# Homework 2

Emre Beray Boztepe

19 06 2024

## Exercise One

For the next two exercises, we consider the LP

$min_{x \in \mathbb{R}} \quad x \; s.t. \quad x \geq 0$

### Option 1

**Write the KKT conditions and verify the optimum solution using those conditions.**

The Karush-Kuhn-Tucker (KKT) conditions are necessary conditions that a solution $x^*$ must satisfy to be optimal, provided that certain regularity conditions hold.

Lagrangian for this exercise:

- We have objective function - f(x) = x
- Constraint function g(x): Given $x \geq 0$, we define $g(x) = -x$ so that the constraint becomes $g(x) \leq 0$. Thus, we will have one Lagrange Multiplier, $\lambda$

General Formula for Lagrangian = $L(x, \lambda) = f(x) - \lambda g(x)$

In our case: $L(x, \lambda) = x - \lambda(-x) = x + \lambda x$

KKT conditions:

- Primal Feasibility: All constraints must be satisfied

$$x \geq 0$$

- Dual Feasibility: The Lagrange multipliers associated with constraints have to be non-negative (zero or positive).

$$\lambda \geq 0$$

- Stationarity: No possible objective improvement at the solution.

We can find it by derivation of Lagrangian with respect to x

$$\frac{\partial L}{\partial x} = 1 - \lambda = 0$$

- Complementary Slackness: The product of the Lagrange multipliers and the corresponding variables must be zero.

$$\lambda_i\big(g_i(x)\big) = 0$$

$$\lambda(-x) = 0$$

simplifies to: $\lambda x = 0$

These conditions are necessary for optimality. Since the objective and constraints are convex, these conditions are also sufficient.

Now, from the conditions we have, if we look for Stationary option:

$1 - \lambda = 0$ and say $\lambda = 1$ to satisfy this option

And, from Complementary option:

$\lambda x = 0$ and substitute $\lambda = 1$, we would have $1 \times x = 0$ and we can say $x = 0$

It satisfies all of the conditions above:

1-) $x \geq 0 \rightarrow 0 \geq 0$

2-) $\lambda \geq 0 \rightarrow 1 \geq 0$

3-) $1 - \lambda = 0 \rightarrow 1 - 1 = 0$

4-) $x = 0  × 0 = 0 $

Hence, x = 0 is an optimal solution.

### Option 2

**Determine the central path C and draw it**

For this, I referenced the figure 14.2 from the book. And I did the drawing by myself using app.diagrams.net website, without using any programming languages.

For our case, since we only have one x value (primal value) and one s value (dual variable),

$$A^T\lambda + s - c = 0$$

$$Ax - b = 0$$

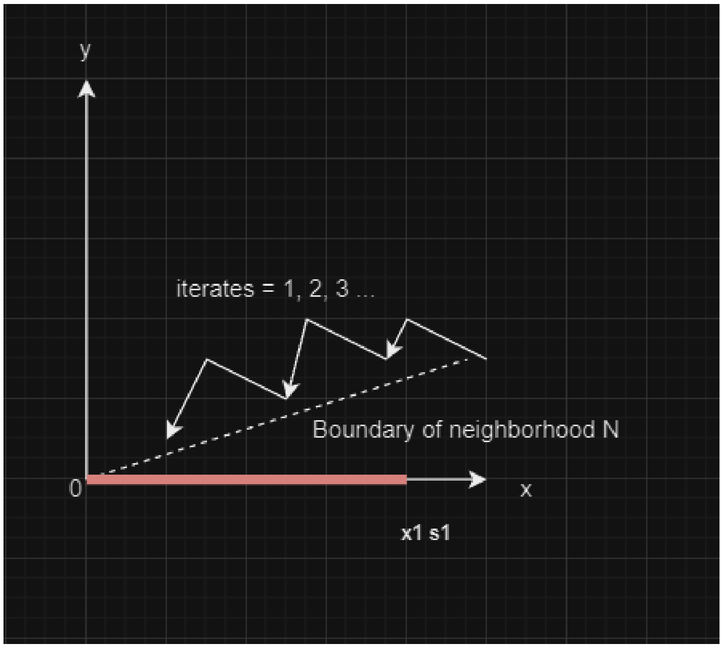
We can find

$$s - 1 = 0$$

$$0 = 0$$

$$xs = x_1 s_1$$

This is how drawing would look like:




Red line represents our central path which goes to 0 in each iterative (which is the optimal solution)

The area between dashed line and x-axis represents the boundary of neighborhood N

and x-axis represents our primal and dual variables.

