IOE 511/Math 562
Homework 10

Yifan Zhou (yzhouum@umich.edu)

March 30, 2017

9

Question 1:

(a) Proof:

- Let $f_c(x) = \frac{1}{2}\|x-c\|_2^2 = \frac{1}{2}\sum_{i=1}^n (x_i - c_i)^2$ be the objective function of (P_c). Then the objective function of (P_d) can be expressed as $f_d(x) = \frac{1}{2}\sum_{i=1}^n (x_i - d_i)^2 = \frac{1}{2}\sum_{i=1}^n \left(x_i - c_i + \dfrac{1-e^T c}{n}\right)^2$.
- Define $\alpha = \dfrac{1-e^T c}{n} = \dfrac{1-\sum_{i=1}^n c_i}{n}$. Note that α is a scalar independent of variable x. By $e^T x = 1$, we know that for all x in the feasible region of (P_c) and (P_d), it must satisfy $\sum_{i=1}^n x_i = 1$.
- $f_d(x) = \frac{1}{2}\sum_{i=1}^n (x_i - c_i + \alpha)^2 = \frac{1}{2}\sum_{i=1}^n \left((x_i - c_i)^2 + \alpha^2 + 2\alpha(x_i - c_i)\right) = f_c(x) + \frac{n}{2}\alpha^2 + \alpha(1 - e^T c)$. In addition, ($\mathrm{P}_c$) and ($\mathrm{P}_d$) have the same feasible region. Therefore, any $\bar{x} \in \mathbb{R}^n$ in the feasible region of (P_c) that can minimize $f_c(x)$ will minimize $f_d(x)$ as well. In other words, they have the same optimal solutions.

(b) Proof: The feasible region of (P_d) is $\mathcal{F} = \{x \in \mathbb{R}^n : e^T x = 1, x \geq 0\}$. If $d \geq 0$, $e^T d = 1$, then $x^* = d$ is feasible. For any $x \in \mathcal{F}$, $f_d(x) \geq 0 = f_d(x^*)$. Therefore, x^* minimizes $f_d(x)$. $x^* = d$ is an optimal solution of (P_d).

You may want to argue that “x >= 0” is the domain of f_0

(c) Proof:

But the catch is that we need f_0 to be differentialbe in its domain. It implies we need an open domian. (Most of the case, we have domian R^n.)

- Let $h(x) = e^T x - 1$ and $X = \{x \in \mathbb{R}^n : x \geq 0\}$. Because $f_d(x)$ is a convex function, $h(x)$ is linear and X is a convex set, (P_d) is in fact a convex problem, where the first order KKT conditions are sufficient and necessary for global optimality.
- $\nabla f_d(x) = x - d = \begin{bmatrix} x_1 - d_1 & \dots & x_n - d_n \end{bmatrix}^T$, $\nabla h(x) = e = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}^T$. Let x^* be an optimal solution of (P_d). According to KKT conditions, it satisfies $\nabla f_d(x^*) + \nabla h(x^*) u = x^* - d + u \cdot e = 0$, where u is a scalar.

  Need to take care of X >= 0 as well. (miss a dual var v for x >= 0)
- Suppose $d_j < 0$ for some index j, but $x_j^* > 0$. Then we have $x_j^* = d_j - u > 0$. Because $e^T d = 1$ and $d_j < 0$, we have $1 = e^T x^* = e^T(d - u \cdot e) = e^T d - nu > e^T d - nd_j > 1$, which results in a contradiction. Therefore, for $d_j < 0$ for some index j, then $x_j^* = 0$ in any optimal solution of (P_d).

(d)
- Initialization: Set $k = 0$ and $c^k = c$. Define $S^0 = \{1, \dots, n\}$.
- Iteration:
  - Compute $d^k = c^k + \left(\dfrac{1 - e^T c^k}{n}\right)e$. Define $I^k = \{j \in S^k : d_j^k < 0\}$. If $I^k = \emptyset$, set $x_i^* = d_i^k$ for all $i \in S^k$. Stop the iteration and the current x^* is the optimal solution to (P_c).
  - Otherwise, set $x_j^* = 0$ for all $j \in I^k$. Set $k = k + 1$ and form a new vector c^k by eliminating all the components c_j^{k-1} for all $j \in I^{k-1}$ from c^{k-1}. Let $S^k = S^{k-1} \backslash I^{k-1}$ denote the indices of each component in c^k. Go back to the above step.

