Homework 10

Luowei Zhou

IOE 511: Continuous Optimization Methods

March 31, 2017

10

Problem 1. (a) Replace d with $c + (\frac{1-e^Tc}{n})e$ in $||x-d||_2^2$:

$$\begin{aligned} ||x-d||_2^2 &= (x-d)^T(x-d) && (1)\\ &= x^Tx - 2d^Tx + d^Td && (2)\\ &= -2(c^T + \frac{1-e^Tc}{n}e^T)x + (c^T + \frac{1-e^Tc}{n}e^T)(c + \frac{1-e^Tc}{n}e) && (3)\\ &= (-2c^Tx + c^Tc) + \frac{(1-e^Tc)(e^Tc-1)}{n} && (4) \end{aligned}$$

Since $\frac{(1-e^Tc)(e^Tc-1)}{n}$ is a constant, optimize $\min \frac{1}{2}||x-d||_2^2 = \frac{1}{2}(||x-c||_2^2 + \frac{(1-e^Tc)(e^Tc-1)}{n})$ is equivalent to optimize $\min \frac{1}{2}||x-c||_2^2$. Also, (P_c) and (P_d) have same constraints. Hence, they have the same optimal solutions.

(b) We know that $||x-d||_2^2 \geq 0$. When $d \geq 0$, we have $e^Td = 1$ and $d \geq 0$, which implies d is in the feasible region. Also, $||d-d||_2^2 = 0$, d minimize (P_d). Hence, $x^* = d$ is an optimal solution of (P_d).

(c) Since the functions are convex and the constrains are linear functions, the KKT necessary conditions hold. Hence, we have

$$\begin{aligned} x^* - d - u + ve &= 0 && (5)\\ e^Tx^* &= 1 && (6)\\ u &\geq 0 && (7)\\ u_ix_i^* = 0 \quad i &= 1, ..., n && (8) \end{aligned}$$

Multiple both sides of Eq. (5) by e^T, we get $e^T u = ve^T e = nv$, which implies $v \geq 0$. When $d_j < 0$, $d_j = x_j^* - u_j + v < 0$. If $x_j > 0$, then $u_j > 0$, which is contradict with the condition $u_i x_i^* = 0$. Hence, $x_j = 0$ if $d_j < 0$.

(d) According to part (a), to can solve (P_c) by solving (P_d). If $d \geq 0$, then let $x^* = d$ and is the optimal solution to (P_c) according to part (b). If $d_j < 0$ for some index j (there might be multiple j), we can eliminate x_j in x and d_j in d, and form a new problem for the rest of x and d, denoted as $x^{(1)}$ and $c^{(1)}$:

$$(P_c^{(1)}) \quad \min \quad \frac{1}{2}||x^{(1)} - c^{(1)}|| \tag{9}$$

$$s.t. \quad e^T x^{(1)} = 1 \tag{10}$$

$$x^{(1)} \geq 0. \tag{11}$$

Then transfer $P_c^{(1)}$ to $P_d^{(1)}$ by taking the same transformation from c to d. If $d^{(1)} \geq 0$, then optimal solution is $x^{(1)*} = d^{(1)}$ and the rest terms are 0. If not, repeat this variable elimination process until we have $d^j \geq$ or $d^j < 0$. Since at each time step, we can at least eliminate one variable, the whole optimization takes at most n steps.

Problem 2. First we prove that, $C = D$ implies $S_C(y) = S_D(y)$. This is obvious according to the definition of $S_{(\cdot)}(y)$.

Then we show $S_C(y) = S_D(y)$ implies $C = D$. We prove this by contrapositive, i.e., if $C \neq D$, then $S_C(y) \neq S_D(y)$. Since, $C \neq D$, without loss of generality, we assume $\exists\, \bar{x} \in C$ but $\bar{x} \notin D$. According to Theorem B.3.1, we can strictly divide point $\bar{x}$ and D by a hyperplane. So, there exists a nonzero vector p and scalar α such that $p^T \bar{x} > \alpha$ and $p^T \bar{x} < \alpha$ for all $x \in S$. Thus, for this specific p, $S_D(p) < \alpha < p^T \bar{x} \leq S_C(p)$ so $S_D(p) \neq S_C(p)$. Hence, proved.

Problem 3. First, according to Weak Duality Theorem, $v^* \leq f^*$ and $\bar{v}^* \leq f^*$ hold. Next, we show that $v^* < \bar{v}^*$. The dual problems for (P) and $(\bar{P})$ are as follows.

$$(D) \quad L(u) = \min_x \quad f(x) + u^T g(x) \tag{12}$$

$$s.t. \quad x \in X \tag{13}$$

$$(\bar{D}) \quad \bar{L}(u) = \min_x \quad f(x) + \bar{u}^T \bar{g}(x) \tag{14}$$

$$s.t. \quad x \in \bar{X} \tag{15}$$

where $g(x) = (g_1(x)...g_m(x))^T$, $\bar{g}(g_1(x)...g_r(x))^T$ and $\bar{X}$ is as defined in the primal problem. Problem (D) can be rewritten as

$$(D) \quad L(u) = \min_x \quad f(x) + \bar{u}^T \bar{g}(x) + \tilde{u}^T \tilde{g}(x) \tag{16}$$

$$s.t. \quad x \in X \tag{17}$$

where $\tilde{g}(x) = (g_{r+1}(x)...g_m(x))^T$. So for arbitrary $\bar{u}$, problem (D) can be regarded as a dual problem of $(\bar{D})$, with $\tilde{x}$ as the slack variables. Hence, $L^*(\bar{u}, x) \leq \bar{L}^*(\bar{u}, x)$ for arbitrary $\bar{u}$. Furthermore, $v^* = \max_{\bar{u}} L^*(\bar{u}, x) = L^*(\bar{u}^*, x) \leq \bar{L}^*(\bar{u}^*, x) \leq \max_{\bar{u}} \bar{L}^*(\bar{u}, x) = \bar{v}^*$. Hence, proved.

