

Homework Assignment 7

Matthew Tiger

April 18, 2016

Problem 1. State all of the KKT conditions for (N -max). More precisely state all of the following results for (N -max): KKT-FONC, KKT-FOSC, KKT-SONC, KKT-SOSC.

Solution. For the following theorems, we assume (N -max) has the following form

$$\begin{aligned} (N\text{-max}) \quad & \text{maximize} && f(\mathbf{x}) \\ & \text{subject to} && \mathbf{h}(\mathbf{x}) = \mathbf{0} \\ & && \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \end{aligned}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, and $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ with $m \leq n$. Additionally, define the following Lagrangian function to be $\mathbf{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) := -f(\mathbf{x}) + \boldsymbol{\lambda}^\top \mathbf{h}(\mathbf{x}) + \boldsymbol{\mu}^\top \mathbf{g}(\mathbf{x})$.

Theorem 1 (KKT-FONC for (N -max)). Let $f, \mathbf{g}, \mathbf{h} \in C^1$ and let \mathbf{x}^* be a regular point and local maximizer for the problem (N -max). Then, there exist $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that:

- i. $\boldsymbol{\mu}^* \geq \mathbf{0}$.
- ii. $D_{\mathbf{x}}\mathbf{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = -Df(\mathbf{x}^*) + \boldsymbol{\lambda}^{*\top} D\mathbf{h}(\mathbf{x}^*) + \boldsymbol{\mu}^{*\top} D\mathbf{g}(\mathbf{x}^*) = \mathbf{0}^\top$.
- iii. $\boldsymbol{\mu}^{*\top} \mathbf{g}(\mathbf{x}^*) = 0$.

Note that there are no explicit first-order conditions that are sufficient in general to show optimality.

Theorem 2 (KKT-SONC for (N -max)). Let $f, \mathbf{g}, \mathbf{h} \in C^2$ and let \mathbf{x}^* be a regular point and local maximizer for the problem (N -max). Then, there exist $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that:

- i. $\boldsymbol{\mu}^* \geq \mathbf{0}$, $D_{\mathbf{x}}\mathbf{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \mathbf{0}^\top$, $\boldsymbol{\mu}^{*\top} \mathbf{g}(\mathbf{x}^*) = 0$.
- ii. For all $\mathbf{y} \in T(\mathbf{x}^*) = \{\mathbf{y} \mid D\mathbf{h}(\mathbf{x}^*)\mathbf{y} = \mathbf{0}, D\mathbf{g}_j(\mathbf{x}^*)\mathbf{y} = 0, j \in J(\mathbf{x}^*)\}$, we have that $\mathbf{y}^\top D_{\mathbf{x}}^2 \mathbf{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \mathbf{y} \leq 0$.

Theorem 3 (KKT-SOSC for (N -max)). Let $f, \mathbf{g}, \mathbf{h} \in C^2$ and suppose there exists a feasible point \mathbf{x}^* and vectors $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that:

- i. $\boldsymbol{\mu}^* \geq \mathbf{0}$, $D_{\mathbf{x}}\mathbf{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \mathbf{0}^\top$, $\boldsymbol{\mu}^{*\top} \mathbf{g}(\mathbf{x}^*) = 0$.

ii. For all

$$\mathbf{y} \in \tilde{T}(\mathbf{x}^*, \boldsymbol{\mu}^*) = \{\mathbf{y} \mid D\mathbf{h}(\mathbf{x}^*)\mathbf{y} = \mathbf{0}, Dg_i(\mathbf{x}^*)\mathbf{y} = 0, \text{ for } i \in \{i \mid g_i(\mathbf{x}^*) = 0, \mu_i^* > 0\}\},$$

with $\mathbf{y} \neq \mathbf{0}$, we have that $\mathbf{y}^\top D_x^2 \mathbf{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\mathbf{y} < 0$.

Then \mathbf{x}^* is a strict local maximizer for the problem (N -max).

