ENGG 5501: Foundations of Optimization **2016–17 First Term**

Homework Set 5

Instructor: Anthony Man–Cho So Due: December 7, 2016

SOLVE THE FOLLOWING PROBLEMS:

Problem 1 (30pts). Let $u \in \mathbb{R}^n$ be a given unit vector (i.e., $\|u\|_2 = 1$). The orthogonal projection of the identity matrix I onto the linear subspace

$$\mathcal{L}_u = \left\{ ux^T + xu^T \in \mathcal{S}^n : x \in \mathbb{R}^n \right\} \subset \mathcal{S}^n$$

is, by definition, the solution to the following optimization problem:

$$\min_{x \in \mathbb{R}^n} \left(I - (ux^T + xu^T) \right) \bullet \left(I - (ux^T + xu^T) \right). \tag{1}$$

(a) **(10pts).** Write down the first–order necessary optimality condition of (1).

(b) **(10pts).** Show that the condition found in (a) is also sufficient for optimality of (1).

(c) **(10pts).** Using the results in (a) and (b), determine the optimal solution x^* to (1).

Problem 2 (20pts). Let $A \in \mathcal{S}^n$ and $b \in \mathbb{R}^n$ be given. Consider the optimization problem

$$\begin{array}{rcll} v_p^* & = & \text{minimize} & x^T A x + 2b^T x \\ & & \text{subject to} & x^T x \leq 1. \end{array} \tag{2}$$

Let $u \geq 0$ be the Lagrangian multiplier associated with the constraint $x^T x \leq 1$. Show that the Lagrangian dual of (2) is equivalent to the following SDP:

$$\begin{array}{rcll} v_d^* & = & \text{maximize} & -t - u \\ & & \text{subject to} & \begin{bmatrix} A + uI & b \\ b^T & t \end{bmatrix} \succeq \mathbf{0}, \\ & & & u \geq 0. \end{array} \tag{3}$$

(*Hint: Consider the Moore–Penrose pseudoinverse of $A + uI$.*)

Remark: By the weak duality theorem, we always have $v_p^* \geq v_d^*$. Thus, we can view problem (3) as a convex relaxation of problem (2). In fact, one can use the Shapiro–Barvinok–Pataki theorem (Theorem 1 of Handout 6) and the SDP strong duality theorem to show that such a relaxation is *tight*; i.e., $v_p^* = v_d^*$. We leave this as an optional exercise to the reader.

Problem 3 (20pts). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable convex function and $C \subset \mathbb{R}^n$ be a non–empty closed convex set. Show that x^* is an optimal solution to the convex optimization problem

$$\min_{x \in C} f(x)$$

iff

$$x^* = \Pi_C(x^* - \nabla f(x^*)),$$

where $\Pi_C(\cdot)$ is the projection operator onto C.

Problem 4 (30pts). Let $A^j \in \mathbb{R}^{m_j \times n}$ be given matrices, $C_j \subset \mathbb{R}^{m_j}$ be given non–empty convex sets, and $f_j : \mathbb{R}^{m_j} \to \mathbb{R} \cup \{+\infty\}$ be given functions that are convex on C_j, where $j = 1, \ldots, J$. Consider the following optimization problem:

$$\begin{array}{ll} \inf & \displaystyle\sum_{j=1}^{J} f_j(A^j x) \\ \text{subject to} & A^j x \in C_j \qquad \text{for } j = 1, \ldots, J. \end{array} \tag{4}$$

(a) **(15pts).** Show that (4) is equivalent to

$$\begin{array}{lll} \inf & \displaystyle\sum_{j=1}^{J} f_j(A^j x^j) & \\ \text{subject to} & A^j x^j = A^j \bar{x} & \text{for } j = 1, \ldots, J, \\ & x^j \in (A^j)^{-1} C_j & \text{for } j = 1, \ldots, J, \\ & \bar{x} \in \mathbb{R}^n. & \end{array} \tag{5}$$

(Recall that $(A^j)^{-1} C_j = \{x \in \mathbb{R}^n : A^j x \in C_j\}$.) Hence, by letting $w^j \in \mathbb{R}^{m_j}$ be the Lagrangian multiplier associated with the constraint $A^j x^j = A^j \bar{x}$, where $j = 1, \ldots, J$, show that the Lagrangian dual of (5) can be expressed as

$$\begin{array}{ll} \sup & \displaystyle\sum_{j=1}^{J} \inf_{x \in C_j} \left\{ f_j(x) + (w^j)^T x \right\} \\ \text{subject to} & \displaystyle\sum_{j=1}^{J} (A^j)^T w^j = \mathbf{0}. \end{array}$$

(b) **(15pts).** Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $\lambda > 0$ be given. Using the result in (a), construct a dual of the following *ridge regression problem*:

$$\min_{x \in \mathbb{R}^n} \left\{ \|Ax - b\|_2^2 + \lambda \|x\|_2^2 \right\}.$$

Simplify your answer as much as possible.