Question 2:

Proof:

- It is obvious that if $C = D$, then for any $y \in \mathbb{R}^n$, $S_C(y) = S_D(y)$.
- Let's show that if $S_C(y) = S_D(y)$, then $C = D$ by contradiction. Suppose $S_C(y) = S_D(y)$ but $C \neq D$. That is, there exists $\bar{x} \in C$ but $\bar{x} \notin D$. By definition, for any $y \in \mathbb{R}^n$, $y^T\bar{x} \leq S_C(y)$. For all $x \in D$, by the separation theorem, $y^T x > S_C(y) = S_D(y)$, which contradicts the assumption that $S_D(y) \geq y^T x$. Therefore, if $S_C(y) = S_D(y)$, then $C = D$.

Question 3:

Proof:

- $L(x,u) = f(x) + \sum_{i=1}^m u_i g_i(x)$, $L^*(u) = \inf_{x\in X} L(x,u)$, and the dual problem is

$$\begin{aligned} (D)\ \ v^* &= \sup L^*(u) \\ \text{s.t. } & u_i \geq 0, \text{for all } i = 1,\ldots,m. \end{aligned}$$

- $\bar{L}(x,u) = f(x) + \sum_{i=1}^r u_i g_i(x)$, $\bar{L}^*(u) = \inf_{x\in\bar{X}} \bar{L}(x,u)$, and the dual problem is

$$\begin{aligned} (\bar{D})\ \ \bar{v}^* &= \sup \bar{L}^*(u) \\ \text{s.t. } & u_i \geq 0, \text{for all } i = 1,\ldots,r. \end{aligned}$$

- Because the dual problems require all $u_i \geq 0$ for $i = 1,\ldots,m$, we can safely assume that in the following discussion without loss of generality. Because for all $x \in \bar{X}$, $g_i(x) \leq 0$, we have $u_i g_i(x) \leq 0$, $i = r+1,\ldots,m$. As a result, $\bar{L}(x,u) = f(x) + \sum_{i=1}^r u_i g_i(x) \geq f(x) + \sum_{i=1}^m u_i g_i(x)$ for the same $x \in \bar{X}$ and u_i, $i = 1,\ldots,r$. Let $\bar{x} \in \bar{X}$ be the solution that attains $\inf_{x\in\bar{X}} \bar{L}(x,u)$. When $u_i \geq 0$, $i = 1,\ldots,m$, we have $\inf_{x\in\bar{X}}(f(x) + \sum_{i=1}^m u_i g_i(x)) \leq f(\bar{x}) + \sum_{i=1}^m u_i g_i(\bar{x}) \leq f(\bar{x}) + \sum_{i=1}^r u_i g_i = \inf_{x\in\bar{X}} \bar{L}(x,u) = \bar{L}^*(u)$, for the same u_i, $i = 1,\ldots,r$. Therefore, $\sup_{u\geq 0} \bar{L}^*(u) \geq \sup_{u\geq 0} \inf_{x\in\bar{X}}(f(x) + \sum_{i=1}^m u_i g_i(x))$.
- $\inf_{x\in X} L(x,u) \leq \inf_{x\in\bar{X}}(f(x) + \sum_{i=1}^m u_i g_i(x))$ for the same u_i, $i = 1,\ldots,m$ because $\bar{X} \subseteq X$. It follows that $\sup_{u\geq 0} L^*(u) \leq \sup_{u\geq 0} \inf_{x\in\bar{X}}(f(x) + \sum_{i=1}^m u_i g_i(x))$.
- From above analysis, we know $\sup_{u\geq 0} \bar{L}^*(u) \geq \sup_{u\geq 0} \inf_{x\in\bar{X}}(f(x) + \sum_{i=1}^m u_i g_i(x)) \geq \sup_{u\geq 0} L^*(u)$. That is, $\bar{v}^* \geq v^*$. Due to the weak duality, $v^* \leq f^*$ and $\bar{v}^* \leq \bar{f}^*$. Since (P) and $(\bar{P})$ are essentially the same problem, $f^* = \bar{f}^*$. Therefore, we have shown that $v^* \leq \bar{v}^* \leq f^*$.

Question 4:

(a) Approach 1:

$$\begin{aligned} (P_1)\quad z_1^* &= \inf f_1(x) \\ \text{s.t.}\quad & g_i(x) \leq 0,\ \ i = 1,\ldots,m \\ & f_j(x) \leq b_j,\ \ j = 2,\ldots,s \\ & x \in \mathbb{R}^n \end{aligned}$$

Approach 2:

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

(b) Proof:

