IOE 511 HWK10

Zhang, Haofan (hfzhang@umich.edu)

March 30, 2017

8.5

Question 1. (a) We have:

$$\|x - c\|_2^2 = \sum_{i=1}^{n}(x_i - c_i)^2$$

where $x = (x_1, \cdots, x_n)^T$ and $c = (c_1, \cdots, c_n)^T$. And we have:

$$\begin{aligned}
\|x - d\|_2^2 &= \sum_{i=1}^{n}(x_i - c_i - (\frac{1 - \sum_{j=1}^{n} c_j}{n}))^2 \\
&= \sum_{i=1}^{n}(x_i - c_i)^2 - 2\sum_{i=1}^{n}(x_i - c_i)(\frac{1 - \sum_{j=1}^{n} c_j}{n}) + \sum_{i=1}^{n}(\frac{1 - \sum_{j=1}^{n} c_j}{n})^2 \\
&= \sum_{i=1}^{n}(x_i - c_i)^2 - \frac{2}{n}(\sum_{i=1}^{n} c_i - 1)^2 + \frac{1}{n}(\sum_{i=1}^{n} c_i - 1)^2 \\
&= \sum_{i=1}^{n}(x_i - c_i)^2 - \frac{1}{n}(\sum_{i=1}^{n} c_i - 1)^2
\end{aligned}$$

Since $\frac{1}{n}(\sum_{i=1}^{n} c_i - 1)^2$ is a constant for a given c, so we can have that (P_c) and (P_d) have the same optimal solutions.

(b) Since $\|x - d\|_2^2 \geq 0$, so if we can show $e^T d = 1$, then $x^* = d$ must be an optimal solution of (P_d). As we have:

$$\begin{aligned}
e^T d &= e^T(c + (\frac{1 - e^T c}{n})e) \\
&= \sum_{i=1}^{n}[c_i + \frac{1 - \sum_{j=1}^{n} c_j}{n}] \\
&= \sum_{i=1}^{n} c_i + 1 - \sum_{j=1}^{n} c_j \\
&= 1
\end{aligned}$$

So we proved the result.

(c) We first calculate the derivative of funtions using in KKT condition:

$$\nabla f(x) = x - d,\ \nabla g(x) = -e,\ \nabla h(x) = e$$

So if we already have an optimal solution x^*, we will have $x^* - d - u + ve = 0$ with $u \geq 0$. Then by multiple e^T on both sides, we have $e^T x^* - e^T d - e^T u + nv = 0$. As from the constraint and (b), we have $e^T x^* = e^T d = 1$, so the above equation changes to $nv = e^T u$, which implies that $v \geq 0$.
Then, for some i which $d_i < 0$, the KKT condition will be $x_i^* - d_i - u_i + ve = 0$, which implies that $u + i = x_i^* - d_i + ve > 0$. By using $x_i^* u_i = 0$ in the KKT condition, we must have $x_i = 0$.

(d) The method will be:
1.Calculate $d = c + (\frac{1-e^T c}{n})e$
x* may be infeasible.
2.Check any $i \in [n]$, if $d_i < 0$, set $x_i^* = 0$, otherwise set $x_i^* = d_i$.
3.Let $x^* = (x_1, \cdots, x_n)^T$ formed by step2, this is an optimal solution for (P_c).
Since d is a n-dimension vector, so step2 needs at most n steps.

Question 2. We divide into two parts.
When $C = D$, it is obvious that $S_C(y) = S_D(y)$ for any $y \in \mathbb{R}^n$.
When $S_C(y) = S_D(y)$ for any $y \in \mathbb{R}^n$, we need to show that $C = D$. We will prove this result by prove $C \subseteq D$ and $D \subseteq C$.
We first show that $C \subseteq D$. Suppose there exists some $x' \in C$ but $x' \notin D$. Since D is closed, x' can be strictly separated from D, which means we can find some $a \neq 0$ and b, with $a^T x' > b$ and $a^T x < b$ for any $x \in D$. This means we can have:

$$\sup_{x \in D} a^T x \leq b < a^T x' \leq \sup_{x \in C} a^T x$$

which implies that $S_C(a) \neq S_D(a)$. So by contradiction, we can have $C \subseteq D$.
As C and D are symmetric, we can repeat the argument and get $D \subseteq C$, which implies that $C = D$.