□

Problem 2. Find local minimizers for

$$\begin{aligned} (N\text{-min}) \quad & \text{minimize} \quad x_1^2 + 6x_1x_2 - 4x_1 - 2x_2 \\ & \text{subject to} \quad x_1^2 + 2x_2 \leq 1 \\ & \quad \quad \quad 2x_1 - 2x_2 \leq 1. \end{aligned}$$

Solution. We begin by rewriting the above problem as follows:

$$\begin{aligned} (N\text{-min}) \quad & \text{minimize} \quad f(\mathbf{x}) = x_1^2 + 6x_1x_2 - 4x_1 - 2x_2 \\ & \text{subject to} \quad g_1(\mathbf{x}) = x_1^2 + 2x_2 - 1 \leq 0 \\ & \quad \quad \quad g_2(\mathbf{x}) = 2x_1 - 2x_2 - 1 \leq 0. \end{aligned}$$

We proceed by using the KKT-FONC to determine the possible local minimizers for this problem. The Lagrangian associated to this problem is given by

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\mu}) &= f(\mathbf{x}) + \boldsymbol{\mu}^\top \mathbf{g}(\mathbf{x}) \\ &= f(\mathbf{x}) + \mu_1 g_1(\mathbf{x}) + \mu_2 g_2(\mathbf{x}) \\ &= x_1^2 + 6x_1x_2 - 4x_1 - 2x_2 + \mu_1(x_1^2 + 2x_2 - 1) + \mu_2(2x_1 - 2x_2 - 1). \end{aligned}$$

This implies that

$$D_{\mathbf{x}}L(\mathbf{x}, \boldsymbol{\mu}) = \begin{bmatrix} 2x_1 + 6x_2 - 4 + 2\mu_1x_1 + 2\mu_2 \\ 6x_1 - 2 + 2\mu_1 - 2\mu_2 \end{bmatrix}^\top = \begin{bmatrix} 0 \\ 0 \end{bmatrix}^\top. \quad (1)$$

Thus, the KKT-FONC are then written as

- i. $\mu_1, \mu_2 \geq 0$.
- ii. $2x_1 + 6x_2 - 4 + 2\mu_1x_1 + 2\mu_2 = 0$.
- iii. $6x_1 - 2 + 2\mu_1 - 2\mu_2 = 0$.
- iv. $\mu_1 g_1(\mathbf{x}) + \mu_2 g_2(\mathbf{x}) = \mu_1(x_1^2 + 2x_2 - 1) + \mu_2(2x_1 - 2x_2 - 1) = 0$.
- v. $g_1(\mathbf{x}) = x_1^2 + 2x_2 - 1 \leq 0$.
- vi. $g_2(\mathbf{x}) = 2x_1 - 2x_2 - 1 \leq 0$.

Solving the system (1) for x_1, x_2 yields that

$$\begin{aligned} x_1 &= \frac{\mu_2 - \mu_1 + 1}{3} \\ x_2 &= \frac{\mu_1^2 - \mu_1\mu_2 - 4\mu_2 + 5}{9} \end{aligned} \quad (2)$$

with $\mu_1, \mu_2 \geq 0$. Using these representations of x_1, x_2 we see that condition iv. yields three possible solutions in terms of μ_1, μ_2 :

$$\begin{aligned} \text{Case 1 : } \mu_2 &= \frac{13 + 12\mu_1 + 6\mu_1^2 - \sqrt{169 + 200\mu_1 + 388\mu_1^2}}{2(14 + 3\mu_1)} \\ \text{Case 2 : } \mu_2 &= \frac{13 + 12\mu_1 + 6\mu_1^2 + \sqrt{169 + 200\mu_1 + 388\mu_1^2}}{2(14 + 3\mu_1)} \\ \text{Case 3 : } \mu_1 &= -\frac{14}{3}, \mu_2 = -\frac{3220}{789} \end{aligned}$$

We readily see that Case 3 cannot happen in light of condition i.

