ENGG 5501: Foundations of Optimization	**2021–22 First Term**
Homework Set 5 Solution	
Instructor: Anthony Man–Cho So	December 18, 2021

Problem 1 (20pts).

(a) **(10pts).** Recall that we have the problem

$$\begin{array}{ll} \max & x^T A x \\ \text{subject to} & \|x\|_2^2 = 1, \\ & v_1^T x = 0. \end{array} \tag{1}$$

Let $\theta, \gamma \in \mathbb{R}$ be the multipliers associated with the constraints $\|x\|_2^2 = 1$ and $v_1^T x = 0$, respectively. Then, the first-order optimality conditions of Problem (1) are given by

$$\begin{array}{rcll} -2Ax + 2\theta x + \gamma v_1 & = & \mathbf{0}, & (i) \\ \|x\|_2^2 & = & 1, & (ii) \\ v_1^T x & = & 0. & (iii) \end{array}$$

Suppose that $\bar{x} \in \mathbb{R}^n$ is an optimal solution to Problem (1). Note that both constraints of Problem (1) are active at $\bar{x}$, and their gradients are given by $2\bar{x}$ and v_1. Since $\bar{x} \neq \mathbf{0}$ and $v_1^T \bar{x} = 0$, the vectors $2\bar{x}$ and v_1 are linearly independent. It follows from Theorem 3 of Handout 7 that the conditions (i)–(iii) above are necessary for optimality.

(b) **(10pts).** Using (i), (iii), and the fact that $Av_1 = \lambda_1 v_1$ with $\|v_1\|_2^2 = 1$, we have

$$\gamma = \gamma v_1^T v_1 = 2v_1^T (A - \theta I)x = 2\lambda v_1^T x = 0.$$

Hence, we obtain from (i), (ii) that $Ax = \theta x$ (i.e., (θ, x) is an eigenpair of A) and $\theta = x^T A x$. Since the eigenvectors of A form an orthonormal basis of $\mathbb{R}^n$, the solution $x = v_2$ is optimal for Problem (1) with an objective value of $\theta = \lambda_2$.

Remark: Here, by "second largest eigenvalue" we allow for the possibility that $\lambda_1 = \lambda_2$, because in this case we still have $\lambda_1 \geq \lambda_2$ and the eigenspace corresponding to λ_1 is at least 2-dimensional.

Problem 2 (15pts). Suppose that $x^* = \Pi_C(x^* - \nabla f(x^*))$. Since C is a non-empty closed convex set, for any $x \in C$, we have

$$0 \geq (x - \Pi_C(x^* - \nabla f(x^*)))^T \left(x^* - \nabla f(x^*) - \Pi_C(x^* - \nabla f(x^*))\right) = -\nabla f(x^*)^T (x - x^*).$$

This, together with the continuous differentiability and convexity of f, implies that for all $x \in C$,

$$f(x) \geq f(x^*) + \nabla f(x^*)^T (x - x^*) \geq f(x^*);$$

i.e., x^* is an optimal solution to the optimization problem

$$\min_{x \in C} f(x). \tag{2}$$

Conversely, suppose that $x^* \neq \Pi_C(x^* - \nabla f(x^*))$. Since

$$(x^* - \Pi_C(x^* - \nabla f(x^*)))^T (x^* - \nabla f(x^*) - \Pi_C(x^* - \nabla f(x^*))) \leq 0,$$

we have

$$\nabla f(x^*)^T (x^* - \Pi_C(x^* - \nabla f(x^*))) \geq \|x^* - \Pi_C(x^* - \nabla f(x^*))\|_2^2 > 0,$$

or equivalently,

$$\nabla f(x^*)^T (\Pi_C(x^* - \nabla f(x^*)) - x^*) < 0.$$

This implies that $d = \Pi_C(x^* - \nabla f(x^*)) - x^*$ is a descent direction of f at x^*. Moreover, since $x^*, \Pi_C(x^* - \nabla f(x^*)) \in C$, we see that $x^* + \alpha d \in C$ for all $\alpha \in [0, 1]$, which implies that d is also a feasible direction at x^*. It follows that x^* is not an optimal solution to (2).

