HW10

Bingjie Liu

March 30, 2017

10

1.

a. The feasible region of (P_c) is the same as that of (P_d). Let $a = d - c = (\frac{1-e^Tc}{n}e)$. The objective function of (P_c) is

$$||x - c||_2^2 = (x - c)^T(x - c)$$

The objective function of (P_d) is

$$||x-d||_2^2 = (x-d)^T(x-d) = (x-c)^T(x-c)-a^T(x-c)+a^Ta = (x-c)^T(x-c)-(\frac{1-e^Tc}{n}e)^Tx+a^T(c+a)$$

From the constraint,$e^Tc = 1$. Thus, to minimize $||x - d||_2^2$ is to minimize $(x - c)^T(x - c)$ if we get rid of the constant terms.

b. If $d \geq 0$, d is a feasible solution to (P_d). $||x - d||_2^2 \geq 0$. When $x^* = d$, $||x - d||_2^2$ can achieve its minimum 0. Thus, if $d \geq 0$, $x^* = d$ is an optimal solution.

c. In (P_d), convexity conditions hold. Thus KKT conditions are sufficient for optimality.

$$\nabla f(x) = \begin{bmatrix} x_1 - d_1 \\ \cdots \\ x_n - d_n \end{bmatrix} \nabla g_i(x) = \begin{bmatrix} 0 \\ \cdots \\ 0 \\ -1(\text{the i th entry}) \\ 0 \\ \cdots \\ 0 \end{bmatrix}, \nabla h(x) = \begin{bmatrix} 1 \\ \cdots \\ 1 \end{bmatrix}$$

$$\nabla f(x) + \sum_{i=1}^n u_i \nabla g_i(x) + v\nabla h(x) = \begin{bmatrix} x_1 - d_1 \\ \cdots \\ x_n - d_n \end{bmatrix} + \sum_{i=1}^n u_i \begin{bmatrix} 0 \\ \cdots \\ 0 \\ -1(\text{the i th entry}) \\ 0 \\ \cdots \\ 0 \end{bmatrix} + v \begin{bmatrix} 1 \\ \cdots \\ 1 \end{bmatrix}$$

For any $i \in \{1, ..., n\}$, $x_i^* - d_i - u_i + v = 0$. Sum it over $\{1, ..., n\}$, $\sum_{i=1}^n x_i^* - \sum_{i=1}^n d_i - \sum_{i=1}^n u_i + nv = 0$. Since $e^Tx = e^Td = 1$. Thus, $v = 1/n \sum_{i=1}^n u_i \geq 0$. For any index j, that

$d_j < 0$, $x_i = d_i + u_i - v$. Since $x_i \geq 0$, $u_i > 0$. Also, $u_j x_j^* = 0$ from KKT condition. Then, $x_j^* = 0$ for any index j, that $d_j < 0$ in optimal solution.

d. From above, from (P_c), (P_d) can be come up with accordingly. If there is any j, that $d_j < 0$, set $x_j^* = 0$, and remove it from the system and come up with $(P_c)'$. Repeat the procedures until for all index j, $d_j \geq 0$. For all remaining x_j^*, set it to d_j. Then x^* is the optimal solution.

10 2. $C = D \Leftrightarrow \forall y \in \mathbb{R}^n, S_C(y) = S_D(y)$.
"$\Rightarrow$", $S_C(y) = sup\{y^T x : x \in C\} = sup\{y^T x : x \in D\} = S_D(y)$.
"$\Leftarrow$", to show $C = D$, we can show that $C \subseteq D$ and $D \subseteq C$.
For any point $\bar{x} \in C$, if $\bar{x} \notin D$, by thm B.3.1, $\exists$ nonzero vector p, and a scalar α, s.t. $p^T \bar{x} > \alpha$ and $p^T x \leq \alpha, \forall x \in D$. However, $\forall y \in \mathbb{R}^n, C(y) = sup\{y^T x : x \in C\} = sup\{y^T x : x \in D\} = S_D(y)$. There is not such p. Thus, $\forall \bar{x} \in C, \bar{x} \in D$. That is, $C \subseteq D$. Similarly, we can show that $D \subseteq C$. Therefore, $D = C$

3.

