ENGG 5501: Foundations of Optimization	**2018–19 First Term**
Homework Set 1 Solution	
Instructor: Anthony Man–Cho So	October 1, 2018

Problem 1 (10pts). For $i = 1, \ldots, n$, define the decision variable y_i by

$$y_i = \begin{cases} 1 & \text{if } x_i \text{ is non–zero,} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Then, the constraint $\|x\|_0 \leq K$ implies that

$$\sum_{i=1}^{n} y_i \leq K. \tag{2}$$

Moreover, by assumption, there exists a constant $M > 0$ such that $\|x^*\|_\infty \leq M$ for some optimal solution x^* to problem (S). Hence, by imposing the constraint

$$y_i = 1 \quad \Longrightarrow \quad |x_i| \leq M \qquad \text{for } i = 1, \ldots, n, \tag{3}$$

we preserve at least one optimal solution to the original problem (S), namely, x^*. Note that in view of (1), the constraint (3) can be written as

$$x_i \leq My_i, \quad x_i \geq -My_i \qquad \text{for } i = 1, \ldots, n. \tag{4}$$

Hence, by putting (2) and (4) together, we obtain the following equivalent formulation of problem (S) that only involves linear and binary constraints:

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|_2^2 \\ \text{subject to} & \displaystyle\sum_{i=1}^{n} y_i \leq K, \\ & x_i \leq My_i \qquad \text{for } i = 1, \ldots, n, \\ & x_i \geq -My_i \qquad \text{for } i = 1, \ldots, n, \\ & y_i \in \{0, 1\} \qquad \text{for } i = 1, \ldots, n. \end{array}$$

Problem 2 (20pts).

(a) **(10pts).** The set S needs not be convex. For instance, take $S = \mathbb{Q}$, the set of rational numbers. Then, we have $(x + y)/2 \in \mathbb{Q}$ whenever $x, y \in \mathbb{Q}$, but $\mathbb{Q}$ is not convex as it is not even connected.

Remark. As it turns out, if S is a *closed* set satisfying the property that $(x + y)/2 \in S$ whenever $x, y \in S$, then S *is* in fact convex. Indeed, let $x, y \in S$ and consider the point $z \in [x, y]$. Construct a sequence $\{z_k\}_{k\geq 0}$, where $z_k = (x_k + y_k)/2$, $x_0 = x$, $y_0 = y$, and

$$x_{k+1} = \begin{cases} z_k & \text{if } z_k \in [x, z], \\ x_k & \text{if } z_k \in (z, y]; \end{cases} \qquad y_{k+1} = \begin{cases} z_k & \text{if } z_k \in (z, y], \\ y_k & \text{if } z_k \in [x, z]. \end{cases} \tag{5}$$

Intuitively, one can view $\{z_k\}_{k\geq 0}$ as the sequence generated by a binary search over the line segment $[x, y]$ for z. Now, let us prove by induction that for all $k \geq 0$,

$$x_k, y_k, z_k \in S, \quad z \in [x_k, y_k], \quad \|z_k - z\|_2 \leq \frac{1}{2}\|x_k - y_k\|_2, \quad \|x_k - y_k\|_2 = \frac{1}{2^k}\|x - y\|_2. \tag{6}$$

The base case (i.e., $k = 0$) follows directly from the fact that $x, y \in S$ and $z_0 = (x+y)/2 \in S$. By the inductive hypothesis and (5), it is clear that $x_{k+1}, y_{k+1}, z_{k+1} \in S$ and $z \in [x_{k+1}, y_{k+1}]$. Since $z_{k+1} = (x_{k+1} + y_{k+1})/2$, it follows that

$$\|z_{k+1} - z\|_2 \leq \frac{1}{2}\|x_{k+1} - y_{k+1}\|_2.$$

Moreover, using (5) and the inductive hypothesis, we have

$$\|x_{k+1} - y_{k+1}\|_2 = \frac{1}{2}\|x_k - y_k\|_2 = \frac{1}{2^{k+1}}\|x - y\|_2.$$

This completes the inductive step.

To complete the proof, observe from (6) that $z_k \to z$ as $k \to \infty$. Since S is closed, we have $z \in S$.

(b) **(10pts).** Let $x, y \in S$ and $\alpha \in [0, 1]$ be arbitrary. Since $x, y \in \mathbb{R}^n_+$, we have

$$\alpha x_i + (1-\alpha)y_i \geq x_i^\alpha y_i^{1-\alpha} \qquad \text{for } i = 1, \ldots, n.$$

This implies that

$$\prod_{i=1}^n (\alpha x_i + (1-\alpha)y_i) \geq \prod_{i=1}^n \left(x_i^\alpha y_i^{1-\alpha}\right) = \left(\prod_{i=1}^n x_i\right)^\alpha \left(\prod_{i=1}^n y_i\right)^{1-\alpha} \geq 1.$$

Hence, we conclude that S is convex.

Problem 3.