Problem 3 (15pts). The given problem is equivalent to

$$\begin{array}{ll} \text{minimize} & z \\ \text{subject to} & g_j(x) \leq z \quad \text{for } j = 1, \ldots, m. \end{array} \tag{P}$$

Note that the objective function $(x, z) \mapsto z$ is convex, and for $i = 1, \ldots, m$, the function $(x, z) \mapsto g_i(x) - z$ is continuously differentiable and convex. Hence, the above formulation is a convex optimization problem. Moreover, given any $\bar{x} \in \mathbb{R}^n$, if we let $\bar{z} = \max\{g_1(\bar{x}), \ldots, g_m(\bar{x})\} + 1$, then $g_i(\bar{x}) < \bar{z}$ for $i = 1, \ldots, m$. This shows that Problem (P) satisfies the Slater condition. Hence, by Theorems 4 and 6 of Handout 7, $(x^*, z^*) \in \mathbb{R}^n \times \mathbb{R}$ is an optimal solution to Problem (P) if and only if there exists a $u^* \in \mathbb{R}^m$ such that

$$\begin{array}{rcll} \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} + \displaystyle\sum_{j=1}^m u_j^* \begin{bmatrix} \nabla g_j(x^*) \\ -1 \end{bmatrix} & = & \mathbf{0}, & & (a) \\ u_j^*(g_j(x^*) - z^*) & = & 0 & \text{for } j = 1, \ldots, m, & (b) \\ g_j(x^*) & \leq & z^* & \text{for } j = 1, \ldots, m, & (c) \\ u^* & \geq & \mathbf{0}. & & (d) \end{array}$$

To complete the proof, it remains to show that $z^* = \max\{g_1(x^*), \ldots, g_m(x^*)\}$. From (c), we clearly have $z^* \geq \max\{g_1(x^*), \ldots, g_m(x^*)\}$. On the other hand, using (a), (b), and (d), we obtain

$$z^* = \sum_{j=1}^m u_j^* g_j(x^*) \leq \max\{g_1(x^*), \ldots, g_m(x^*)\} \sum_{j=1}^m u_j^* = \max\{g_1(x^*), \ldots, g_m(x^*)\}.$$

Problem 4 (30pts).

(a) **(5pts).** The semidefinite relaxation of the given QCQP is given by

$$\begin{array}{ll} \inf & A \bullet X \\ \text{subject to} & X_{ii} = 1 \quad \text{for } i = 1, \ldots, n, \\ & X \succeq \mathbf{0}. \end{array} \tag{SDR}$$

(b) **(10pts).** The dual of (SDR) is given by

$$\begin{array}{ll} \sup & e^T y \\ \text{subject to} & A - \mathrm{Diag}(y) \succeq \mathbf{0}. \end{array} \tag{SDD}$$

Note that $\bar{X} = I$ is strictly feasible for (SDR). It follows from the CLP strong duality theorem that the duality gap between (SDR) and (SDD) is zero.

(c) **(15pts).** By a simple manipulation, we have

$$\theta(w) = e^T w + \inf_{x \in \mathbb{R}^n} \left\{ x^T \left(A - \mathrm{Diag}(w) \right) x \right\}.$$

For any given $w \in \mathbb{R}^n$, we claim that

$$\inf_{x \in \mathbb{R}^n} \left\{ x^T \left(A - \mathrm{Diag}(w) \right) x \right\} = \begin{cases} 0 & \text{if } A - \mathrm{Diag}(w) \succeq \mathbf{0}, \\ -\infty & \text{otherwise.} \end{cases} \tag{3}$$

Indeed, if $A - \mathrm{Diag}(w) \not\succeq \mathbf{0}$, then $\lambda_{\min}(A - \mathrm{Diag}(w)) < 0$. Let $u \in \mathbb{R}^n$ be the unit eigenvector corresponding to the smallest eigenvalue of $A - \mathrm{Diag}(w)$. Then, as $\alpha \nearrow +\infty$, we have

$$(\alpha u)^T \left(A - \mathrm{Diag}(w) \right) (\alpha u) = \alpha^2 \lambda_{\min}(A - \mathrm{Diag}(w)) \searrow -\infty,$$

which implies that

$$\inf_{x \in \mathbb{R}^n} \left\{ x^T \left(A - \mathrm{Diag}(w) \right) x \right\} = -\infty.$$

On the other hand, if $A - \mathrm{Diag}(w) \succeq \mathbf{0}$, then $x^T \left(A - \mathrm{Diag}(w) \right) x \geq 0$ for any $x \in \mathbb{R}^n$. In particular, we have

$$\inf_{x \in \mathbb{R}^n} \left\{ x^T \left(A - \mathrm{Diag}(w) \right) x \right\} = 0.$$

This establishes (3). Consequently, the Lagrangian dual is equivalent to

$$\begin{array}{ll} \sup & e^T w \\ \text{subject to} & A - \mathrm{Diag}(w) \succeq \mathbf{0}, \end{array}$$

which is exactly the same problem as (SDD).