## Exercise Two

**We continue with the same LP from the previous exercise. Assuming the complementarity condition we wish is $XSe = \sigma\mu e$ (the same as in Framework 14.1), write the specific formula for F and its Jacobian J for our problem. Then, compute one step of the Newton method for finding F(x) = 0 (a full step, i.e., with $\alpha$ = 1) and show that it jumps directly to the central path from any initial point that is strictly feasible.**

We know that:

- S: diagonal matrix of dual variables s
- X: diagonal matrix of primal variables x
- e: vector of ones
- $\sigma$: centering parameter
- $\mu$: duality measure

our LP:

$minimize \quad x$ subject to $x \geq 0$

First, we need to find its dual problem.

Dual:

$maximize 0$ subject to $\lambda + s = 1, s \geq 0$

Checking KKT conditions:

- $x = 0$, objective function achieves its minimum value
- $\lambda + s = 1$, dual variables
- $x_i \times s_i = 0$, complementary slackness
- $(x_i, s_i) \geq 0$, both x and s for each i is non-negative

The specific formula for F:

We know that F in general case for dual problem of $min \quad c^T x$ subject to $Ax = b, x \geq 0$:

$max \quad b^T \lambda$ subject to $A^T \lambda + s = c, s \geq 0$

$$F(x, \lambda, s) = \begin{bmatrix} A^T \lambda + s - c \\ Ax - b \\ XSe \end{bmatrix} = \vec{0}$$

For $(x, s) \geq 0$

In our case given complementary condition $XSe = \sigma \mu e$, since we have one dual variable $(x \geq 0)$ and one primal variable (x) we can simplify it as:

$$xs = \sigma \mu$$

Now, we can define F:

$$F(x, \lambda, s) = \begin{bmatrix} \lambda + s - 1 \\ x \\ xs - \sigma \mu \end{bmatrix} = \vec{0}$$

for $(x, s) \geq 0$

where:

- $\lambda + s - 1$ = 0: enforces dual feasibility
- $x$ = 0: enforces primal feasibility
- $xs - \sigma \mu$ = 0: modified complementarity condition

Jacobian of F is basically partial derivatives of F wit respect to each variable.

Derivatives for $F_1(x, \lambda, s)$ = $\lambda + s - 1$

$\frac{\partial F_1}{\partial x}$ = 0 (no x in $F_1$)

$\frac{\partial F_1}{\partial \lambda} = 1$

$\frac{\partial F_1}{\partial s} = 1$

Derivatives for $F_2(x, \lambda, s) = x$

$\frac{\partial F_2}{\partial x} = 1$

$\frac{\partial F_2}{\partial \lambda} = 0$ (no $\lambda$ in $F_2$)

$\frac{\partial F_2}{\partial s} = 0$ (no s in $F_2$)

Derivatives for $F_3(x, \lambda, s) = xs - \sigma\mu$

$\frac{\partial F_3}{\partial x} = \mathrm{s}$

$\frac{\partial F_3}{\partial \lambda} = 0$ (no $\lambda$ in $F_3$)

$\frac{\partial F_3}{\partial s} = \mathrm{x}$

So, Jacobian is:

$$J = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ s & 0 & x \end{bmatrix}$$

Now, to perform one step Newton Method to find $\Delta x, \Delta\lambda, \Delta s$, we can use this equation:

$$J(x, \lambda, s) \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{bmatrix} = -F(x, \lambda, s)$$

For simplicity, we can use pre-defined initial points which satisfies strictly feasible point $x = 1, \lambda = 1, s = 1$ and $\sigma = 0.9, \mu = 1$:

$$F = \begin{bmatrix} 1 + 1 - 1 \\ 1 \\ 1 \times 1 - 0.9 \times 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0.1 \end{bmatrix}$$

and

$$J = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Get back to the formula:

