Advanced Linear Programming Homework 7

Michael Nelson

Problem 1

For this problem, let $f(\boldsymbol{x}) = -x_1x_2x_3$ and let $h(\boldsymbol{x}) = x_1 + x_2 + x_3 - 9$. We consider the following NLP

$$\begin{aligned} &\text{minimize} && f(\boldsymbol{x}) \\ &\text{subject to} && h(\boldsymbol{x}) = 0 \end{aligned} \tag{1}$$

(the author sets this up as a maximization problem with $f(\boldsymbol{x}) = x_1x_2x_3$ however we convert it to a minimization problem as usual by setting $f(\boldsymbol{x}) = -x_1x_2x_3$).

Problem 1.a

Exercise 1. Write the KKT FONC to this NLP.

Solution 1. A KKT point for this NLP is a pair $(\boldsymbol{x}, \lambda) = (x_1, x_2, x_3, \lambda) \in \mathbb{R}^4$ which satisfies the following:

$$\begin{aligned} \nabla f(\boldsymbol{x}) + \lambda \nabla h(\boldsymbol{x}) &= 0 && \text{Stationary} \\ h(\boldsymbol{x}) &= 0 && \text{Primal Feasibility} \end{aligned} \tag{2}$$

Let $\mathcal{K} \subseteq \mathbb{R}^4$ denote the set of all KKT points for this NLP. The KKT FONC says that if $\boldsymbol{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$ is a local minimizer and the NLP satisfies some regularity conditions, then there exists $\alpha \in \mathbb{R}$ such that $(\boldsymbol{a}, \alpha) \in \mathcal{K}$. The regularity condition that we will use is the LCQ condition, which says if h is an affine function, then no other condition is needed (here, h is affine, so the LCQ regularity condition is satisfied). Now observe that

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} -x_2x_3 \\ -x_1x_3 \\ -x_1x_2 \end{pmatrix} \quad \text{and} \quad \nabla h(\boldsymbol{x}) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Thus (2) can be rewritten as:

$$\begin{aligned} -x_2x_3 + \lambda &= 0 \\ -x_1x_3 + \lambda &= 0 \\ -x_1x_2 + \lambda &= 0 \\ x_1 + x_2 + x_3 - 9 &= 0 \end{aligned} \tag{3}$$

Problem 1.b

Exercise 2. Find all solutions $(\boldsymbol{x}, \lambda)$ to the KKT FONC for this NLP.

Solution 2. Let $(\boldsymbol{x}, \lambda) \in \mathcal{K}$. First we use the computer algebra system Singular to give us a "nicer" set of equations which is equivalent to (3):

```
ring A=0,(x1,x2,x3,lambda),lp;
ideal I = lambda-x2*x3,lambda-x1*x3,lambda-x1*x2,x1+x2+x3-9;
std(I);

_[1]=lambda^2-9*lambda
_[2]=x3*lambda-3*lambda
_[3]=x3^2-9*x3+2*lambda
_[4]=2*x2*lambda+x3*lambda-9*lambda
_[5]=x2*x3-lambda _
_[6]=x2^2+x2*x3-9*x2+lambda
_[7]=x1+x2+x3-9
```

Singular tells us that the system of equations (3) is equivalent to the following system of equations

$$
\begin{aligned}
(\lambda-9)\lambda &= 0\\
(x_3-3)\lambda &= 0\\
x_3^2-9x_3+2\lambda &= 0\\
(2x_2+x_3-9)\lambda &= 0\\
x_2x_3-\lambda &= 0\\
x_2^2+x_2x_3-9x_2+\lambda &= 0\\
x_1+x_2+x_3-9 &= 0
\end{aligned}
\tag{4}
$$

What's nice about the system of equations (4) is that we immediately see from $\lambda(\lambda-9)=0$ that either $\lambda=0$ or $\lambda=9$. First suppose $\lambda=9$. Then from $\lambda=9$, the equation $(x_3-3)\lambda=0$ implies $x_3=3$. Then from $\lambda=9$ and $x_3=3$, the equation $x_2x_3-\lambda=0$ implies $x_2=3$. Finally from $\lambda=9$, $x_3=3$, and $x_2=3$, the equation $x_1+x_2+x_3-9=0$ implies $x_1=3$. This produces the following solution $(\boldsymbol{a},\alpha)\in\mathcal{K}$ where we set

$$(\boldsymbol{a},\alpha)=(a_1,a_2,a_3,\alpha)=(3,3,3,0)$$

Now suppose that $\lambda=0$. Then the system of equations (3) reduces to the following system of equations

$$
\begin{aligned}
x_2x_3 &= 0\\
x_1x_3 &= 0\\
x_1x_2 &= 0\\
x_1+x_2+x_3 &= 9
\end{aligned}
\tag{5}
$$

