : \mathbb{R}^n \to \mathbb{R}$ be given by $g(x) = f(x) + \frac{\rho}{2}\|x\|_2^2$.

(1)$\Longleftrightarrow$(2): We compute

$$\begin{aligned} &\alpha\|x\|_2^2 + (1-\alpha)\|y\|_2^2 - \|\alpha x + (1-\alpha)y\|_2^2 \\ &= \alpha\|x\|_2^2 + (1-\alpha)\|y\|_2^2 - \alpha^2\|x\|_2^2 - 2\alpha(1-\alpha)x^T y - (1-\alpha)^2\|y\|_2^2 \\ &= \alpha(1-\alpha)(\|x\|_2^2 + \|y\|_2^2 - 2x^T y) \\ &= \alpha(1-\alpha)\|x - y\|_2^2. \end{aligned}$$

Hence, we have

$$
\begin{aligned}
& f \text{ is } \rho\text{-convex for some } \rho \in \mathbb{R} \\
\iff & f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y) + \frac{\rho\alpha(1-\alpha)}{2}\|x-y\|_2^2, \quad \forall x, y \in \mathbb{R}^n;\ \alpha \in [0,1] \\
\iff & f(\alpha x + (1-\alpha)y) \\
& \quad \le \alpha f(x) + (1-\alpha) f(y) + \frac{\rho}{2}\left(\alpha\|x\|_2^2 + (1-\alpha)\|y\|_2^2 - \|\alpha x + (1-\alpha)y\|_2^2\right), \ \forall x, y \in \mathbb{R}^n;\ \alpha \in [0,1] \\
\iff & g(\alpha x + (1-\alpha)y) \le \alpha g(x) + (1-\alpha) g(y), \quad \forall x, y \in \mathbb{R}^n;\ \alpha \in [0,1] \\
\iff & g \text{ is convex.}
\end{aligned}
$$

(2)$\iff$(3): Since $\nabla g(x) = \nabla f(x) + \rho x$, we invoke Theorem 9 of Handout 2 to get

$$
\begin{aligned}
& g \text{ is convex} \\
\iff & g(y) \ge g(x) + \nabla g(x)^T (y-x), \quad \forall x, y \in \mathbb{R}^n \\
\iff & f(y) \ge f(x) + \nabla f(x)^T (y-x) + \frac{\rho}{2}(\|x\|_2^2 - \|y\|_2^2) + \rho x^T (y-x), \quad \forall x, y \in \mathbb{R}^n \\
\iff & f(y) \ge f(x) + \nabla f(x)^T (y-x) + \frac{\rho}{2}(x+y-2x)^T(x-y), \quad \forall x, y \in \mathbb{R}^n \\
\iff & f(y) \ge f(x) + \nabla f(x)^T (y-x) - \frac{\rho}{2}\|x-y\|_2^2, \quad \forall x, y \in \mathbb{R}^n.
\end{aligned}
$$