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{bmatrix} = -\begin{bmatrix} 1 \\ 1 \\ 0.1 \end{bmatrix}$$

When we do the matrix calculation, we would get these 3 equations:

$$\Delta\lambda + \Delta s = -1$$

$$\Delta x = -1$$

$$\Delta x + \Delta s = -0.1$$

Since $\Delta x = -1$ from the second equation, when we substitute it with our third equation:

$\Delta x + \Delta s = -0.1 = -1 + \Delta s = -0.1$

$$\Delta s = 0.9$$

And from the first equation:

$\Delta\lambda + \Delta s = -1 = \Delta\lambda + 0.9 = -1$

$$\Delta\lambda = -1.9$$

Now that we have:

$$\begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -1 \\ -1.9 \\ 0.9 \end{bmatrix}$$

We can calculate the new values for $x, \lambda, s$ by using ($\alpha = 1$):

$$(x^1, \lambda^1, s^1) = (x^0, \lambda^0, s^0) + \alpha(\Delta x^0, \Delta\lambda^0, \Delta s^0)$$

$$(x^1, \lambda^1, s^1) = (1,1,1) + 1(-1, -1.9, 0.9)$$

$$x^1 = 0$$

$$\lambda^1 = 0.9$$

$$s^1 = 1.9$$

To validate the updated solution places the variables on the central path, we would check if the conditions for being on the central path are met:

The new values must satisfy the primal and dual feasibility conditions. Since our new x is non-negative, it meets the primal feasibility. And our new $\lambda$ and s are also non-negative, they meet the dual feasibility.

In moving along the central path, we expect $x_i \times s_i$ (for each i) to be positive and decrease towards zero as we approach the solution. For our case, we have $x^1 \times s^1 = 0 \times 1.9 = 0$, this meets the ideal complementarity at he solution point.

And our objective x remains minimized at zero.

## Exercise Three

**Suppose we are computing one Newton step for the interior point path-following method, meaning we solve:**

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe + \sigma \mu e \end{bmatrix}$$

where $\mu = x^T \frac{s}{n}$

**Suppose we already start from a strictly feasible point $(x, \lambda, s) \in F_0$, meaning we also have $-r_c = 0$ and $-r_b = 0$ above.**

**Prove a simple but quite important observation that our next direction $(\Delta x, \Delta \lambda, \Delta s)$ satisfies that $(\Delta x)^T (\Delta s) = 0$, so the directions for $\Delta x$ and $\Delta s$ are orthogonal.**

**Hint: No difficult math or any extra knowledge is required. It is really a one-line proof. Just try to play with the system of equations above until you get $(\Delta x)^T (\Delta s)$ on the left.**

From the matrix calculations, we can derive these equations (knowing that $-r_c = 0$ and $-r_b = 0$):

$$A^T \Delta \lambda + \Delta s = 0 \Rightarrow \Delta s = -A^T \Delta \lambda$$

$$A \Delta x = 0$$

$$S \Delta x + X \Delta s = -XSe + \sigma \mu e$$

So, for the first equation we have ($\Delta s = -A^T \Delta \lambda$), if we multiply both sides with $(\Delta x)^T$, we would get:

$$(\Delta x)^T \Delta s = (\Delta x)^T (-A^T \Delta \lambda)$$

Right hand side:

$$\Rightarrow (\Delta x)^T (-A^T \Delta \lambda) = -(\Delta x)^T A^T \Delta \lambda = -(A \Delta x)^T \Delta \lambda$$

Previously, we found out that $A \Delta x = 0$:

$$\Rightarrow -(0)^T \Delta \lambda = 0$$

$$(\Delta x)^T \Delta s = 0$$

So, 0 confirms orthogonality. This is ensuring that the changes in primal and dual variables do not counteract each other destructively.

