1 Problem 1

7

Solution:

(a)

Let's denote the objective functions in (P_c) and (P_d) as

$$
\begin{aligned}
f_c(x) &= \frac{1}{2} \parallel x-c \parallel_2^2 = \frac{1}{2}\sum_{i=1}^{n}(x_i-c_i)^2 \\
&= \frac{1}{2}\sum_{i=1}^{n}(x_i^2-2c_ix_i+c_i^2) = \frac{1}{2}\sum_{i=1}^{n}c_i^2 + \boldsymbol{\frac{1}{2}\sum_{i=1}^{n}(x_i^2-2c_ix_i)} \\
f_d(x) &= \frac{1}{2} \parallel x-d \parallel_2^2 = \frac{1}{2}\sum_{i=1}^{n}(x_i-d_i)^2 \\
&= \frac{1}{2}\sum_{i=1}^{n}(x_i^2-2d_ix_i+d_i^2) \\
&= \frac{1}{2}\sum_{i=1}^{n}[x_i^2-2x_i(c_i+\frac{1-e^Tc}{n})] + \frac{1}{2}\sum_{i=1}^{n}[c_i^2+2c_i\frac{1-e^Tc}{n}+(\frac{1-e^Tc}{n})^2] \\
&= \frac{1}{2}\sum_{i=1}^{n}(x_i^2-2x_ic_i) - \frac{1-e^Tc}{2n}\sum_{i=1}^{n}x_i + \frac{1}{2}\sum_{i=1}^{n}[c_i^2+2c_i\frac{1-e^Tc}{n}+(\frac{1-e^Tc}{n})^2 \\
&= \boldsymbol{\frac{1}{2}\sum_{i=1}^{n}(x_i^2-2x_ic_i)} - \frac{1-e^Tc}{2n} + \frac{1}{2}\sum_{i=1}^{n}[c_i^2+2c_i\frac{1-e^Tc}{n}+(\frac{1-e^Tc}{n})^2
\end{aligned}
$$

It's obvious that both $f_c(x)$ and $f_d(x)$ can be denoted as a constant plus $\boldsymbol{\frac{1}{2}\sum_{i=1}^{n}(x_i^2-2c_ix_i)}$, that is to say, both (P_c) and (P_d) are trying to minimize $\frac{1}{2}\sum_{i=1}^{n}(x_i^2-2c_ix_i)$ with the exactly same constraints $e^Tx=1$ and $x \geq 0$.

Therefore, (P_c) and (P_d) have the same solution.

(b) and (c)

Let's write down the (P_c) as the following format first, we have

$$
\begin{aligned}
\min\ & f(x) \\
\text{s.t.}\ & h(x)=0 \\
& x \in X, \text{ where}
\end{aligned}
$$

$$
f(x) = \frac{1}{2}\sum_{i=1}^{n}(x_i-c_i)^2, h(x) = \sum_{i=1}^{n}x_i - 1, X = \{x_i \geq 0, i=1,...n\}
$$

Then we calculate the Lagrangian function as

$$
\begin{aligned}
L_c(x,u) &= f(x) + u_c h(x) = \frac{1}{2}\sum_{i=1}^{n}(x_i - c_i)^2 + u_c(\sum_{i=1}^{n} x_i - 1) \\
\frac{\partial L_c}{\partial x_i} &= (x_i - c_i) + u_c = 0, \text{ which makes } x_i = c_i - u_c \geq 0, i = 1, ...n \\
L_c^*(u_c) &= -u_c + \min_x \frac{1}{2}\sum_{i=1}^{n}(x_i - c_i)^2 + u\sum_{i=1}^{n} x_i, \text{ plug in } x_i \\
&= -u_c + \frac{1}{2}\sum_{i=1}^{n}(u_c)^2 + u_c\sum_{i=1}^{n}(c_i - u_c) \\
&= -\frac{nu_c^2}{2} + (\sum_{i=1}^{n} c_i - 1)u_c
\end{aligned}
$$

The dual prblem (D_c) can be written as

$$
\begin{aligned}
v_c^* = \max_u \ L_c^*(u_c) &= \max_{u_c} \ -\frac{nu_c^2}{2} + (\sum_{i=1}^{n} c_i - 1)u_c \\
s.t. \ u_c &\leq c_i, i = 1, ...n
\end{aligned}
$$