$$\begin{aligned} L(x,u) &= f(x) + \textstyle\sum_{i=1}^m u_i g_i(x) \\ \bar{L}(x,\bar{u}) &= f(x) + \textstyle\sum_{i=1}^r \bar{u}_i g_i(x) \end{aligned}$$

$$\begin{aligned} L^*(u) &= \inf_{x \in X} L(x,u) \\ &= \inf_{x \in X} [f(x) + \textstyle\sum_{i=1}^m u_i g_i(x)] \\ \bar{L}^*(\bar{u}) &= \inf_{x \in \bar{X}} \bar{L}(x,\bar{u}) \\ &= \inf_{x \in \bar{X}} [f(x) + \textstyle\sum_{i=1}^r \bar{u}_i g_i(x)] \end{aligned}$$

Suppose $x^* \in X$ minimize $(P)(\bar{P})$. Thus, the duals are

$$\begin{aligned} v^* = \sup \quad & \inf_{x \in X}[f(x) + \textstyle\sum_{i=1}^m u_i g_i(x)] = f(x^*) + \textstyle\sum_{i=1}^m u_i g_i(^*) \\ \text{s.t.} \quad & u \in U \\ & u \geq 0 \\ \bar{v^*} = \sup \quad & \inf_{x \in \bar{X}}[f(x) + \textstyle\sum_{i=1}^r \bar{u}_i g_i(x)] = f(x^*) + \textstyle\sum_{i=1}^r \bar{u}_i g_i(x^*) \\ \text{s.t.} \quad & \bar{u} \in \bar{U} \\ & \bar{u} \geq 0 \end{aligned}$$

U and $\bar{U}$ can be derived from $L^*(u)$ and $\bar{L}^*(\bar{u})$. $U \subseteq \bar{U}$, since there are less constraints in $\bar{L}^*(\bar{u})$ than in $L^*(u)$. Since $u \geq 0$, $\bar{u} \geq 0$ and $\forall i \in \{1, ..., m\}, g_i(x) \leq 0$, we can have $u_i g_i(x) \leq 0$ and $\bar{u}_i g_i(x) \leq 0 \forall i \in \{1, ..., m\}$. Thus

$$\begin{aligned} v^* &= \sup_{u \in U, u \geq 0} [f(x^*) + \textstyle\sum_{i=1}^m u_i g_i(^*)] \\ &\leq \sup_{u \in U, u \geq 0} [f(x^*) + \textstyle\sum_{i=1}^r u_i g_i(^*)] \\ &\leq \sup_{u \in \bar{U}, u \geq 0} [f(x^*) + \textstyle\sum_{i=1}^r u_i g_i(^*)] \\ &= \bar{v^*} \end{aligned}$$

Thus, $v^* \leq \bar{v^*}$. By weak duality, $v^* \leq \bar{v^*} \leq f^*$.

4.a. For Approach 1

$$\begin{array}{ll} \inf & f_1(x) \\ \text{s.t.} & g_i(x) \leq 0, i = 1, ..., m \\ & f_j(x) \leq b_j, j = 2, ..., s \\ & x \in \mathbb{R}^n \end{array}$$

For Approach 2

$$\begin{array}{ll} \inf & f(x) = \sum_{j=1}^{s} w_j f_j(x) \\ \text{s.t.} & g_i(x) \leq 0, i = 1, ..., m \\ & x \in \mathbb{R}^n \end{array}$$

b. Let x^* be the solution obtained by Approach 2. Then for every solution obtained by Approach 2, we can set $b_j = f_j(x^*), j = 2, ..., s$

Then Approach 1 can be rewritten as

$$\begin{array}{ll} \inf & f_1(x) \\ \text{s.t.} & g_i(x) \leq 0, i = 1, ..., m \\ & f_j(x) \leq f_j(x^*), j = 2, ..., s \\ & x \in \mathbb{R}^n \end{array}$$

The objective function can be written as $w_1 f_1(x) + \sum_{j=1}^{s} w_j f_j(x^*)$ for some selection of weights $w_j, j = 1, ..., s$ with $w_1 > 0$, which will not change x value for optimal solution. Then x^* is the optimal solution to the new system.
c. Since a constraint qualification is satisfied for both problems, KKT conditions can serve as the first order necessary conditions.