Assume that $x_3\neq 0$. Then from the equations $x_2x_3=0$ and $x_1x_3=0$, we see that $x_1=0=x_2$. Then from $x_1=0=x_2$, the equation $x_1+x_2+x_3=9$ implies $x_3=9$. A similar argument shows that if $x_2\neq 0$, then $x_1=0=x_3$ and $x_2=9$, and if $x_1\neq 0$, then $x_2=0=x_3$ and $x_1=9$. Note that we can't have $x_1=x_2=x_3=0$ since this would violate the equation $x_1+x_2+x_3=9$. Thus we have produced all solutions in this case, they are $(\boldsymbol{b},\beta),(\boldsymbol{c},\gamma),(\boldsymbol{d},\delta)\in\mathcal{K}$ where we set

$$
\begin{aligned}
(\boldsymbol{b},\beta) &= (9,0,0,0)\\
(\boldsymbol{c},\gamma) &= (0,9,0,0)\\
(\boldsymbol{d},\delta) &= (0,0,9,0)
\end{aligned}
$$

Altogether we have $\mathcal{K}=\{(\boldsymbol{a},\alpha),(\boldsymbol{b},\beta),(\boldsymbol{c},\gamma),(\boldsymbol{d},\delta)\}$.

Problem 1.c

Exercise 3. Find all optimal solutions to this NLP. Give their properties (choose from local, global, strict local, unique global).

Solution 3. We only need to check the points $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, and $\boldsymbol{d}$ by the KKT FONC. If we set $x_1=-x_2-x_3+9$ then

$$
\begin{aligned}
f(\boldsymbol{x}) &= -x_1x_2x_3\\
&= -(9-x_2-x_3)x_2x_3.
\end{aligned}
$$

In particular, the constrained NLP (1) is equivalent to the unconstrained NLP below:

$$
\begin{aligned}
&\text{minimize} && \widetilde{f}(\widetilde{\boldsymbol{x}})=-(9-\widetilde{x}_1-\widetilde{x}_2)\widetilde{x}_1\widetilde{x}_2\\
&\text{subject to} && \widetilde{\boldsymbol{x}}\in\mathbb{R}^2
\end{aligned}
\tag{6}
$$

where $\widetilde{\boldsymbol{x}}=(\widetilde{x}_1,\widetilde{x}_2)=(x_2,x_3)$. Thus it suffices to find the optimal solutions of the new NLP (6) (and classify them accordingly). In these new coordinates, the KKT points become

$$
\begin{aligned}
\widetilde{\boldsymbol{a}} &= (3,3)\\
\widetilde{\boldsymbol{b}} &= (0,0)\\
\widetilde{\boldsymbol{c}} &= (9,0)\\
\widetilde{\boldsymbol{d}} &= (0,9)
\end{aligned}
$$

Observe that

$$\nabla\widetilde{f}(\widetilde{\boldsymbol{x}}) = \begin{pmatrix} \widetilde{x}_2(2\widetilde{x}_1+\widetilde{x}_2-9) \\ \widetilde{x}_1(\widetilde{x}_1+2\widetilde{x}_2-9) \end{pmatrix} \quad \text{and} \quad \mathrm{H}_{\widetilde{f}}(\widetilde{\boldsymbol{x}}) = \begin{pmatrix} 2\widetilde{x}_2 & 2\widetilde{x}_1+2\widetilde{x}_2-9 \\ 2\widetilde{x}_1+2\widetilde{x}_2-9 & 2\widetilde{x}_1 \end{pmatrix}$$

In particular $\widetilde{\boldsymbol{a}}$, $\widetilde{\boldsymbol{b}}$, $\widetilde{\boldsymbol{c}}$, and $\widetilde{\boldsymbol{d}}$ are all critical points and

$$\mathrm{H}_{\widetilde{f}}(\widetilde{\boldsymbol{a}}) = \begin{pmatrix} 6 & 3 \\ 3 & 6 \end{pmatrix}, \quad \mathrm{H}_{\widetilde{f}}(\widetilde{\boldsymbol{b}}) = \begin{pmatrix} 0 & -9 \\ -9 & 0 \end{pmatrix}, \quad \mathrm{H}_{\widetilde{f}}(\widetilde{\boldsymbol{c}}) = \begin{pmatrix} 0 & 9 \\ 9 & 9 \end{pmatrix}, \text{and} \quad \mathrm{H}_{\widetilde{f}}(\widetilde{\boldsymbol{d}}) = \begin{pmatrix} 9 & 9 \\ 9 & 0 \end{pmatrix}.$$

Only $\mathrm{H}_{\widetilde{f}}(\widetilde{\boldsymbol{a}})$ is positive semidefinite (in fact positive definite since its eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 9$ are strictly positive). The other matrices are mixed definite (each has one strictly positive eigenvalue and one strictly negative eigenvalue). It follows that $\widetilde{\boldsymbol{a}}$ is a *strict* local minimum whereas the other points are all saddle points. Note that $\widetilde{\boldsymbol{a}}$ can't be a global minimum since $\widetilde{f}$ takes arbitrary small values (for instance $\widetilde{f}(t,t) = -(9-2t)t^2 = 2t^3 +$ lower terms in t, so taking $t \to -\infty$ gives us $\widetilde{f}(t,t) \to -\infty$). In other words, $\boldsymbol{a}$ is a strict local optimal solution for the NLP (1) (and the only one at that).