Assuming Case 1 is true and using the representations of x_1, x_2 in (2), we see that $g_1(\mathbf{x}) < 0$ for $\mu_1, \mu_2 \geq 0$ implying that this constraint is inactive and that $\mu_1 = 0$. This implies that $\mu_2 = 0$ which in turn implies that $x_1 = 1/3, x_2 = 5/9$. However, $g_1(x_1, x_2) = 2/9 \not\leq 0$ violating condition v. Thus, Case 1 cannot happen.

Assuming Case 2 is true and using the representations of x_1, x_2 in (2), we again see that $g_1(\mathbf{x}) < 0$ for $\mu_1, \mu_2 \geq 0$ implying that this constraint is inactive and that $\mu_1 = 0$. This implies that $\mu_2 = 13/14$ which in turn implies that $x_1 = 9/14, x_2 = 1/7$. These values of x_1, x_2 satisfy conditions v. and vi.

Therefore, the only vector \mathbf{x}^* that satisfies conditions i. - vi., i.e. the only possible local minimizer for this problem is

$$\mathbf{x}^* = \begin{bmatrix} 9/14 \\ 1/7 \end{bmatrix}$$

with associated KKT multiplier

$$\boldsymbol{\mu}^* = \begin{bmatrix} 0 \\ 13/14 \end{bmatrix}.$$

To verify whether or not this vector is a strict local minimizer, we check the KKT-SOSC. For the vectors \mathbf{x}^* and $\boldsymbol{\mu}^*$ defined above, we see that $\{i \mid g_i(\mathbf{x}^*) = 0, \mu_i^* > 0\} = \{2\}$ and that

$$\begin{aligned} \tilde{T}(\mathbf{x}^*, \boldsymbol{\mu}^*) &= \{\mathbf{y} \in \mathbb{R}^2 \mid Dg_2(\mathbf{x}^*)\mathbf{y} = 0\} \\ &= \{\mathbf{y} \in \mathbb{R}^2 \mid [2, -2]\mathbf{y} = 0\} \\ &= \{\mathbf{y} = [y_1, y_2]^\top \in \mathbb{R}^2 \mid y_1 = y_2\}. \end{aligned}$$

Further, we have, for these vectors, that

$$D_x^2 L(\mathbf{x}^*, \boldsymbol{\mu}^*) = \begin{bmatrix} 2 + 2\mu_1 & 6 \\ 6 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 6 & 0 \end{bmatrix}.$$

Combining we see that for $\mathbf{0} \neq \mathbf{y} \in \tilde{T}(\mathbf{x}^*, \boldsymbol{\mu}^*)$, we have that

$$\mathbf{y}^\top D_x^2 L(\mathbf{x}^*, \boldsymbol{\mu}^*) \mathbf{y} = \begin{bmatrix} y_1 \\ y_1 \end{bmatrix}^\top \begin{bmatrix} 2 & 6 \\ 6 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_1 \end{bmatrix} = 14y_1^2 > 0$$

for $y_1 \neq 0$. Therefore, $\mathbf{x}^* = [9/14, 1/7]^\top$ is a strict local minimizer. □

Problem 3. Consider the problem of optimizing

$$(N) \quad \begin{array}{ll} \text{minimize (maximize)} & (x_1 - 2)^2 + (x_2 - 1)^2 \\ & x_2 - x_1^2 \geq 0 \\ \text{subject to} & 2 - x_1 - x_2 \geq 0 \\ & x_1 \geq 0. \end{array}$$

Let $\mathbf{x}^* = [0, 0]^\top$.

- Does \mathbf{x}^* satisfy the KKT-FONC for minimization or maximization? What are the KKT multipliers?
- Does \mathbf{x}^* satisfy the KKT-SOSC? Justify your answer.