Problem 4. (a) For Approach 1, the problem is formulated as

$$\min_x \quad f_1(x) \tag{18}$$

$$s.t. \quad f_j(x) \leq b_j \quad j = 2, 3, ..., s \tag{19}$$

$$g_i(x) \leq 0 \quad i = 1, 2, ..., m \tag{20}$$

For Approach 2, the problem is formulated as

$$\min_x \quad \sum_{j=1}^{s} w_j f_j(x) \tag{21}$$

$$s.t. \quad g_i(x) \leq 0. \tag{22}$$

(b) Let $\bar{x}$ be the optimal solution for Approach 2 and let $b_j = f_j(\bar{x}) \;\; j = 2, ..., s$. We prove by contradiction that $\bar{x}$ is the optimal solution for Approach 1. Assume $\bar{x}$ is not the optimal solution and we can find $\tilde{x}$ such that $f_1(\tilde{x}) < f_1(\bar{x})$ and it satisfy all the constraints. Then, owing to $w_1 > 0$, we have $\sum_{j=1}^{s} w_j f_j(\tilde{x}) < \sum_{j=1}^{s} w_j f_j(\bar{x})$ and also $\tilde{x}$ holds all the constraints of Approach 2. So, $\bar{x}$ is not optimal solution for Approach 2, contradict!

(c) Since all the constraints are satisfied, we skipped these conditions in the following formulations. The first order necessary condition for Approach 1 is:

$$f_1(x) + \sum_{j=2}^{s} u_j \nabla f_j(x) + \sum_{i=1}^{m} v_i \nabla g_i(x) = 0 \tag{23}$$

$$u_j \geq 0 \quad j = 2, 3, ..., s \tag{24}$$

$$u_j(f_j(x) - b_j) = 0 \quad j = 2, 3, ..., s \tag{25}$$

$$v_i \geq 0 \quad i = 1, 2, ..., m \tag{26}$$

$$v_i g_i(x) = 0 \quad i = 1, 2, ..., m \tag{27}$$

The condition for Approach 2 is:

$$\sum_{j=1}^{s} w_j \nabla f_j(x) + \sum_{i=1}^{m} v_i \nabla g_i(x) = 0 \tag{28}$$

$$v_i \geq 0 \quad i = 1, 2, ..., m \tag{29}$$

$$v_i g_i(x) = 0 \quad i = 1, 2, ..., m \tag{30}$$

(d) Assume x is the optimal solution for the problem in Approach 1, then x, u_j and v_i satisfy the first order necessary condition, as shown in Eq. (23)-(27). Let $w_j = \frac{u_j}{w_1}, j = 2, ..., s$ and $v_i' = \frac{v_i}{w_1}, i = 1, ..., m$ where $w_1 > 0$. Then Eq. (23)-(27) become:

$$w_1 f_1(x) + \sum_{j=2}^{s} w_j \nabla f_j(x) + \sum_{i=1}^{m} v_i' \nabla g_i(x) = 0 \tag{31}$$

$$w_j \geq 0 \quad j = 2, 3, ..., s \tag{32}$$

$$w_j(f_j(x) - b_j) = 0 \quad j = 2, 3, ..., s \tag{33}$$

$$v_i' \geq 0 \quad i = 1, 3, ..., m \tag{34}$$

$$v_i' g_i(x) = 0 \quad i = 1, 2, ..., m \tag{35}$$

Let v_j' be v_j and we notice that x is also a solution to Eq. (28)-(30), which is the KKT conditions for problem in Approach 2. Moreover, all functions are convex and a constraint qualification is satisfied, both of the problems are convex. Hence, the KKT sufficient condition holds and x is an optimal solution to the problem in Approach 2.

Problem 5. Assume the number of nodes are N. In this problem, $f(x) = \sum_{(i,j)\in A}(\frac{1}{2}R_{ij}x_{ij}^2 - t_{ij}x_{ij})$, $h_i(x) = \sum_{j:(i,j)\in A} x_{ji} - \sum_{j:(i,j)\in A} x_{ij}$ where $i = 1, 2, ..., N$ and x is the vector consists of

x_{ij} $(i,j) \in A$. Since $f(x)$ is a convex function regarding x_{ij} and $h_i(x)$ are linear functions, this optimization problem is a convex problem. We demonstrate that voltages of the nodes satisfy the KKT conditions for the nonlinear programming problem. Hence, if the currents satisfy the Kirchhoff's low and Ohm's law, they are the optimal solution to the target problem. Suppose for x_{ij} $(i,j) \in A$ together with u_1, u_2,...,u_N, the following condition hold:

$$\nabla f(x) + \sum_{i=1,\ldots,N} u_i \nabla h_1(x) = 0. \tag{36}$$

Plug in $f(x)$ and $h_i(x)$ and we have the following equations for each of x_{ij}

$$R_{ij}x_{ij} + (u_j - u_i) = 0. \tag{37}$$

Apparently, $u_i = v_i$ where $i = 1, 2, ..., N$ and x_{ij} following Kirchhoff's law and Ohm's law is a solution. Moreover, this solution is unique since the objective function is strict convex (Reason: Hessian matrix is diagonal with entries $R_{ij} > 0 \rightarrow$ positive definite).