ENGG 5501: Foundations of Optimization	**2016–17 First Term**
Homework Set 4	
Instructor: Anthony Man–Cho So	Due: November 21, 2016

SOLVE THE FOLLOWING PROBLEMS:

Problem 1 (25pts). Let $A_1, \ldots, A_m \in \mathcal{S}^n_+$, $\alpha_1, \ldots, \alpha_m > 0$, and $\beta_1, \ldots, \beta_m > 0$ be given. Consider the following SDP:

$$\begin{array}{lll} \inf & \displaystyle\sum_{i=1}^m \mathrm{tr}(Z_i) & \\ \text{subject to} & \mathrm{tr}\left[A_i \left(\alpha_i Z_i - \displaystyle\sum_{j \neq i} Z_j \right) \right] \geq \beta_i \quad \text{for } i = 1, \ldots, m, & \\ & Z_1, \ldots, Z_m \in \mathcal{S}^n_+. & \end{array} \tag{Q}$$

(a) **(15pts).** Write down the dual of (Q).

(b) **(10pts).** Using the result in (a), show that the dual is strictly feasible.

Problem 2 (45pts).

(a) **(10pts).** Let $u \in \mathbb{R}^n$ and $U \in \mathcal{S}^n$ be given. Define

$$W = \begin{bmatrix} 1 & u^T \\ u & U \end{bmatrix}.$$

Show that $\mathrm{rank}(W) = 1$ if and only if $W = (1, u)(1, u)^T$.

Let $\mathcal{CP}_n = \mathrm{conv}\left(\{ vv^T : v \in \mathbb{R}^n_+ \} \right)$. The results in Homework 3, Problem 3 imply that $\mathcal{CP}_n$ is a closed pointed cone with non–empty interior. Now, given $Q \in \mathcal{S}^n$ and $c \in \mathbb{R}^n$, consider the following problem:

$$\begin{array}{lll} \text{minimize} & \dfrac{1}{2} x^T Q x + c^T x & \\ \text{subject to} & x_i \in \{0, 1\} & \text{for } i = 1, \ldots, n. \end{array} \tag{1}$$

Our goal now is to show that the discrete optimization problem (1) can be reformulated as a linear optimization problem over the pointed cone $\mathcal{CP}_{2n+1}$. In other words, problem (1) has a convex reformulation![1]

[1] Since problem (1) is NP–hard in general, we see that not all convex optimization problems are easy to solve.

(b) **(5pts).** Using the result in (a), show that problem (1) is equivalent to

$$
\begin{array}{lll}
\text{minimize} & \frac{1}{2}Q \bullet X + c^T x & \\
\text{subject to} & Z = \begin{bmatrix} 1 & x^T & s^T \\ x & X & Y \\ s & Y^T & S \end{bmatrix} \in \mathcal{CP}_{2n+1}, & \text{(I)} \\
& x_i + s_i = 1 \quad \text{for } i = 1, \ldots, n, & \text{(II)} \\
& X_{ii} + 2Y_{ii} + S_{ii} = 1 \quad \text{for } i = 1, \ldots, n, & \text{(III)} \\
& x_i = X_{ii},\ s_i = S_{ii} \quad \text{for } i = 1, \ldots, n, & \text{(IV)} \\
& \text{rank}(Z) = 1. &
\end{array} \tag{2}
$$

(c) **(20pts).** Let

$$
\begin{aligned}
S_1 &= \text{conv}\left(\left\{(1,x,s)(1,x,s)^T : x_i, s_i \in \{0,1\},\ x_i + s_i = 1 \text{ for } i = 1, \ldots, n\right\}\right), \\
S_2 &= \left\{Z \in \mathcal{S}^{2n+1} : Z \text{ satisfies (2–I) — (2–IV)}\right\}.
\end{aligned}
$$

Show that $S_1 = S_2$.

(d) **(10pts).** By relaxing the rank constraint in (2), we obtain a convex relaxation of problem (1), which is a CLP. Using the result in (c), show that this convex relaxation is in fact equivalent to problem (1); i.e., (i) the optimal values of both problems are equal, and (ii) if

$$
Z^* = \begin{bmatrix} 1 & (x^*)^T & (s^*)^T \\ x^* & X^* & Y^* \\ s^* & (Y^*)^T & S^* \end{bmatrix}
$$

is an optimal solution to the convex relaxation, then x^* is in the convex hull of the optimal solutions to problem (1).

Problem 3 (30pts). Consider the ℓ_1–regularized ℓ_2–regression problem

$$
\min_{x \in \mathbb{R}^n,\, t \in \mathbb{R}} \left\{ \|b - Ax - te\|_2 + \lambda \|x\|_1 \right\}, \tag{3}
$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $\lambda > 0$ are given, and $e \in \mathbb{R}^m$ is the vector of all–ones. A common interpretation of the ℓ_1–regularizer $x \mapsto \|x\|_1$, which is based on heuristic arguments, is that it promotes sparsity in the optimal solution to (3). In this problem, we will show in a rigorous manner that problem (3) is actually equivalent to a robust ℓ_2–regression problem.

To begin, consider the following robust optimization problem:

$$
\min_{x \in \mathbb{R}^n,\, t \in \mathbb{R}} \ \max_{\Delta A \in \mathcal{U}_\lambda} \|b - (A + \Delta A)x - te\|_2. \tag{4}
$$

Here, the uncertainty set $\mathcal{U}_\lambda$ is defined as

$$
\mathcal{U}_\lambda = \left\{ X \in \mathbb{R}^{m \times n} : \|X\|_{1,2} \le \lambda \right\},
$$

where $\|X\|_{1,2} = \max_{\|v\|_1 = 1} \|Xv\|_2$. Note that $\|\cdot\|_{1,2}$ defines a matrix norm (see Handout B).

(a) **(10pts).** Show that for any $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ and $\Delta A \in \mathcal{U}_\lambda$,

$$\|b - (A + \Delta A)x - te\|_2 \leq \|b - Ax - te\|_2 + \lambda \|x\|_1.$$

(b) **(20pts).** Show that for any $(x, t) \in \mathbb{R}^n \times \mathbb{R}$, there exists a $\Delta A^* \in \mathcal{U}_\lambda$ such that the inequality in (a) holds as equality; i.e.,

$$\|b - (A + \Delta A^*)x - te\|_2 = \|b - Ax - te\|_2 + \lambda \|x\|_1.$$

Hence, conclude that problems (3) and (4) are equivalent.