7

1 Question 1

1.1 Part (a)

Considering the problem where $e \in \mathbb{R}^n$ is a vector of all ones

$$\begin{aligned} (P_c) \quad \min_x \ & \frac{1}{2}||x-c||_2^2 \\ \text{s.t. } & e^\intercal x = 1 \\ & x \geq 0 \end{aligned} \tag{1}$$

Our objective function can also be seen as $f_c(x) = \frac{1}{2}(x-c)^\intercal(x-c)$.

Now lets look at the problem:

$$\begin{aligned} (P_d) \quad \min_x \ & \frac{1}{2}||x-d||_2^2 \\ \text{s.t. } & e^\intercal x = 1 \\ & x \geq 0 \end{aligned} \tag{2}$$

where $d = c + (\frac{1-e^\intercal c}{n})e$. This problem is very similar to Equation 1, except that the objective function here is defined as $f_d(x) = \frac{1}{2}(x-d)^\intercal(x-d)$. Substituting d into this and expanding we can see:

$$f_d(x) = \frac{1}{2}(x - (c + (\frac{1-e^\intercal c}{n})e))^\intercal(x - (c + (\frac{1-e^\intercal c}{n})e))$$

$$f_d(x) = \frac{1}{2}(x - c - (\frac{1-e^\intercal c}{n})e)^\intercal(x - c - (\frac{1-e^\intercal c}{n})e)$$

$$f_d(x) = \frac{1}{2}(x^\intercal - c^\intercal - (\frac{1-e^\intercal c}{n})e^\intercal)(x - c - (\frac{1-e^\intercal c}{n})e)$$

$$f_d(x) = \frac{1}{2}[(x^\intercal - c^\intercal) - ((\frac{1-e^\intercal c}{n})e^\intercal)][(x-c) - ((\frac{1-e^\intercal c}{n})e)]$$

$$f_d(x) = \frac{1}{2}[(x^\intercal - c^\intercal)(x-c) - 2((\frac{1-e^\intercal c}{n})e^\intercal)(x-c) + ((\frac{1-e^\intercal c}{n})^2 e^\intercal e)]$$

$$f_d(x) = \frac{1}{2}(x-c)^\intercal(x-c) - (\frac{1-e^\intercal c}{n})e^\intercal(x-c) + \frac{1}{2}((\frac{1-e^\intercal c}{n})^2 e^\intercal e)$$

$$f_d(x) = \frac{1}{2}(x-c)^\intercal(x-c) - (\frac{1-e^\intercal c}{n})(e^\intercal x - e^\intercal c) + \frac{1}{2}((\frac{1-e^\intercal c}{n})^2 e^\intercal e)$$

Noticing that our constraint says that $e^\intercal x = 1$, we can say that:

$$f_d(x) = \frac{1}{2}(x-c)^\intercal(x-c) - (\frac{1-e^\intercal c}{n})(1 - e^\intercal c) + \frac{1}{2}((\frac{1-e^\intercal c}{n})^2 e^\intercal e)$$

$$f_d(x) = f_c(x) - (\frac{1-e^\intercal c}{n})(1 - e^\intercal c) + \frac{1}{2}((\frac{1-e^\intercal c}{n})^2 e^\intercal e)$$

Since P_c and P_d have the same constraints and we know that for the objective function of P_d is $f_d(x) = f_c(x) + K$, where K is a constant that does not depend on x, we can say that Equations 1 and 2 will give the exact same solution of x^*.

1.2 Part (b) Correct. But look at solution for a simple argument

Consider Equation 2. We can work to obtain the dual to help us find a solution for x. First, assign u to our equality constraint $g(x) = e^\intercal x - 1$ (and assume $X = \{x : x \geq 0\}$), and then we can obtain the Lagrangian:

$$L_d(x, u) = -u + \frac{1}{2}(x^\intercal - d^\intercal)(x - d) + ue^\intercal x$$

Then we can use this to form the Lagrangian Dual:

$$L_d^*(u) = -u + \inf_x \frac{1}{2}x^\intercal x - d^\intercal x + \frac{1}{2}d^\intercal d + ue^\intercal x$$

This function is obviously convex, so we can take the derivative and set it equal to zero to get the solution: $x = d - ue$ (where $d \geq ue$ because we restricted $x \geq 0$). Substituting this back into our Lagrangian Dual:

$$L_d^*(u) = -u + \frac{1}{2}(d - ue)^\intercal(d - ue) - d^\intercal(d - ue) + \frac{1}{2}d^\intercal d + ue^\intercal(d - ue)$$

Thus now we can further simplify this and form the dual to the primal:

$$\begin{aligned} v_d^* = \quad & \sup_u \quad -u + ud^\intercal e - \frac{1}{2}u^2 n \\ & \text{s.t. } ue \leq d \end{aligned}$$

We know that $d^\intercal e = 1$, thus we can rewrite this dual to be:

$$\begin{aligned} v_d^* = \quad & \sup_u \quad -\frac{n}{2}u^2 \\ & \text{s.t. } ue \leq d \end{aligned}$$

This objective function is clearly concave, so if we assume $d \geq 0$, we can say that our optimal solution is clearly $u^* = 0$. Since we know the relationship between u and x is given by $x = d - ue$, we can say that $x^* = d$.