Question 3. We have:

$$L(x, u) = f(x) + \sum_{i=1}^{m} u_i g_i(x),\ L^\star(u) = \inf_{x \in X} f(x) + \sum_{i=1}^{m} u_i g_i(x)$$

and

$$\overline{L}(x, u) = f(x) + \sum_{i=1}^{r} u_i g_i(x),\ \overline{L}^\star(u) = \inf_{x \in \overline{X}} f(x) + \sum_{i=1}^{r} u_i g_i(x)$$

Since $g_i(x) \leq 0$ for all $i \in [m]$, so we have

$$v^* = L^\star(u) = \inf_{x \in X} f(x) + \sum_{i=1}^{m} u_i g_i(x) \leq \inf_{x \in \overline{X}} f(x) + \sum_{i=1}^{r} u_i g_i(x) = \overline{L}^\star(u) = \overline{v}^*$$

As the two dual problem are formed from the same primal problem, so by apply weak duality, we have:

$$f^* \geq \overline{L}^*(u) = \overline{v}^*$$

which implies that $v^* \leq \overline{v}^* \leq f^*$.

Question 4. (a) For Approach 1, the problem is:

$$\begin{aligned} \text{minimize} \quad & f_1(x) \\ \text{subject to} \quad & f_j(x) \le b_j, j = 2, \cdots, s \\ & g_i(x) \le 0, i = 1, \cdots, m \\ & x \in \mathbb{R}^n \end{aligned}$$

For Approach 2, the problem is:

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{s} w_j f_j(x) \\ \text{subject to} \quad & g_i(x) \le 0, i = 1, \cdots, m \\ & x \in \mathbb{R}^n \end{aligned}$$

(b) For some selection of weight w_j, $j = 1, \cdots, s$, we can assume that x^* is one optimal solution obtained by Approach 2. We can set $b_j = f_j(x^*)$, $j = 2, \cdots, s$. Now we only need to prove that by using such b_j, x^* is an optimal solution for Approach 1.
We first show that x^* is feasible in Approach 1. As x^* is a feasible solution in Approach 2, we can have that $g_i(x^* \le 0$ for $i = 1, \cdots, m$. By the definition of b_j, we can have that $f_j(x^*) = b_j$ for $j = 2, \cdots, s$. So all constraints in Approach 1 is satisfied by x^*, which implies that it is a feasible solution for Approach 1.
Then we show it is actually an optimal solution for Approach 1. Assume there exists one $x^{'}$ which $f_1(x^{'}) < f_1(x^*)$. Due to the constraints in Approach 1, we can have that $f_j(x^{'}) \le b_j = f_j(x^*)$ for $j = 2, \cdots, s$. So, we have:

$$\sum_{j=1}^{s} w_j f_j(x^{'}) < \sum_{j=1}^{s} w_j f_j(x^*) \text{ with } g_i(x^{'}) \le 0 \text{ for } i = 1, \cdots, m$$

which implies that x^* is not an optimal solution for Approach 2, and this leads to a contradiction.