Similarily, we can apply exactly same procedure for (P_d), and replacing all u_c, c by u_d, d, we will get corresponding Lagrangian and dual problem (D_d) as

$$
L_d^*(u_d) = -\frac{nu_d^2}{2} + (\sum_{i=1}^{n} c_i - 1)u_d, \text{ with } x_i = d_i - u_d \geq 0, i = 1, ...n
$$

$$
\begin{aligned}
v_d^* = \max_u \ L_d^*(u_d) &= \max_{u_c} \ -\frac{nu_d^2}{2} + (\sum_{i=1}^{n} d_i - 1)u_d \\
&= \max_{u_c} \ -\frac{nu_d^2}{2}, \text{ since} \sum_{i=1}^{n} d_i = 1 \\
s.t. \ u_d &\leq d_i, i = 1, ...n
\end{aligned}
$$

The dual function is a negative quadratic function, that is to say, if $d \geq 0$, $u_d = 0$ can maximize $L_d^*(u_d)$, corresponding $x^* = d - u_d = d$. Otherwise, if $d_j = \min\{d_i, ...d_n\} < 0$ for some index j, then $u_d = d_j$ can maximize $L_d^*(u_d)$ and corresponding $x_j^* = d_j - u_d = 0$.

Didn't see why here

(d)

Step one: calculate $d = c + (\frac{1-e^T c}{n})e$

Incorrect alg

Step two: if $d \geq 0$, $x^* = d$

Step three: otherwise, if $d_j = \min\{d_i, ...d_n\} < 0$, let $u_d = d_j = \min\{d_i, ...d_n\}$, $x^* = d - u_d e$, with $x_j^* = 0$.

10

2 Problem 2

Proof:

Part I: C=D $\Rightarrow \forall \boldsymbol{y} \in \boldsymbol{R^n}, \boldsymbol{S_C(y) = S_D(y)}$

We know that the support function for set C and D can be written as

$$S_C(y) = \sup\{y^T x : x \in C\}$$

$$S_D(y) = \sup\{y^T x : x \in D\}$$

$\forall y \in R^n$, let $S_C(y) = y^T x_C^*$ where $x_C^* \in C$. Since C=D, $x_C^* \in D$, then $S_D(y) = y^T x_D^* \geq y^T x_C^* = S_C(y)$ holds for sure.

Assume $S_C(y) \neq S_D(y)$, then $S_D(y) > S_C(y)$ at $x_D^* \in D$. However, we know that $C = D$, $x_D^* \in C$, $S_C(y) \geq y^T x_D = S_D(y)$. Contradiction.

Therefore, $C = D \Rightarrow \forall y \in R^n, S_C(y) = S_D(y)$.

Part II: $\forall \boldsymbol{y} \in \boldsymbol{R^n}, \boldsymbol{S_C(y) = S_D(y)} \Rightarrow$ C=D

Let's prove it by contradiction. We want to prove $\forall y \in R^n, S_C(y) = S_D(y) \Rightarrow C = D$. Suppose $C \neq D$ holds for true, that is to say, $\exists x_C \in C$ such that $x_C \notin D$ or $\exists x_D \in D$ such that $x_D \notin C$.

If $\exists x_C \in C$ such that $x_C \notin D$ holds for true, according to Theorem B.3.1, if D is a closed and convex set and $x_C \notin D$, then there exists a nonzero vector y and scalar α such that

$$y^T x_C > \alpha \text{ and } y^T x_D < \alpha, \forall x_D \in D$$

It's obvious that if $y^T x_D < \alpha, \forall x_D \in D$, then

$$S_D(y) = \sup\{y^T x : x \in D\} \leq \alpha$$

According to the definition of support function and our assumption $S_C(y) = S_D(y), \forall y \in R^n$, we have

$$\boldsymbol{S_D(y) = S_C(y) \geq y^T x_C > \alpha} \text{ and } \boldsymbol{S_D(y) \leq \alpha}$$

Contradiction. Similarly, if we assume that $\exists x_D \in D$ such that $x_D \notin C$, we can also derive the contradiction following the same method.

Therefore, $\forall y \in R^n, S_C(y) = S_D(y) \Rightarrow C = D$.

In conclusion, combine the result from part I and part II, $C = D \; iff text \forall y \in R^n, S_C(y) = S_D(y)$.