- Suppose $\bar{x}$ is the optimal solution obtained by Approach 2. For all x in the feasible region of (P), there is $z^* = w_1 f_1(\bar{x}) + \sum_{j=2}^{s} w_j f_j(\bar{x}) \le w_1 f_1(x) + \sum_{j=2}^{s} w_j f_j(x)$.
- Let $b_j = f_j(\bar{x})$ for all $j = 2, \ldots, s$. For all x in the feasible region of (P_1), $f_j(x) \le b_j$, $j = 2, \ldots, s$. It follows that $\sum_{j=2}^{s} w_j f_j(\bar{x}) \ge \sum_{j=2}^{s} w_j f_j(x)$ because $w_j \ge 0$, $j = 2, \ldots, s$.
- From above, it can be deduced that for all feasible x of (P_1), there is $f_1(\bar{x}) \le f_1(x)$ since $w_1 > 0$. In addition, $\bar{x}$ is feasible for (P_1) because it is feasible for (P) and it satisfies $f_j(\bar{x}) \le b_j$, $j = 2, \ldots, s$ as well. Therefore, $\bar{x}$ is also the optimal solution for Approach 1.

(c) Suppose both problems satisfy a constraint qualification for KKT first order necessary conditions for optimality.

- Approach 1: Let $\bar{x}$ be a feasible solution of (P_1). If $\bar{x}$ is a local optimum, $\exists (u, v)$, such that

$$\nabla f_1(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) + \sum_{j=2}^{s} v_j \nabla f_j(\bar{x}) = 0$$

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

$$v_j \ge 0, \;\; v_j f_j(\bar{x}) = 0, \;\; j = 2, \ldots, s$$

- Approach 2: Let $\hat{x}$ be a feasible solution of (P). If $\hat{x}$ is a local minimum, $\exists \hat{u}$, such that

$$\sum_{j=1}^{s} w_j \nabla f_j(\hat{x}) + \sum_{i=1}^{m} \hat{u}_i \nabla g_i(\hat{x}) = 0$$

$$\hat{u}_i \ge 0, \;\; \hat{u}_i g_i(\hat{x}) = 0, \;\; i = 1, \ldots, m$$

(d) Proof:

- Because all functions are convex and a constraint qualification is satisfied, KKT first order necessary conditions are sufficient and necessary for global optimality.
- Let $\bar{x}$ be a solution obtained by Approach 1. It must satisfy the optimality conditions in (c). Let $w_1 = 1 > 0$, $w_j = v_j \ge 0$ for $j = 2, \ldots, s$ and $\hat{u}_i = u_i$ for $i = 1, \ldots, m$. Then the optimality conditions for Approach 2 is in fact exactly the same as the ones for (P_1).
- $\bar{x} \in \mathbb{R}^n$ is in a feasible region of (P) since it satisfies $g_i(\bar{x}) \le 0$, $i = 1, \ldots, m$. And it also satisfy the optimality conditions for Approach 2. Therefore, Approach 2 can have the same solution as in Approach 1 with the weights we find for w_j, $j = 1, \ldots, s$.

Question 5:

Proof:

- Let x be a vector that consists of all x_{ij} for all $(i, j) \in A$. Let $f(x) = \sum_{(i,j)\in A} (\frac{1}{2} R_{ij} x_{ij}^2 - t_{ij} x_{ij})$ and
$$\nabla f(x) = \begin{bmatrix} \vdots \\ R_{ij} x_{ij} - t_{ij} \\ \vdots \end{bmatrix}, \forall (i, j) \in A.$$
- Let $h_i(x) = \sum_{j:(j,i)\in A} x_{ji} - \sum_{j:(i,j)\in A} x_{ij}$, $\forall i \in N$. $h(x)$ can be written as $h(x) = Bx$. For the k^{th} row of B, its entries are $\dfrac{\partial h_k(x)}{\partial x_{ij}} = \begin{cases} -1 & \text{if } i = k, \\ 1 & \text{if } j = k, \\ 0 & \text{otherwise,} \end{cases} \quad \forall (i, j) \in A, \forall k \in N.$

- It is obvious that f is a convex function and h_i, $i \in N$ are linear, so it is a convex problem where KKT first order conditions are sufficient and necessary for global optimality.
- Let $\bar{x}$ be a feasible solution of the problem, which means it satisfies Kirchhoff's law. It is an optimal solution if and only if $\nabla f(\bar{x}) + B^T v = 0$, where v is a vector with the same number of components as the nodes. That is, $R_{ij}\bar{x}_{ij} - t_{ij} + v_j - v_i = 0$, $\forall (i, j) \in A$. Rearranging the equations, we have $v_i - v_j = R_{ij}\bar{x}_{ij} - t_{ij}$, $\forall (i, j) \in A$, which is Ohm's law. Therefore, the current x_{ij}, $(i, j) \in A$ satisfying laws (1) and (2) is the unique solution of the problem.