(c) For Approach 1, the first order necessary condition is:

$$\begin{gathered} \nabla f_1(x) + \sum_{j=2}^{s} \nabla v_j f_j(x)^T + \sum_{i=1}^{m} \nabla u_i g_i(x)^T = 0 \\ u_i g_i(x) = 0, \ i = 1, \cdots, m \\ v_j(f_j(x) - b_j) = 0, \ j = 2, \cdots, s \\ u \ge 0, \ v \ge 0 \end{gathered}$$

For Approach 2, the first order necessary condition is:

$$\begin{gathered} \sum_{j=1}^{s} w_j \nabla f_j(x) + \sum_{i=1}^{m} \nabla u_i g_i(x)^T = 0 \\ u_i g_i(x) = 0, \ i = 1, \cdots, m \\ u \ge 0 \end{gathered}$$

(d) As all functions are convex and continuous, $\mathbb{R}^n$ is a convex set, so both problems are convex problems. So we can have that the first order conditions we gave in (c) is necessary and sufficient conditions.
Assume x^* is an optimal solution attained by Approach 1 for selecting b_j, $j = 2, \cdots, s$. By the result of (c), we have:

$$\nabla f_1(x^*) + \sum_{j=2}^{s} \nabla v_j f_j(x^*)^T + \sum_{i=1}^{m} \nabla u_i g_i(x^*)^T = 0$$
$$u_i g_i(x^*) = 0,\ i = 1, \cdots, m$$
$$v_j(f_j(x^*) - b_j) = 0,\ j = 2, \cdots, s$$
$$u \geq 0,\ v \geq 0$$

have a solution v and u. So if we set $w_1 = 1$, $w_j = v_j$ for $j = 2, \cdots, s$, we can change the above result to:

$$\sum_{j=1}^{s} w_j \nabla f_j(x^*) + \sum_{i=1}^{m} \nabla u_i g_i(x^*)^T = 0$$
$$u_i g_i(x^*) = 0,\ i = 1, \cdots, m$$
$$u \geq 0$$

which is exactly the first order condition for Approach 2, which implies that x^* is also an optimal solution for Approach 2 if we set w_j as we used above.

Question 5. Let $f(x) = \sum_{(i,j)\in A}(\frac{1}{2}R_{ij}x_{ij}^2 - t_{ij}x_{ij})$, $h_i(x) = \sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(i,j)\in A} x_{ji}$, $i \in N$. As $f(x)$ and all $h_i(x)$ are convex functions, and $\mathbb{R}^n$ is a convex set, (3) is a convex problem. So KKT condition is a necessary and sufficient condition. We write it as:

$$\nabla f(x) + \sum_{i\in N} u_i \nabla h_i(x) = 0$$

For a selected $(i', j') \in A$, the corresponding row in $\nabla f(x)$ will be $R_{i'j'}x_{i'j'} - t_{i'j'}$, and the corresponding row in $\nabla h_{i'}(x)$ will be $u_{i'}$ wile the corresponding row in $\nabla h_{j'}(x)$ will be $-u_{j'}$(when having $j = i'$). So the equation in this selected (i', j') is:

$$R_{i'j'}x_{i'j'} - t_{i'j'} + u_{i'} - u_{j'} = 0$$

Use the same analysis, we can have for (j', i'), the equation is:

$$R_{j'i'}x_{j'i'} - t_{j'i'} + u_{j'} - u_{i'} = 0$$

which implies that by adding these two equations, we get:

$$(R_{j'i'}x_{j'i'} - t_{j'i'}) + (R_{i'j'}x_{i'j'} - t_{i'j'}) = 0$$

As (i', j') is arbitrary selected, we have that $\forall (i,j) \in A$:

$$(R_{ji}x_{ji} - t_{ji}) + (R_{ij}x_{ij} - t_{ij}) = 0$$

So, if there exists some solution x_0 doesn't satisfy (2), there must be at least one $(i_0, j_0) \in A$ where $(R_{j_0i_0}x_{j_0i_0} - t_{j_0i_0}) + (R_{i_0j_0}x_{i_0j_0} - t_{i_0j_0}) \neq 0$, which implies that x_0 can't satisfy KKT condition, so x_0 is not an optimal solution. And if not satisfy (1), it means this point even not in the domain of (3), which must not be an optimal solution.
So, the current x_{ij}, $(i,j) \in A$ that satisfies both law (1) and (2) is the unique solution of (3).