Solution. We begin by rewriting the problem (N) as

$$(N_1) \quad \begin{array}{ll} \text{minimize (maximize)} & f(\mathbf{x}) = (x_1 - 2)^2 + (x_2 - 1)^2 \\ & g_1(\mathbf{x}) = -x_2 + x_1^2 \leq 0 \\ \text{subject to} & g_2(\mathbf{x}) = -2 + x_1 + x_2 \leq 0 \\ & g_3(\mathbf{x}) = -x_1 \leq 0. \end{array}$$

For both problems, the vector $\mathbf{x}^* = [0, 0]^\top$ is a regular point. To see this, we note that \mathbf{x}^* is feasible and the constraints $g_1(\mathbf{x}^*) \leq 0$ and $g_3(\mathbf{x}^*) \leq 0$ are both active for this vector. Since $\nabla g_1(\mathbf{x}^*) = [0, -1]^\top$ and $\nabla g_3(\mathbf{x}^*) = [-1, 0]^\top$ are linearly independent, we have that \mathbf{x}^* is a regular point as desired.

The Lagrangian function associated to problem (N_1 -min) is given by

$$\begin{aligned} L_{\min}(\mathbf{x}, \boldsymbol{\mu}) &= f(\mathbf{x}) + \mu_1 g_1(\mathbf{x}) + \mu_2 g_2(\mathbf{x}) + \mu_3 g_3(\mathbf{x}) \\ &= (x_1 - 2)^2 + (x_2 - 1)^2 + \mu_1(-x_2 + x_1^2) + \mu_2(-2 + x_1 + x_2) + \mu_3(-x_1) \end{aligned}$$

while the Lagrangian associated to the problem (N_1 -max) is given by

$$\begin{aligned} L_{\max}(\mathbf{x}, \boldsymbol{\mu}) &= -f(\mathbf{x}) + \mu_1 g_1(\mathbf{x}) + \mu_2 g_2(\mathbf{x}) + \mu_3 g_3(\mathbf{x}) \\ &= -(x_1 - 2)^2 - (x_2 - 1)^2 + \mu_1(-x_2 + x_1^2) + \mu_2(-2 + x_1 + x_2) + \mu_3(-x_1). \end{aligned}$$

- Note that for problem (N_1 -min), we have that

$$D_{\mathbf{x}} L_{\min}(\mathbf{x}, \boldsymbol{\mu}) = \begin{bmatrix} 2(x_1 - 2) + 2\mu_1 x_1 + \mu_2 - \mu_3 \\ 2(x_2 - 1) + \mu_2 - \mu_1 \end{bmatrix}^\top,$$

while for the problem (N_1 -max), we have that

$$D_{\mathbf{x}} L_{\max}(\mathbf{x}, \boldsymbol{\mu}) = \begin{bmatrix} -2(x_1 - 2) + 2\mu_1 x_1 + \mu_2 - \mu_3 \\ -2(x_2 - 1) + \mu_2 - \mu_1 \end{bmatrix}^\top.$$

The KKT-FONC for problem (N_1 -min) then require that the following conditions hold

- $\mu_1, \mu_2, \mu_3 \geq 0$.

- ii a. $2(x_1 - 2) + 2\mu_1 x_1 + \mu_2 - \mu_3 = 0.$
- iii a. $2(x_2 - 1) + \mu_2 - \mu_1 = 0.$
- iv a. $\mu_1 g_1(\mathbf{x}) + \mu_2 g_2(\mathbf{x}) + \mu_3 g_3(\mathbf{x}) = \mu_1(-x_2 + x_1^2) + \mu_2(-2 + x_1 + x_2) + \mu_3(-x_1) = 0.$
- v a. $g_1(\mathbf{x}) = -x_2 + x_1^2 \leq 0.$
- vi a. $g_2(\mathbf{x}) = -2 + x_1 + x_2 \leq 0.$
- vii a. $g_3(\mathbf{x}) = -x_1 \leq 0.$

while the KKT-FONC for problem (N_1 -max) require that the following similar conditions hold

