IOE 511 HW X

Cheng Sun

March 31, 2017

10

1 Question 1

(a) Because any feasible solution x for (P_c) and (P_d) satisfying that $e^Tx = 1$, we can rewrite $d = c + \frac{e^T(x-c)}{n}e$, then the objective function of (P_d) can be written as $\frac{1}{2}\|x - c + e^T(x - c)e/n\|_2$. If we can prove that $x - c + e^T(x - c)e/n$ have the norm proportion to $x - c$, then we are done.
For the simplicity of notation, let $a = x - c$, then the objective function of (P_c) can be written as $\frac{1}{2}a^Ta$, and the objective function of (P_d) can be written as $\frac{1}{2}(a - \frac{e^Tae}{n})^T(a - \frac{e^Tae}{n}) = \frac{1}{2}(a^T - \frac{(e^Ta)e^T}{n})(a - \frac{e^Tae}{n})$
$= \frac{1}{2}(a^Ta - \frac{2(e^Ta)^2}{n} + \frac{(e^Ta)^2(e^Te)}{n^2}) = \frac{1}{2}(a^Ta - \frac{2(e^Ta)^2}{n})$. Note that for every feasible x, $e^Ta = 1 - e^Tc$ is a constant. So the goal of (P_d) is also minimize $\frac{1}{2}a^Ta$.
Given that two problems have same constraints, which means the feasible region is same. We can be sure that optimal solutions for (P_c) and (P_d) are the same.

(b) Since the objective value is non-negative, we can show that d is optimal solution by showing that d is feasible (i) and takes value 0 in objective function(ii).
It is clear that (ii) stands. If (i) is true, then the proof is done.
Since $d \geq 0$, the second constraint fits. As for the first constraint, $e^Td = e^T(c + \frac{1-e^Tc}{n}e) = e^Tc + \frac{e^Te - e^T(e^Tce)}{n} = 1$, for $e \in \mathbb{R}$, $e^Te = n$.

(c) According to Theorem 3.1.8, if all constraints are linear, the KKT conditions are necessary to characterize a local optimal solution. Constraints in this problem are both linear, we can apply KKT condition to optimal solution $\bar{x}$.
$\nabla f(\bar{x}) = \bar{x} - d$, $\nabla h(\bar{x}) = e$, $\nabla g(\bar{x}) = -e$.
From KKT, we have $\bar{x} - d + ve - u = 0$, with $\bar{x} \geq 0, u \geq 0$. Times e^T on both sides we have $e^T\bar{x} - e^Td + vn - e^Tu = 0$, for $e^T\bar{x} = 1$ (constraint) $e^Td = 1$ (proved in (b)), we have $vn - e^Tu = 0$, this leads to $v \geq 0$.
From KKT, we have $\bar{x}_j - d_j + v - u_j = 0$ and given $d_j < 0, v > 0, \bar{x}_j] \geq 0$, we know that $u_j = \bar{x}_j - d_j + v > 0$. From KKT, we have $\bar{x}_ju_j = 0$, then there is $\bar{x}_j = 0$

(d) We can run such algorithm in following way.
Note that in each iteration, if the algorithm doesn't stop, there will be at least one component fixed to be 0. So there will be at most n steps. So after at most n iteration, the algorithm will stop.

Initialization	$x = 0, t = e$
Iteration	For $k = 1, \cdots, n$, $d_k = t_k(c_k + \frac{1-t^t c}{t^T t})$.
	The subscript here refers to the position of component in a vector.
	If $d \geq 0$, set such that $x_k = d_k \times t_k$, and stop
	Else, for any k such that $d_k < 0$, set $t_k = 0$

2 Question 2

- C=D to equal function
  It is quite straight forward. We can prove this by contradiction, without loss of generality, let assume that there exist a y such that $S_C(y) > S_D(y)$, then it violates the definition of sup function, so $S_C(y) = S_D(y), \forall y \in \mathbb{R}^n$.
- equal function to C=D
  Without loss of generality, let assume that there exist a point a such that $a \in C, a \notin D$. Because that C and D are all convex set, there is a hyperplane strictly divide a and D, such that $y'^T a > 0$ $y'^T b < 0, \forall b \in D$. Then we know $S_C(y') > S_D(y')$, which contradicts to the condition that their support functions are equal.