(a) Let $x_1, x_2 \in S$ and $\alpha \in (0, 1)$. Then, we have

$$\begin{aligned} x_1^T A x_1 + b^T x_1 + c &\leq 0, & (7)\\ x_2^T A x_2 + b^T x_2 + c &\leq 0. & (8) \end{aligned}$$

Now, we compute

$$\begin{aligned}
& (\alpha x_1 + (1-\alpha)x_2)^T A (\alpha x_1 + (1-\alpha)x_2) + b^T (\alpha x_1 + (1-\alpha)x_2) + c \\
= \; & (\alpha x_1 + (1-\alpha)x_2)^T A (\alpha x_1 + (1-\alpha)x_2) + \alpha \left(b^T x_1 + c\right) + (1-\alpha)\left(b^T x_2 + c\right) \\
\leq \; & (\alpha x_1 + (1-\alpha)x_2)^T A (\alpha x_1 + (1-\alpha)x_2) - \alpha x_1^T A x_1 - (1-\alpha)x_2^T A x_2 & (9) \\
= \; & -\alpha(1-\alpha)x_1^T A x_1 - (1-\alpha)(1-(1-\alpha))x_2^T A x_2 + 2\alpha(1-\alpha)x_1^T A x_2 \\
= \; & -\alpha(1-\alpha)\left(x_1^T A x_1 - 2x_1^T A x_2 + x_2^T A x_2\right) \\
= \; & -\alpha(1-\alpha)(x_1 - x_2)^T A (x_1 - x_2) \\
\leq \; & 0, & (10)
\end{aligned}$$

where (9) follows from the fact that $b^T x_i + c \leq -x_i^T A x_i$ for $i = 1, 2$ (by (7) and (8)), and (10) follows from the assumption that $A \succeq \mathbf{0}$. This proves that S is convex if $A \succeq \mathbf{0}$.

Note that the converse of the claim need not be true. Indeed, let $n = 1$, and let $A = -1$, $b = c = 0$. Then, we have $S = \{x \in \mathbb{R} : -x^2 \leq 0\} = \mathbb{R}$, which is trivially convex.

(b) Let $x_1, x_2 \in S \cap H$ and $\alpha \in (0, 1)$. From the calculations in part (a), we have

$$\begin{aligned} & (\alpha x_1 + (1-\alpha)x_2)^T A (\alpha x_1 + (1-\alpha)x_2) + b^T (\alpha x_1 + (1-\alpha)x_2) + c \\ \leq \;& -\alpha(1-\alpha)(x_1 - x_2)^T A (x_1 - x_2). \end{aligned} \tag{11}$$

Since $A + \gamma g g^T \succeq \mathbf{0}$, we have

$$0 \leq (x_1 - x_2)^T \left(A + \gamma g g^T\right)(x_1 - x_2) = (x_1 - x_2)^T A (x_1 - x_2) + \gamma \left(g^T (x_1 - x_2)\right)^2.$$

It follows from (11) that

$$\begin{aligned} & (\alpha x_1 + (1-\alpha)x_2)^T A (\alpha x_1 + (1-\alpha)x_2) + b^T (\alpha x_1 + (1-\alpha)x_2) + c \\ \leq \;& -\alpha(1-\alpha)(x_1 - x_2)^T A (x_1 - x_2) \\ \leq \;& \alpha(1-\alpha)\gamma \left(g^T (x_1 - x_2)\right)^2 \\ = \;& 0, \end{aligned}$$

where the last equality follows from the fact that $g^T x_1 + h = g^T x_2 + h = 0$.

Problem 4.

(a) Intuitively, the vector $x - \Pi_{H(s,c)}(x)$ should be normal to the hyperplane $H(s, c)$. Hence, we should have $x - \Pi_{H(s,c)}(x) = \alpha s$ for some $\alpha \in \mathbb{R}$. Since $\Pi_{H(s,c)}(x) \in H(s, c)$, this requires that $s^T(x - \alpha s) = c$, which implies that $\alpha = (s^T x - c)/s^T s$. This yields the following candidate for $\Pi_{H(s,c)}(x)$:

$$\Pi_{H(s,c)}(x) = x - \frac{s^T x - c}{s^T s} s. \tag{12}$$

To prove the correctness of the above formula, we use Theorem 5 of Handout 2. Let $y \in H(s, c)$ be arbitrary. Since $s^T y = c$, we obtain

$$\begin{aligned} \left(y - \Pi_{H(s,c)}(x)\right)^T \left(x - \Pi_{H(s,c)}(x)\right) &= \left(y - x + \frac{s^T x - c}{s^T s} s\right)^T \left(\frac{s^T x - c}{s^T s} s\right) \\ &= \frac{s^T x - c}{s^T s} s^T y - \frac{s^T x - c}{s^T s} s^T x + \frac{(s^T x - c)^2}{s^T s} \\ &= 0. \end{aligned}$$

This establishes the correctness of the formula in (12).

(b) Let $A \in \mathcal{S}^n$ be arbitrary and $A = U\Lambda U^T$ be its spectral decomposition. Observe that

$$(\Lambda - \Lambda^+)_{ii} = \min\{\Lambda_{ii}, 0\} \quad \text{for } i = 1, \ldots, n.$$

Hence, for any $Q \in \mathcal{S}^n_+$, we have

$$\begin{aligned}
(Q - \Pi_{\mathcal{S}^n_+}(A)) \bullet (A - \Pi_{\mathcal{S}^n_+}(A)) &= (Q - \Pi_{\mathcal{S}^n_+}(A)) \bullet U(\Lambda - \Lambda^+)U^T \\
&= (U^T QU - \Lambda^+) \bullet (\Lambda - \Lambda^+) \\
&= \sum_{i=1}^n \left[(U^T QU)_{ii} \cdot \min\{\Lambda_{ii}, 0\} - \Lambda^+_{ii} \cdot \min\{\Lambda_{ii}, 0\} \right] \\
&= \sum_{i:\Lambda_{ii} \le 0} (U^T QU)_{ii} \cdot \Lambda_{ii} \\
&\le 0,
\end{aligned}$$

where the last inequality follows from the fact that $U^T QU \in \mathcal{S}^n_+$ and the diagonal entries of a psd matrix are non–negative. This, together with Theorem 5 of Handout 2, completes the proof.