1.3 Part (c) Shouldn't assume this

Additionally, if any $d_j < 0$ (assume all other $d_{i \neq j} \geq 0$), then we know that $u < 0$. Since we know that our dual's objective function is concave, then our $u^* = d_j$. Thus since $x_j = d_j - u$, we know that $x_j^* = 0$.

1.4 Part (d) Incorrect alg.

Step 1: Given c from problem P_c, calculate $d = c + (\frac{1 - e^\intercal c}{n})e$.
Step 2: If $d_j \geq 0$, then $x_j^* = d$. Else if any $d_j < 0$, assume $\bar{d} = \{d_j : d_j < 0\}$, then $x_i^* = d_i - \min(\bar{d})$.

2 Question 2

10

We have defined the support function on the set $C \subseteq \mathbb{R}^n$ to be:

$$S_C(y) = \sup_y \{y^\intercal x : x \in C\}$$

Likewise we can define the support function on the set $D \subseteq \mathbb{R}^n$ to be:

$$S_D(y) = \sup_y \{y^\intercal x : x \in D\}$$

If $C = D$ then we can say the following:

$$S_C(y) = \sup_y \{y^\intercal x : x \in C\}$$

$$S_C(y) = \sup_y \{y^\intercal x : x \in D\}$$
$$S_C(y) = S_D(y)$$

If $S_C(y) = S_D(y)$ then suppose there is a point $x_D \in D$ that is not in C. If this is true, given that C and D are closed convex sets, then we should be able to separate x_D from the set C with a hyperplane according to Theorem B.3.1. Therefore let us assume there is some nonzero vector p and scalar α such that $p^\intercal x_D > \alpha$ and $p^\intercal x < \alpha$ for all $x \in C$. Using this notation, we can see that:

$$S_C(y) = \sup_x \{p^\intercal x : x \in C\} \leq \alpha$$

Where we also know that:

$$\alpha < p^\intercal x_D \leq \sup_x \{p^\intercal x \in D\} = S_D(y)$$

This implies that:

$$S_C(y) < S_D(y)$$

Which contradicts our assumption of $S_C(y) = S_D(y)$.

Likewise: If $S_C(y) = S_D(y)$ then suppose there is a point $x_C \in C$ that is not in D. If this is true, given that C and D are closed convex sets, then we should be able to separate x_C from the set D with a hyperplane according to Theorem B.3.1. Therefore let us assume there is some nonzero vector p and scalar α such that $p^\intercal x_C > \alpha$ and $p^\intercal x < \alpha$ for all $x \in D$. Using this notation, we can see that:

$$S_D(y) = \sup_x \{p^\intercal x : x \in D\} \leq \alpha$$

Where we also know that:

$$\alpha < p^\intercal x_C \leq \sup_x \{p^\intercal x \in C\} = S_C(y)$$

This implies that:

$$S_D(y) < S_C(y)$$

Which contradicts our assumption of $S_C(y) = S_D(y)$.Thus, we can see that if $S_C(y) = S_D(y)$, then $C = D$.

The above results show that $S_C(y) = S_D(y)$ if and only if $C = D$

3 Question 3

Considering:

$$\begin{aligned}(P) \quad \min_x \ & f(x) \\ \text{s.t. } & g_i(x) \leq 0, \ i = 1, \ldots, m \\ & x \in X\end{aligned} \tag{1}$$

and having r be an integer satisfying $1 \leq r < m$ and $\bar{X} := \{x \in X : g_{r+1} \leq 0, \ldots, g_m \leq 0\}$ for the equation:

$$\begin{aligned}(\bar{P}) \quad \min_x \ & f(x) \\ \text{s.t. } & g_i(x) \leq 0, \ i = 1, \ldots, r \\ & x \in \bar{X}\end{aligned} \tag{2}$$

We need to show that $v^* \leq \bar{v}^* \leq f^*$.