3 Question 3

We form the Lagrangian with r constraints:

$$\bar{L}(x, u) = f(x) + u'^T g'(x)$$

Then the dual function $\bar{L}^\star(u')$ is:

$$\bar{L}^\star(u') = \inf f(x) + u'^T g'(x)$$

$$s.t.\ x \in \bar{X}$$

The dual problem is then defined to be:

$$\bar{D} : \bar{v}^\star = \sup \inf f(x) + u'^T g'(x)$$

$$s.t.\ u' \geq 0$$

We form the Lagrangian with m constraints:

$$L(x, u) = f(x) + u^T g(x)$$

Then the dual function $L^\star(u)$ is:

$$L^\star(u) = \inf f(x) + u^T g(x)$$

$$s.t.\ x \in X$$

The dual problem is then defined to be:

$$D : v^\star = \sup \inf f(x) + u^T g(x)$$

$$s.t.\ u \geq 0$$

Note that $u' \in \mathbb{R}^r, u \in \mathbb{R}^m, 1 \leq r < m$. $g'(\cdot)$ corresponds to first r constraints, and $g(\cdot)$ corresponds to first m constraints.
Given the primal problem is feasible and not unbounded. We know that the dual problems won't be unbounded. We can write $f(x) + u'^T g'(x) \geq f(x) + u'^T g'(x) + u''^T g''(x) = f(x) + u^T g(x) \geq \inf_{x \in X} f(x) + u^T g(x) = L^\star(u)$, where $g''(\cdot)$ corresponds to the r+1 to m constraints, and $u'' \in \mathbb{R}^{m-r}$.
The first inequality stands for $u'' \geq 0, g''(\cdot) \leq 0, \bar{X} \subseteq X$, the second inequality comes from the definition of inf. Therefore, $\bar{v}^\star \geq v^\star$.
As for $f^\star \geq \bar{v}^\star$, it is based on weak duality. $f(x) \geq f(x) + u'^T g'(x) \geq \inf_{x \in \bar{X}} f(x) + u^T g(x) = \bar{L}^\star(u)$.

4 Question 4

(a) the corresponding nonlinear optimization problem are written as follow.

Approach 1	min	$f_1(x)$	
	s.t.	$g_i(x) \leq 0$	$i = 1, \cdots, m$
		$f_j(x) - b_j \leq 0$	$j = 2, \cdots, s$
		$x \in \mathbb{R}^n$	
Approach 2	min	$\sum_{j=1}^{s} w_j f_j(x)$	
	s.t.	$g_i(x) \leq 0$	$i = 1, \cdots, m$
		$x \in \mathbb{R}^n$	

(b) We can show that the same solution is optimal in Approach 1 with certain target levels by proving those two things. (1) the solution obtained from approach 2 is feasible in approach 1; (2) there is no better solution in Approach 1 that gives better value than the solution obtained from approach 2.
Given a selection of weights $w \in \mathbb{R}^s \geq 0$, with $w_1 > 0$, let x_0 represents the optimal solution we obtained by Approach 2. We can then set the target levels in Approach 1 to be $b_j = f_j(x_0), j = 2, \cdots, s$.

- Feasibility
  Since x is the optimal value in Approach 2, we have $g_i(x_0) \leq 0, i = 1, \cdots, m$.
  As for $f_j(x) \leq b_j, j = 2, \cdots, s$, we know that $f_j(x_0) = b_j, j = 2, \cdots, s$, then we know all the constraints are satisfied. x_0 is a feasible solution in Approach 1.
- Optimality
  We can show this by contradiction. Let's assume that there is a feasible solution x_1 in Approach 1 that gives lower value. It means that $g_i(x_1) \leq 0, i = 1, \cdots, m$, $f_j(x_1) \leq b_j, j = 2, \cdots, s$, and $f_1(x_1) < f_1(x_0)$.
  For $g_i(x_1) \leq 0, i = 1, \cdots, m$, we know that x_1 is also a feasible solution for Approach 2. Given $w \in \mathbb{R}^s \geq 0$, with $w_1 > 0$, we can find that $\sum_{j=2}^{s} w_j f_j(x_1) \leq \sum_{j=2}^{s} w_j f_j(x_0)$,

so $\sum_{j=1}^{s} w_j f_j(x_1) < \sum_{j=1}^{s} w_j f_j(x_0)$. It leads to that x_0 is not the optimal solution for Approach 2. This is a contradiction, which means there is no such x_1. Then x_0 is the optimal solution for Approach 1.

(c) Given that a constraint qualification is satisfied for both problems, we can then claim that the KKT conditions are necessary to characterize a local optimal solution for both problems.