Problem 2

For this problem, let $f(\boldsymbol{x}) = x_1 - x_2$ and let $h(\boldsymbol{x}) = x_1^2 + x_2^2 - 1$. We consider the following NLP:

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & h(\boldsymbol{x}) = 0 \end{aligned}$$

Problem 2.a

Exercise 4. Write the KKT FONC to this NLP.

Solution 4. A KKT point for this NLP is a pair $(\boldsymbol{x}, \lambda) = (x_1, x_2, \lambda) \in \mathbb{R}^4$ which satisfies the following:

$$\begin{aligned} \nabla f(\boldsymbol{x}) + \lambda \nabla h(\boldsymbol{x}) &= 0 && \text{Stationary} \\ h(\boldsymbol{x}) &= 0 && \text{Primal Feasibility} \end{aligned} \tag{7}$$

Let $\mathcal{K} \subseteq \mathbb{R}^3$ denote the set of all KKT points for this NLP. The KKT FONC says that if $\boldsymbol{a} = (a_1, a_2) \in \mathbb{R}^2$ is a local minimizer and the NLP satisfies some regularity conditions, then there exists $\alpha \in \mathbb{R}$ such that $(\boldsymbol{a}, \alpha) \in \mathcal{K}$. The regularity condition that we will use is the LICQ condition, which says that the gradients of the active inequality constraints and the gradients of the equality constraints are linearly independent at $\boldsymbol{a}$. What this means in this case is that if $h(\boldsymbol{a}) = 0$, then $\nabla h(\boldsymbol{a}) \neq 0$. Now observe that

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \nabla h(\boldsymbol{x}) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}.$$

In particular, we see that the regularity condition is always satisfied since $\nabla h(\boldsymbol{a}) = 0$ if and only if $\boldsymbol{a} = (0,0)$, and $h(0,0) \neq 0$. Now (2) can be rewritten as:

$$\begin{aligned} 2\lambda x_1 + 1 &= 0 \\ 2\lambda x_2 - 1 &= 0 \\ x_1^2 + x_2^2 &= 1 \end{aligned} \tag{8}$$

Problem 2.b

Exercise 5. Find all solutions $(\boldsymbol{x}, \lambda)$ to the KKT FONC for this NLP.

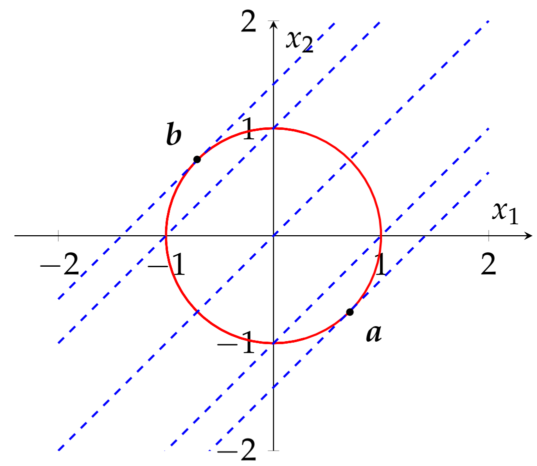
Solution 5. Let $(\boldsymbol{x}, \lambda) \in \mathcal{K}$. Then adding the first two equations in (8) together gives us $4\lambda(x_1 + x_2) = 0$. It follows that either $\lambda = 0$ or $x_1 = -x_2$. We can't have $\lambda = 0$ since this would contradict the equation $2\lambda x_1 + 1 = 0$. Thus we must have $x_1 = -x_2$. From $x_1 = -x_2$, the third equaiton in (8) becomes $2x_1^2 = 1$. Thus we must have $x_1 = \pm\sqrt{2}/2$, and since $x_2 = -x_1$, this implies $x_2 = \mp\sqrt{2}/2$. Finally, From first equation in (8), we see that $\lambda = \mp\sqrt{2}/2$. Thus we have produced all solutions in this case, they are $\mathcal{K} = \{(\boldsymbol{a}, \alpha), (\boldsymbol{b}, \beta)\}$ where we set

$$\begin{aligned} (\boldsymbol{a}, \alpha) &= (\sqrt{2}/2, -\sqrt{2}/2, -\sqrt{2}/2) \\ (\boldsymbol{b}, \beta) &= (-\sqrt{2}/2, \sqrt{2}/2, \sqrt{2}/2) \end{aligned}$$

Problem 2.c

Exercise 6. Find all optimal solutions to this NLP. Give their properties (choose from local, global, strict local, unique global).