Let us form the Lagrangians of (P) and $\bar{P}$:

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

$$\bar{L}(x, u) = f(x) + \sum_{i=1}^{r} u_i g_i(x)$$

We can then notice that the following is also true given the above:

$$L(x, u) = \bar{L}(x, u) + \sum_{i=r+1}^{m} u_i g_i(x)$$

We know also know that:

$$L^*(u) = \inf_x \ \{L(x, u)\} = \inf_x \ \{\bar{L}(x, u) + \sum_{i=r+1}^{m} u_i g_i(x)\}$$

$$L^*(u) = \bar{L}^*(u) + \inf_x \ \{\sum_{i=r+1}^{m} u_i g_i(x)\}$$

Since we have inequality constraints we know that in addition to $g_i(x) \leq 0$, that we have defined $u \geq 0$. Using this information with the above statement implies:

$$L^*(u) \leq \bar{L}^*(u)$$

Which then we can see that the following is true:

$$\begin{pmatrix} v^* = & \sup_u \ L^*(u) \\ & \text{s.t. } u \geq 0 \end{pmatrix} \leq \begin{pmatrix} \sup_u \ \bar{L}^*(u) & = \bar{v}^* \\ \text{s.t. } u \geq 0 & \end{pmatrix}$$

Thus we can see that $v^* \leq \bar{v}^*$. From weak duality we know that $\bar{v}^* \leq f^*$ and can thus say that $v^* \leq \bar{v}^* \leq f^*$.

4 Question 4

4.1 Part (a)

We can formulate Approach 1 as the following:

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

We can formulate Approach 2 as the following:

$$\begin{aligned} (P_2) \quad \min_x \ & \sum_{j=1}^{s} w_j f_j(x) \\ \text{s.t. } & g_i(x) \le 0, \ i = 1, \dots, m \\ & x \in \mathbb{R}^n \end{aligned}$$

4.2 Part (b)

Let x^{*2} be the optimal solution to (P_2) and assume it is not the optimal solution for (P_1). If this is true, then that means there exists some $\hat{x}$ such that $f_1(\hat{x}) < f_1(x^{*2})$. We know the following about the objective function from (P_2):

$$\sum_{j=1}^{s} w_j f_j(x^{*2}) = w_1 f_1(x^{*2}) + \sum_{j=2}^{s} w_j f_j(x^{*2}) \tag{1}$$

We know if $\hat{x}$ is feasible in (P_1), this it must also be feasible for (P_2). Since we also know that $f_1(\hat{x}) < f_1(x^{*2})$ and assumed $w_1 > 0$, then we can establish a relationship between the two:

$$w_1 f_1(x^{*2}) + \sum_{j=2}^{s} w_j f_j(x^{*2}) > w_1 f_1(\hat{x}) + \sum_{j=2}^{s} w_j f_j(x^{*2})$$

Assume $\hat{x}$ is the optimal solution to (P_1). Now assume that for $j = 2, \dots, s$, that $f_j(x^{*2}) = b_j$. Thus we can say that for $j = 2, \dots, s$ that $f_j(\hat{x}) \le f_j(x^{*2})$. Then we can say that since $w_j \ge 0$ for $j = 1, \dots, s$ that the following is also true:

$$w_1 f_1(\hat{x}) + \sum_{j=2}^{s} w_j f_j(x^{*2}) \ge w_1 f_1(\hat{x}) + \sum_{j=2}^{s} w_j f_j(\hat{x})$$

Which would then imply that:

$$\sum_{j=1}^{s} w_j f_j(x^{*2}) > \sum_{j=1}^{s} w_j f_j(\hat{x})$$

This contradicts our assumption that x^{*2} is the optimal solution to (P_2). Thus we can say that x^{*2} should also be the optimal solution to (P_1) and we should be able to find some b_j's to satisfy (P_1).

4.3 Part (c)

Assuming x^{*1} is an optimal solution for (P_1), then our KKT conditions are:

1. $f_j(x^{*1}) - b_j \le 0,\ j = 2, \dots, s,\ g_i(x^{*1}) \le 0,\ i = 1, \dots, m,\ x^{*1} \in \mathbb{R}^n$
2. $u_j \ge 0,\ j = 2, \dots, s,\ v_i \ge 0,\ i = 1, \dots, m$
3. $(f_j(x^{*1}) - b_j)u_j = 0,\ g_i(x^{*1})v_i = 0$
4. $\nabla f_1(x^{*1}) + \sum_{j=2}^{s}\{\nabla(f_j(x^{*1}) - b_j)u_j\} + \sum_{i=1}^{m}\{\nabla g_i(x^{*1})v_i\} = 0$