- Approach 1
  $\nabla f_1(x) + \nabla g_(x)^T u_1 + \nabla f'(x)^T u_2 = 0$, where $f'(x) = [f_2(x) \cdots f_j(x)]^T$, we can rewrite the equation as $\nabla f(x)^T u' + \nabla g_(x)^T u = 0$
  $u \geq 0$
  $u' \geq 0, u'_1 = 1$
  $u_i g_i = 0, i = 1, \cdots, m$
  $u'_j(f_j(x) - b_j) = 0, j = 2, \cdots, s$

- Approach 2
  $\sum_{j=1}^{s} w_j \nabla f_j(x) + \nabla g_(x)^T u = 0$
  $u \geq 0$
  $u_i g_i = 0, i = 1, \cdots, m$

Given that $w \in \mathbb{R}^s \geq 0$, with $w_1 > 0$, we know those two condition will be equivalent if we add constraint $w_j(f_j(x) - b_j) = 0, j = 2, \cdots, s$ to the second one. Then the u in the first KKT conditions can be $\frac{u}{w_1}$ in the second KKT conditions. And u' in the first KKT conditions is $[1, \frac{w_2}{w_1}, \cdots, \frac{w_s}{w_1}]^T$.
Thus the first order necessary conditions for optimality for both problems could be

$$\begin{aligned} \sum_{j=1}^{s} w_j \nabla f_j(x) + \nabla g_(x)^T u = 0 & \\ u \geq 0 & \\ u_i g_i = 0 & \qquad i = 1, \cdots, m \\ w_j(f_j(x) - b_j) = 0 & \qquad j = 2, \cdots, s \end{aligned}$$

(d) Since a constraint qualification is satisfied, the optimal solution x in Approach 1 for some selection of target levels $b_j, j = 2, \cdots, s$, x should satisfy KKT conditions. So we have

$$\begin{gathered} \nabla f(x)^T u' + \nabla g_(x)^T u = 0 \\ u \geq 0 \\ u' \geq 0, u'_1 = 1 \\ u_i g_i = 0, i = 1, \cdots, m \\ u'_j(f_j(x) - b_j) = 0, j = 2, \cdots, s \end{gathered}$$

We can set weights in Approach 2 based on solution u and u' on the above equation system, $w = u'$. If first problem satisfies a constraint qualification, the constraints in second problem is a subset of first problem, then the second problem satisfies the same constraint qualification. So we know that optimal solution $\bar{x}$ for second problem should also satisfy KKT conditions.

$$\sum_{j=1}^{s} w_j \nabla f_j(\bar{x}) + \nabla g_(\bar{x})^T u'' = 0$$
$$u'' \geq 0$$
$$u_i'' g_i = 0, i = 1, \cdots, m$$

Since $w = u'$, we can set $u'' = u$, we will find $\bar{x} = x$, satisfying the equation system. Given that all functions are convex, the linear combination of convex functions is still convex. Then the objective function and inequality constraints are all convex functions, $\mathbb{R}^n$ is a convex set. The second problem is a convex problem. According to Theorem 3.2.1 (KKT Sufficient Conditions for Convex Problems) in the textbook, we know that x satisfying KKT condition is the global optimal solution for second problem.

5 Question 5

Note that all the constraints in the nonlinear programming problem are linear functions. Based on Theorem 3.1.8 (Linear constraints), if all constraints are linear, the KKT conditions are necessary to characterize a local optimal solution. We can find the candidates for the optimal solution by KKT condition.

For every edge $(i,j) \in A$, $\frac{\partial f(x)}{\partial x_{ij}} = R_{ij}x_{ij} - t_{ij}$, $\frac{\partial h_i(x)}{\partial x_{ij}} = \begin{cases} 1 & k = i \\ -1 & k = j \\ 0 & otherwise \end{cases}$.

$\nabla f(x) + \sum_{i \in V} v_i \nabla h_i(x) = 0$, this equation holds for every component, so there is $\frac{\partial f(x)}{\partial x_{ij}} + \sum_{k \in N} v_k \frac{\partial h_k(x)}{\partial x_{ij}} = R_{ij}x_{ij} - t_{ij} + v_i - v_j = 0$

Let $(\bar{x}, \bar{v})$ be the solution that satisfies both laws (1) and (2). so we have $\sum_{j:(i,j) \in A} \bar{x}_{ij} = \sum_{j:(j,i) \in A} \bar{x}_{ij}, \forall i \in N$ and $\bar{v}_i - \bar{v}_j = R_{ij}\bar{x}_{ij} - t_{ij}$.

It is clear that if we set $x = \bar{x}$, and $v = -\bar{v}$, we can find that $\bar{x}$ satisfies the KKT conditions. And $\bar{x}$ satisfies the constraint in the problem, so it is a feasible solution.

The Hessian of the objective function is a matrix whose components on the diagonal $\frac{\partial^2 f(x)}{\partial x_{ij}^2} = R_{ij} > 0$, and the rest components are all 0, so the Hessian matrix is positive definite, the objective function is strictly convex. Then we know the solution is unique. Not to mention that KKT is also the sufficient condition for finding the optimal solution. Then we are sure the unique optimal solution for the problem is $\bar{x}$.