Problem 5 (20pts). Recall that we have the problem

$$\begin{array}{ll} \inf & x_1^2 + 4x_2^2 + 16x_3^2 \\ \text{subject to} & x_1 x_2 x_3 = 1. \end{array} \tag{4}$$

(a) **[Extra Credit (15pts).]** Since $(x_1, x_2, x_3) = (1, 1, 1)$ is feasible for Problem (4) with objective value 21, we see that the problem

$$\begin{array}{ll} \inf & x_1^2 + 4x_2^2 + 16x_3^2 \\ \text{subject to} & x_1 x_2 x_3 = 1, \\ & x_1^2 + 4x_2^2 + 16x_3^2 \leq 21 \end{array} \tag{5}$$

has the same set of optimal solutions as Problem (4). Now, observe that the feasible set of Problem (5) is closed (since it is the intersection of the closed sets $S_1 = \{x \in \mathbb{R}^3 : x_1x_2x_3 = 1\}$ and $S_2 = \{x \in \mathbb{R}^3 : x_1^2 + 4x_2^2 + 16x_3^2 \leq 21\}$) and bounded (since the set S_2 is bounded). Moreover, the objective function of Problem (5) is continuous. Hence, by Weierstrass' theorem, Problem (5) has an optimal solution. This in turn implies that Problem (4) has an optimal solution.

(b) **(10pts).** Let $h(x) = x_1x_2x_3 - 1$. Then, we have

$$\nabla h(x) = (x_2x_3, x_1x_3, x_1x_2).$$

In particular, for any (x_1, x_2, x_3) satisfying $x_1x_2x_3 = 1$, we have $\nabla h(x) \neq \mathbf{0}$. It follows that Problem (4) satisfies the linear independence constraint qualification, which implies that the first-order conditions are necessary for optimality. Those conditions are given by

$$\begin{aligned} x_1x_2x_3 &= 1, && (6) \\ \begin{bmatrix} 2x_1 \\ 8x_2 \\ 32x_3 \end{bmatrix} + w \begin{bmatrix} x_2x_3 \\ x_1x_3 \\ x_1x_2 \end{bmatrix} &= \mathbf{0}. && (7) \end{aligned}$$

(c) **(10pts).** Using (6) and (7), we have

$$\begin{aligned} 2x_1^2 &= -w, \\ 8x_2^2 &= -w, \\ 32x_3^2 &= -w. \end{aligned}$$

Upon multiplying the above equations together and using (6) again, we obtain $-w^3 = 512$, which yields $w = -8$. It follows that the first-order optimality conditions (6)–(7) admit the following four solutions:

$$x^1 = \begin{bmatrix} 2 \\ 1 \\ 1/2 \end{bmatrix}, \quad x^2 = \begin{bmatrix} 2 \\ -1 \\ -1/2 \end{bmatrix}, \quad x^3 = \begin{bmatrix} -2 \\ 1 \\ -1/2 \end{bmatrix}, \quad x^4 = \begin{bmatrix} -2 \\ -1 \\ 1/2 \end{bmatrix}.$$

Since all of the above solutions give the same objective value, they must all be optimal for Problem (4).

Problem 6 [Extra Credit (15pts).] Consider the following problem, which is obtained by relaxing the constraint $X \in \mathcal{S}_+^n$ in the original problem:

$$\begin{array}{ll} \sup & \operatorname{tr}(AX) \\ \text{subject to} & \|X - X_0\|_F^2 \leq r^2, \\ & X \in \mathcal{S}^n. \end{array} \qquad (8)$$

Let $X^* \in \mathcal{S}^n$ be an optimal solution to Problem (8), which is guaranteed to exist because the objective function is continuous and the feasible set is compact. Our goal now is to show that $X^* \in \mathcal{S}_+^n$. This would imply that X^* is also optimal for the original problem.

To begin, observe that Problem (8) satisfies the Slater condition (take, e.g., the point X_0). Thus, the first-order optimality conditions of Problem (8), which are given by

$$\begin{aligned}
\|X - X_0\|_F^2 &\leq r^2, && (a)\\
-A + u(X - X_0) &= \mathbf{0}, && (b)\\
u(\|X - X_0\|_F^2 - r^2) &= 0, && (c)\\
u &\geq 0, && (d)
\end{aligned}$$

are necessary for optimality. Since $A \neq \mathbf{0}$, we see from (b) and (d) that $u > 0$ and $X = \frac{1}{u}A + X_0$. These, together with (c), yield $u = \frac{1}{r}\|A\|_F$. It follows that

$$X^* = \frac{r}{\|A\|_F}A + X_0$$

is the optimal solution to Problem (8). Note that since $A, X_0 \in \mathcal{S}_+^n$, we have $X^* \in \mathcal{S}_+^n$. It follows that X^* is the optimal solution to the original problem.