- i b. $\mu_1, \mu_2, \mu_3 \geq 0.$
- ii b. $-2(x_1 - 2) + 2\mu_1 x_1 + \mu_2 - \mu_3 = 0.$
- iii b. $-2(x_2 - 1) + \mu_2 - \mu_1 = 0.$
- iv b. $\mu_1 g_1(\mathbf{x}) + \mu_2 g_2(\mathbf{x}) + \mu_3 g_3(\mathbf{x}) = \mu_1(-x_2 + x_1^2) + \mu_2(-2 + x_1 + x_2) + \mu_3(-x_1) = 0.$
- v b. $g_1(\mathbf{x}) = -x_2 + x_1^2 \leq 0.$
- vi b. $g_2(\mathbf{x}) = -2 + x_1 + x_2 \leq 0.$
- vii b. $g_3(\mathbf{x}) = -x_1 \leq 0.$

Now suppose that $\mathbf{x}^* = [0, 0]^T$. For both problems, since \mathbf{x}^* is a regular point, conditions v a. - vii a. and v b. - vii b. are satisfied. Also, for both problems, since the constraint $g_2(\mathbf{x}^*)$ is inactive we have that by condition iv a. and iv b. that $\mu_2 = 0$.

For the problem (N_1 -min), conditions ii a. and iii a. imply that $\mu_2 - \mu_3 = -\mu_3 = 4$ and $\mu_2 - \mu_1 = -\mu_1 = 2$ or that $\mu_1 = -2$, $\mu_2 = 0$, and $\mu_3 = -4$. However, this violates condition i a. so the point \mathbf{x}^* does not satisfy the KKT-FONC for the problem (N_1 -min).

For the problem (N_1 -max), conditions ii a. and iii a. imply that $\mu_2 - \mu_3 = -\mu_3 = -4$ and $\mu_2 - \mu_1 = -\mu_1 = -2$ or that $\mu_1 = 2$, $\mu_2 = 0$, and $\mu_3 = 4$. Therefore, the vector $\mathbf{x}^* = [0, 0]^T$ satisfies the KKT-FONC for the problem (N_1 -max) with associated KKT multiplier $\boldsymbol{\mu}^* = [2, 0, 4]^T$.

- b. We now check to see if $\mathbf{x}^* = [0, 0]^T$ satisfies the KKT-SOSC for the problem (N_1 -max). Note that for $\mathbf{x}^* = [0, 0]^T$, we have that

$$D_{\mathbf{x}}^2 L_{\max}(\mathbf{x}^*, \boldsymbol{\mu}^*) = \begin{bmatrix} -2 + 2\mu_1 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

We also see that for the vectors \mathbf{x}^* and $\boldsymbol{\mu}^*$ defined above, $\{i \mid g_i(\mathbf{x}^*) = 0, \mu_i^* > 0\} = \{1, 3\}$, and that

$$\begin{aligned} \tilde{T}(\mathbf{x}^*, \boldsymbol{\mu}^*) &= \{\mathbf{y} \in \mathbb{R}^2 \mid Dg_1(\mathbf{x}^*)\mathbf{y} = 0, Dg_3(\mathbf{x}^*)\mathbf{y} = 0\} \\ &= \{\mathbf{y} \in \mathbb{R}^2 \mid [0, -1]\mathbf{y} = 0, [-1, 0]\mathbf{y} = 0\} \\ &= \{\mathbf{0} \in \mathbb{R}^2\}. \end{aligned}$$

Therefore, we trivially have that the second condition in the KKT-SOSC is satisfied and $\mathbf{x}^* = [0, 0]^T$ is a strict local maximizer.

□

Problem 4. Consider the problem with equality constraint

$$\begin{array}{ll}\text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{h}(\mathbf{x}) = \mathbf{0}.\end{array}$$

We can convert the above into the equivalent optimization problem

$$\begin{array}{ll}\text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \frac{1}{2} \|\mathbf{h}(\mathbf{x})\|^2 \leq 0.\end{array}$$

Write down the KKT condition for the equivalent problem and explain why the KKT theorem cannot be applied in this case.

Solution.

□