## Exercise Four

**This exercise illustrates the fact that the bounds $x \geq 0, s \geq 0$ are essential in relating solutions of the F(x,λ,s) = 0 (as defined in the interior-point method) to solutions of our starting linear program and its dual.**

**Consider the following linear program in $\mathbb{R}^2$:**

$min \quad x_1$, subject to $x_1 + x_2 = 1$, $(x_1, x_2) \geq 0$

**Show that the primal-dual solution is**

$$x^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \lambda^* = 0, s^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

**Also verify that the system F(x,λ,s) = 0 also has the solution**

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \lambda = 1, s = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

**which has no relation to the solution of the linear program.**

Primal Solution: Our objective function is to minimize $x_1$ and we want it to be as small as possible and we have condition $x \geq 0$. So, the smallest $x_1$ is 0.

And from the constraint $(x_1 + x_2 = 1)$, we can say:

$$0 + x_2 = 1 \Rightarrow x_2 = 1$$

Therefore, the primal optimal solution is:

$$x^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Dual solution:

We have one constraint $(x_1 + x_2 = 1)$. Lagrangian for this constraint is:

$$L(x, \lambda) = x_1 + \lambda(1 - x_1 - x_2)$$

and our dual objective is to maximize $\lambda$. Deriving dual function:

$\frac{\partial L}{\partial x_1} = 1 - \lambda = 0 \Rightarrow \lambda = 1$

$\frac{\partial L}{\partial x_2} = -\lambda = 0 \Rightarrow \lambda = 0$

This discrepancy indicates that there is no positive value of $\lambda$ that can simultaneously satisfy both conditions. Therefore, the only value of $\lambda$ that does not contradict any derivative condition is $\lambda = 0$.

So, $\lambda^* = 0$

Since, $x_1 = 0$, the dual constraint associated with $x_1$ is tight. That means the corresponding slack variable $s_1$ could be 1.

And $x_2 = 1$, the dual constraint associated with $x_2$ is non-binding. That means the corresponding slack variable $x_2$ could be 1.

Therefore, we can say dual optimal solution is:

$$s^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

To check the feasibility and optimality of another given solution with $x, \lambda, s$

For primal feasibility, when we check only constraint we have, $x_1 + x_2 = 1$ for $x_1 = 0, x_2 = 1$:

$1 + 0 = 1$, So, constraint is satisfied.

For dual feasibility, we must have non-negativity of the slack variables $s \geq 0$:

and we have $s_1 = 0, s_2 = -1$. It is not feasible as $s_2 < 0$

If we check complementary slackness condition which requires $x_i \times s_i = 0$ for each i:

$x_1 \times s_1 = 1 \times 0 = 0$, it is satisfied

$x_2 \times s_2 = 0 \times (-1) = 0$, it is also satisfied

While the complementary slackness condition holds (satisfies the system $F(x, \lambda, s) = 0$), the dual feasibility condition $s \geq 0$ is violated because $s_2 = -1$. So, it can be said that it is not valid for the LP problem.

## Exercise Five

**State the definition of the neighborhood $N_{-\infty}(\gamma)$ of the central path from the lectures on primal-dual constrained optimization. In other primal-dual approaches, a different neighborhood can be important, called $N_2(\theta)$, parametrized by $\theta \in [0,1)$:**

$$N_2(\theta) = \big((x, \lambda, s) | (x, \lambda, s) \in F_0, ||XSe - \mu e||_2 \leq \theta\mu\big)$$

**Prove that for $\gamma \leq 1 - \theta$, we have $N_2(\theta) \subseteq N_{-\infty(\gamma)}$.**

$N_{-\infty}(\gamma)$: This is a neighborhood of the central path defined by:

$N_{-\infty}(\gamma) = \{(x, \lambda, s) \in \overline{f_0} | x_i s_i \geq \mu\gamma\} \; \forall i$

This defines a set $N_{-\infty}(\gamma)$ consisting of triples $(x, \lambda, s)$ that belong to the set $\overline{f_0}$, where each product $x_i s_i$ is at least $\mu$. Here, $\mu$ depends on $\gamma$ and represents a scaled measure of centrality in the feasible region.