Assuming x^{*2} is an optimal solution for (P_2), then our KKT conditions are:

1. $g_i(x^{*2}) \le 0,\ i = 1, \dots, m,\ x^{*2} \in \mathbb{R}^n$
2. $z_i \ge 0,\ i = 1, \dots, m$
3. $g_i(x^{*2})z_i = 0$
4. $\sum_{j=1}^{s}\{w_j \nabla f_j(x^{*2})\} + \sum_{i=1}^{m}\{\nabla g_i(x^{*2})z_i\} = 0$

4.4 Part (d)

First we should notice that x^{*1} will also be able to satisfy (1) and (3) from the KKT conditions for (P_2) as (P_1)'s has the same constraints as (P_2) and more. Secondly, let us notice the following about (4) from (P_1)'s KKT conditions:

$$[\nabla f_1(x^{*1}) + \sum_{j=2}^{s}\{\nabla(f_j(x^{*1}) - b_j)u_j\} + \sum_{i=1}^{m}\{\nabla g_i(x^{*1})v_j\} = \nabla f_1(x^{*1}) + \sum_{j=2}^{s}\{\nabla f_j(x^{*1})u_j\} + \sum_{i=1}^{m}\{\nabla g_i(x^{*1})v_j\}] = 0$$

If we assume $u_1 = 1$, then we can also say the following about (4) from the KKT conditions from (P_1):

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

Notice that this is now of the exact form of (4) from (P_2)'s KKT conditions if we take $u_j = w_j$ and $v_j = z_j$. This holds true because $u_1 > 0$ and $u_j \ge 0,\ j = 2, \dots, s$ and also because $v_i \ge 0,\ i = 1, \dots, m$ too. Thus x^{*1} would also have a solution to Approach 2 as these are satsified.

5 Question 5

$$\begin{aligned} \min_x \quad & \sum_{(i,j)\in A} \left(\frac{1}{2}R_{ij}x_{ij}^2 - t_{ij}x_{ij}\right) \\ \text{s.t.} \quad & \sum_{j:(i,j)\in A} x_{ij} = \sum_{j:(j,i)\in A} x_{ji},\ \forall i \in N \end{aligned} \tag{1}$$

We need to show that a $x_{i,j},\ (i,j) \in A$ that satisfies both of the following equations is a unique solution to Equation 1:

$$\textbf{Kirchhoff's Law:} \sum_{j:(i,j)\in A} x_{ij} = \sum_{j:(j,i)\in A} x_{ji},\ \forall i \in N \tag{2}$$

$$\textbf{Ohm's Law:}\ v_i - v_j = R_{ij}x_{ij} - t_{ij},\ \forall (i,j) \in A \tag{3}$$

If we get a solution $\hat{x}$ to Equation 1, we already know given the constraint that it will satisfy Kirchhoff's law. Now let us define the following for Equation 1:

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

$$g_i(x) = \sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji},\ \forall i \in N$$

Define the set of arcs to be stored the matrix $A \in \mathbb{R}^{m\times 2}$, and have a_k where $k = 1, \ldots, m$ be each row of A defining the arc set of (i,j). We can use this to better define each element of our gradient vectors, where each element corresponds to an a_k:

$$\nabla_{a_k} f(x) = R_{ij}x_{ij} - t_{ij} : (i,j) \in a_k$$

$$\nabla_{a_k} g_i(x) = \begin{cases} 1 & (i,j) = a_k \\ -1 & (j,i) = a_k \\ 0 & \text{otherwise} \end{cases}$$

We can notice that $f(x)$ is strictly convex as it is simply a quadratic equation and that $R_{ij} > 0$ means it will have a positive definite Hessian. We know all $g_i(x)$ also convex as it is just a linear function and that they are all linearly independent as each corresponds to a different node. Additionally, since X is an open set, we can say that Equation 1 is a convex problem. Thus, we can use Theorem 3.2.1 to say the KKT Conditions for this problem are able to give us our globally optimal solution to the problem. That is we should be able to satisfy the following problem:

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

Rearranging we can see that:

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

If we take $u_i = -v_i$ and $u_j = -v_j$, then we can see that this is exactly Ohm's Law:

$$R_{ij}x_{ij} - t_{ij} = v_i - v_j, \forall (i,j) \in A$$

Thus we can see that the KKT conditions for Equation 1 tell us that our optimal solution will satsify Ohm's Law. The constraints for Equation 1 tell us that Kirchhoff's Law will also be satisfied. Additionally, since we know our constraints are convex, X is an open set, and $f(x)$ is strictly convex, then we can say that our solution will also be unique.