$$\nabla f(\bar{x}) + \nabla g(\bar{x})^T u + \nabla h(\bar{x})^T v = 0$$

$$u \geq 0$$

$$u_i g_i(\bar{x}) = 0, i = 1, ..., m$$

For Approach 1,

$$\nabla f(\bar{x}) + \nabla g(\bar{x})^T u + \nabla h(\bar{x})^T v = \nabla f_1(\bar{x}) + \sum_{i=1}^{m} u_{gi} \nabla g_i(\bar{x}) + \sum_{j=2}^{s} u_{fj} \nabla f_j(\bar{x}) = 0$$

$$u_g \geq 0, u_f \geq 0$$

$$u_{gi} g_i(\bar{x}) = 0, i = 1, ..., m$$

$$u_{fj}(f_j(\bar{x}) - b_j) = 0, j = 2, ..., s$$

For Approach 2,

$$\nabla f(\bar{x}) + \nabla g(\bar{x})^T u + \nabla h(\bar{x})^T v = \sum_{j=1}^{s} w_j \nabla f_j(\bar{x}) + \sum_{i=1}^{m} u_{gi} \nabla g_i(\bar{x}) = 0$$

$$u_g \geq 0$$

$$u_{gi} g_i(\bar{x}) = 0, i = 1, ..., m$$

d. If all functions are convex and a constraint qualification is satisfied, KKT conditions are sufficient for optimality. For every solution obtained by Approach 1, let $\bar{x}$ be that solution. Then it should satisfy the KKT conditions for Approach 1. That is, we have

For Approach 1,

$$\nabla f(\bar{x}) + \nabla g(\bar{x})^T u + \nabla h(\bar{x})^T v = \nabla f_1(\bar{x}) + \sum_{i=1}^{m} u_{gi} \nabla g_i(\bar{x}) + \sum_{j=2}^{s} u_{fj} \nabla f_j(\bar{x}) = 0$$

$$u_g \geq 0, u_f \geq 0$$

$$u_{gi} g_i(\bar{x}) = 0, i = 1, ..., m$$

$$u_{fj}(f_j(\bar{x}) - b_j) = 0, j = 2, ..., s$$

.

$$\begin{aligned} \textstyle\sum_{i=1}^{m} u_{gi} \nabla g_i(\bar{x}) &= -(\nabla f_1(\bar{x}) + \textstyle\sum_{j=2}^{s} u_{fj} \nabla f_j(\bar{x}))(From) \\ &= -\textstyle\sum_{j=1}^{s} w_j \nabla f_j(\bar{x}) \end{aligned}$$

where $w_j = u_{fj}, j = 1, ..., s$ and $w_1 = 1 > 0$. Then we can find the same solution is optimal in Approach 2.

5. First show that the current $x_{ij}, (i,j) \in A$ that satisfies both laws (1) and (2) is a solution to the problem.

The current $x_{ij}, (i,j) \in A$ that satisfies (1), and then it is a feasible solution to the problem. Since $R_{ij} \geq 0, (i,j) \in A$, $f(x)$ is convex function. Also, $h(x)$ are convex functions. Thus the convexity condition is satisfied and KKT conditions are sufficient for optimality.

$$\nabla f(\bar{x}) + \nabla g(\bar{x})^T u + \nabla h(\bar{x})^T v = 0$$

That is, when taking the derivative w.r.t. each $x_{ij}, (i,j) \in A$

$$\forall x_{ij}, (i,j) \in A, R_{ij} x_{ij} - t_{ij} + v_j - v_i = 0$$

Then we can set the $v =$ voltage at each node and the current $x_{ij}, (i,j) \in A$ that satisfies both laws (1) and (2) also satisfies KKT conditions, which makes it to be optimal solution to (3).

Then show that the current $x_{ij}, (i,j) \in A$ that satisfies both laws (1) and (2) is the unique solution to the problem.

Let $n =$ number of element in set N and $a =$ number of element in set A. To solve the system, we have $n + a$ of equations as following

$$\forall x_{ij}, (i,j) \in A, R_{ij} x_{ij} - t_{ij} + v_j - v_i = 0$$

$$\forall i, i \in N, \sum_{j:(i,j) \in A} x_{ij} = \sum_{j:(j,i) \in A} x_{ij}$$

Within the system, we have $n + a$ unknowns, $x_{ij}, (i,j) \in A$ and $v_i, i \in N$. Based on the basic properties of electric network (e.g. currents are one-way), the equations are linear independent from each other. Thus, if the system can be solved, the solution is unique.