3 Problem 3

Proof:

According to the Theorem 7.6.1 (week duality theorem), we have $\boldsymbol{f^* \geq v^*}$ and $\boldsymbol{f^* \geq \bar{v}^*}$ hold for true. Let's wirte down the primal problem (P) and modified primial problem $(\bar{P})$ as follows,

$$
\begin{aligned}
(P) \;\; \min = & f(x) \\
\text{s.t. } & g_i(x) \leq 0, i = 1, ...m \\
& x \in X \\
(\bar{P}) \;\; \min = & f(x) \\
\text{s.t. } & g_i(x) \leq 0, i = 1, ...r \\
& x \in \bar{X}, \text{ where } \bar{X} = \{x \in X : g_{r+1} \leq 0, ...g_m(x) \leq 0\}
\end{aligned}
$$

Corresponding Lagrangain for $(\bar{P})$ and (P) can be written as

$$
\begin{aligned}
\bar{L}(x, u) &= f(x) + \sum_{i=1}^{r} u_i g_i(x) \\
\bar{L}^*(u) &= \min_{x \in \bar{X}} \bar{L}(x, u) = \min_{x \in \bar{X}} f(x) + \sum_{i=1}^{r} u_i g_i(x) \\
(\bar{D})\text{: } \bar{v}^* &= \max_{u \geq 0} \bar{L}^*(u) \\
L(x, u) &= f(x) + \sum_{i=1}^{m} u_i g_i(x) = f(x) + \sum_{i=1}^{r} u_i g_i(x) + \sum_{i=r+1}^{m} u_i g_i(x) \\
&= \bar{L}(x, u) + \sum_{i=r+1}^{m} u_i g_i(x) \\
&\leq \bar{L}(x, u), \text{ since } u_i \geq 0, g_i(x) \leq 0, i = r + 1, ...m \\
L^*(u) &= \min_{x \in X} L(x, u) \leq \min_{x \in X} \bar{L}(x, u) = \bar{L}^*(u) \\
\boldsymbol{(D)\text{: } v^*} &\boldsymbol{= \max_{u \geq 0} L^*(u) \leq \max_{u \geq 0} \bar{L}^*(u) = \bar{v}^*}
\end{aligned}
$$

Therefore, we conclude that $\boldsymbol{v^* \leq \bar{v}^* \leq f^*}$.

4 Problem 4

Solution

(a)

Approach 1

$$
\begin{aligned}
\min_{x} \ & f_1(x) \\
s.t. \ & f_j(x) \le b_j, j = 2, ...s \\
& g_i(x) \le 0, i = 1, ...m
\end{aligned} \tag{1}
$$

Approach 2

$$
\begin{aligned}
\min_{x} \ & \sum_{j=1}^{s} w_j f_j(x) \\
s.t. \ & g_i(x) \le 0, i = 1, ...m
\end{aligned} \tag{2}
$$

(b)

Let's denote the optimal value obtained from (2) is x^*, let's also assume that it is not an optimal solution for (1). Instead, the optimal solution for (1) is $\bar{x}$. Then we have

$$f_1(\bar{x}) < f_1(x^*)$$

$$\boldsymbol{\sum_{j=1}^{s} w_j f_j(x^*) \le \sum_{j=1}^{s} w_j f_j(\bar{x})}$$

Simply let $b_j = f_j(x^*)$ in (1), then $f_j(\bar{x}) \le f_j(x^*) = b_j \ \forall j = 2, ...s$, x^* will remain in feasible region for (1), that is to say,

$$f_1(\bar{x}) < f_1(x^*)$$

$$f_j(\bar{x}) \le f_j(x^*), j = 2, ...s$$

Since $w_j \ge 0$, for j=2,...s and $w_1 > 0$, we have

$$\boldsymbol{\sum_{j=1}^{s} w_j f_j(\bar{x}) < \sum_{j=1}^{s} w_j f_j(x^*)}$$

Then x^* is not an optimal solution for (2), contradiction. Therefore, for every solution obtained by Approach 2 for some selection of weights $w_j, j = 1, ...s$ with $w_1 > 0$, we can find target levels $b_j, j = 2, ...s$ such that the same solution is optimal in Approach 1.

(c)

First order necessary condition for optimal solution x of problem (1) and (2) are displayed as below,

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

$$u_j f_j(x) = 0, u_j \geq 0, j = 2, ...s, \text{ and } v_i g_i(x) = 0, v_i \geq 0, i = 1, ...m$$

and

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

$$v_i g_i(x) = 0, v_i \geq 0, i = 1, ...m$$

(d)

If all functions are convex and contraint qualification is satisfied, then both problem (1) and (2) are convex problem. According to Theorem 3.2.1, first order KKT condition is the sufficient condition for global optimal solution. KKT condition of problem (1) is displayed as below.