- $\gamma \in [0,1]$

- $\overline{f_0}$: strict feasibility region, which is $\overline{f} \cap \{x > \vec{0}, s > \vec{0}\}$
- $\overline{f}$: feasible region, which we can represent as

$$\left\{(x, \lambda, s) \middle| \begin{array}{cccc} Ax & = b & x & \geq \vec{0} \\ A^T\lambda + s & = c & s & \geq \vec{0} \end{array}\right\}$$

- C is central path which is $\{(x, \lambda, s) \in \overline{f_0} \mid \exists\tau > 0$ such that $\forall i\ s_i x_i = \tau$

and $N_2(\theta)$ another type of neighborhood defined by

$$N_2(\theta) = \{(x, \lambda, s) \mid (x, \lambda, s) \in f_0 | |XSe - \mu e| |_2 \leq \theta\mu$$

where $| |.| |_2$ denotes the euclidean norm, $\theta \in [0,1]$ and e is vector of ones.

Proof:

We need to show that any point $(x, \lambda, s)$ in $N_2(\theta)$ also satisfies the condition for $N_{-\infty}(\gamma)$ given $\gamma \leq 1 - \theta$

for $\gamma \leq 1 - \theta$, we have $N_2(\theta) \subseteq N_{-\infty}(\gamma)$

from $(x, \lambda, s) \in N_2(\theta)$

Below condition implies that the vector of products $(XS = (x_1 s_1, x_2 s_2, \ldots, x_n s_n))$ is close to the vector where all elements are $\mu$, within a Euclidean distance of $\theta\mu$

$| |XSe - \mu e| |_2 = \sqrt{\sum_{i=1}^{n}(x_i s_i - \mu)^2} \leq \theta\mu$

By squaring both sides:

$$\Rightarrow \sum_{i=1}^{n}(x_i s_i - \mu)^2 \leq (\theta\mu)^2$$

$$\Rightarrow \sum_{i=1}^{n}( (x_i s_i)^2 - 2\mu x_i s_i + (\mu)^2 \leq (\theta\mu)^2$$

$$\Rightarrow \sum_{i=1}^{n}(x_i s_i)^2 - 2\mu \sum_{i=1}^{n} x_i\, s_i + n\mu^2 \leq \theta^2\mu^2$$

If we substitute by using the fact that $\sum_{i=1}^{n}(x_i s_i) = n\mu$:

$$\Rightarrow \sum_{i=1}^{n}(x_i s_i)^2 - 2\mu(n\mu) + n\mu^2 \leq \theta^2\mu^2$$

$$\Rightarrow \sum_{i=1}^{n}(x_i s_i)^2\,(-2\mu^2 n + n\mu^2) \leq \theta^2\mu^2$$

$$\Rightarrow \sum_{i=1}^{n}(x_i s_i)^2 \leq n\mu^2 + \theta^2\mu^2$$

When we take a look into inequality that we derive earlier: $\sum_{i=1}^{n}(x_i s_i - \mu)^2 \leq (\theta\mu)^2$, Using this:

We know

$(x_i s_i - \mu)^2 \leq (\theta\mu)^2\ \forall i$

This implies:

$$\Rightarrow |x_i s_i - \mu| \leq \theta\mu$$

For the absolute value, we would have:

$$\Rightarrow -\theta\mu \leq x_i s_i - \mu \leq \theta\mu$$

$$\Rightarrow \mu - \theta\mu \leq x_i s_i \leq \mu + \theta\mu$$

$$\Rightarrow (1-\theta)\mu \leq x_i s_i \leq \mu(1+\theta)$$

For $N_{-\infty}(\gamma)$ we require: $x_i s_i \geq \gamma_\mu$

Since $x_i s_i \geq (1-\theta)\mu$ and $\gamma \leq 1-\theta$ it follows that:

$$x_i s_i \geq \mu\gamma$$

Finally, after establishing that $\mu - \theta\mu \leq x_i s_i$, we conclude that $(1-\theta)\mu \leq x_i s_i$ for each i. Since each $x_i s_i$ is no less than $(1-\theta)\mu$ and $\gamma\mu$ does not exceed $(1-\theta)\mu$, it follows that every $x_i s_i \geq \gamma\mu$. This confirms that all points in $N_2(\theta)$ satisfy the conditions for $N_2(\theta) \subseteq N_{-\infty}(\gamma)$