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

$$u_j f_j(x) = 0, u_j \geq 0, j = 2, ...s$$

$$v_i g_i(x) = 0, v_i \geq 0, i = 1, ...m$$

For every solution $\bar{x}$ obtained by Approach 1 with given $b_j, j = 2, ...s$, we can figure out corresponding $u_j, j = 2, ...s$ such that the above KKT condition is held. Also, since $g_i(\bar{x}) \leq 0, i = 1, ...m$, $\bar{x}$ is also feasible for problem (2).

Furthermore, KKT condition for problem (2) is written as

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

$$v_i g_i(x) = 0, v_i \geq 0, i = 1, ...m$$

Since $w_1 > 0, w_j, j = 2, ...s \geq 0$, we can re-write it as,

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

$$v_i' = \frac{v_i}{w_1}, v_i' g_i(x) = 0, v_i' \geq 0, i = 1, ...m$$

If we assign $w_1 = 1$, $\frac{w_j}{w_1} = u_j$ for $j = 2, ...s$, the feasible solution $\bar{x}$ and corresponding $v_i' = v_i w_1 = v_1$ satisfy the KKT condition for problem (2). That is to say, optimal solution for problem (1) $\bar{x}$ is also an optimal solution for problem (2) when $w_1 = 1$ and $w_j = u_j w_1 = u_j$ for $j = 2, ...s$.

Therefore, if all functions are convex and a constraint qualification is satisfied, for every solution obtained by Approach 1 for some selection of target levels $b_j, j = 2, ...s$, we can find weights $w_j, j = 1, ...s$ with $w_1 > 0$ such that the same solution is optimal in Approach 2.

5 Problem 5

Proof:
We want to show that the current $x_{ij}, (i,j) \in A$ that satisfies both (1) and (2) is the unque solution of the following nonliear programming problem:

$$(P)\ \min f$$
$$\text{s.t. } h_i = 0, i \in N, \text{ where}$$

$$f = \sum_{(i,j)\in A} (\frac{1}{2} R_{ij} x_{ij}^2) - t_{ij} x_{ij}$$

$$h_i = \sum_{(i,j)\in A} x_{ij} - \sum_{(j,i)\in A} x_{ji}, i \in N$$

We can calculate the first order derivative:

$$\nabla f = \begin{bmatrix} ... \\ R_{ij} x_{ij} - t_{ij} \\ ... \end{bmatrix}, \forall (i,j) \in A$$

$$\nabla h_i = [h_{ij}], h_{ij} = 1 \text{ if } (i,j) \in A \text{ and } h_{ij} = -1 \text{ if } (j,i) \in A, \forall i \in N$$

We know that $R_{ij} > 0, (i,j) \in A$, Hessian matrix of f is $f'' = diag(\{R_{ij}\}), (i,j) \in A$, it is a diagonal matrix with diagonal elements equal to $R_{ij} > 0$. Hessian matrix of f is a SPD matrix, therefore f is a convex function. $h_i, i \in N$ are linear function of $x_{ij}, (i,j) \in A$. We also know that there is no nonnegativity constraints on $x_{ij}, (i,j) \in A$, we can conclude that (P) is a convex problem.

According to Theorem 3.2.1, the following condition is the sufficient condition for global optimal solution of convex problem,

$$\nabla f + \sum_{i=1}^{N} u_i \nabla h_i = 0, \text{ which can be written explicitly as}$$

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

If current $x_{ij}, (i,j) \in A$ satifies Kirchhoff's law (1), then it automatically satisfies the feasible constraints of (P), it is a feasible solution.

If current $x_{ij}, (i,j) \in A$ further satisies Ohm's law (2), then it solves the KKT condition with $u_i = -v_i, i \in N$, it is a global optimal solution of (P).

Combine (1) and (2), both the number of linear equations and variables (x_{ij}, u_i) equal to number of nodes plus arcs. Then there is should be at most one solution to this system of linear equations.

If we are given the current $x_{ij}, (i,j) \in A$, and calculate corresponding $u_i, i \in N$, then the set of solution x_{ij}, u_i is unique.

In conclusion, we can claim that the current $x_{ij}, (i,j) \in A$ that satisfies both law (1) and (2) is the unique solution of the nonlinear programming problem (P).