Solution 6. We solve this problem geometrically:



The feasible region $\{h = 0\}$ is given by the red circle above. The blue dashed lines above are level sets $\{f = c\}$ for various values of $c \in \mathbb{R}$ (namely for $c = -\sqrt{2}, -1, 0, 1, \sqrt{2}$). We also plotted the points $\boldsymbol{a}$ and $\boldsymbol{b}$ because we know that these are the only possible points that can be local optimizers. Note that the vector $-\nabla f(\boldsymbol{b}) = (-1, 1)^\top$ is normal to the circle $\{h = 0\}$ at the point $\boldsymbol{b}$. This implies that the level set $\{f = -\sqrt{2}\}$ is tangent to the circle $\{h = 0\}$, which implies $\boldsymbol{b}$ is a strict local minimizer with optimal objective value $f(\boldsymbol{b}) = -\sqrt{2}$ (in fact $\boldsymbol{b}$ is unique global minimizer). Similarly, that the vector $\nabla f(\boldsymbol{a}) = (1, -1)^\top$ is normal to the circle $\{h = 0\}$ at the point $\boldsymbol{a}$. This implies that the level set $\{f = \sqrt{2}\}$ is tangent to the circle $\{h = 0\}$, which implies $\boldsymbol{a}$ is a strict local maximizer with optimal objective value $f(\boldsymbol{a}) = \sqrt{2}$ (in fact $\boldsymbol{a}$ is unique global maximizer). Note however that in the NLP we consider is a minimization problem, so $\boldsymbol{b}$ is the optimal solution whereas $\boldsymbol{a}$ is not.

Problem 3

For this problem, let $f(\boldsymbol{x}) = 4x_1^2 - 9x_1 + x_2^2$. We consider the following unconstrained problem

$$\begin{aligned} &\text{minimize} && f(\boldsymbol{x}) \\ &\text{subject to} && \boldsymbol{x} \in \mathbb{R}^2. \end{aligned}$$

Exercise 7. Let $\boldsymbol{x}^0 = (1, 1)^\top$ be the starting point for the method of steepest descent. Use the algorithm to find $\boldsymbol{x}^1$.

Solution 7. 1. We set $\boldsymbol{x}^1 = \boldsymbol{x}^0 - \gamma \nabla f(\boldsymbol{x}^0)$ where the step size $\gamma > 0$ is chosen to minimize $f(\boldsymbol{x}^0 - \gamma \nabla f(\boldsymbol{x}^0))$. We first we calculate

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 8x_1 - 9 \\ 2x_2 \end{pmatrix}.$$

In particular, we have $\nabla f(\boldsymbol{x}^0) = (-1, 2)^\top$. Thus we have

$$\begin{aligned} \boldsymbol{x}^1 &= \boldsymbol{x}^0 - \gamma \nabla f(\boldsymbol{x}^0) \\ &= (1, 1)^\top - \gamma(-1, 2)^\top \\ &= (1 + \gamma, 1 - 2\gamma)^\top \end{aligned}$$

Now to find γ, we first calculate

$$\begin{aligned} f(\boldsymbol{x}^1) &= f(\boldsymbol{x}^0 - \gamma \nabla f(\boldsymbol{x}^0)) \\ &= f((1 + \gamma, 1 - 2\gamma)^\top) \\ &= 4(1 + \gamma)^2 - 9(1 + \gamma) + (1 - 2\gamma)^2 \\ &= 8\gamma^2 - 5\gamma - 4 \\ &= \frac{1}{32}(16\gamma - 5)^2 - \frac{153}{32}. \end{aligned}$$

Clearly this is minimized when $\gamma = 5/16$ with objective value being $f(\boldsymbol{x}^1) = -153/32$. Thus we have

$$\boldsymbol{x}^1 = \left(\frac{21}{16}, \frac{3}{8}\right)^\top.$$

Problem 4

Exercise 8. Keep the same notation as in question 3.

1. Find $\widetilde{\boldsymbol{x}}^1$ for the unconstrained problem in question 3 above with the Newton Method.
2. Compare the results you obtained with the steepest descent method and Newton method.
3. Did you obtain an optimal solution? Explain.

Solution 8. 1. First we calculate

$$\mathrm{H}_f(\boldsymbol{x}) = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}.$$

Next we set

$$\begin{aligned}
\widetilde{\boldsymbol{x}}^1 &= \boldsymbol{x}^0 - \mathrm{H}_f^{-1}(\boldsymbol{x}^0)\nabla f(\boldsymbol{x}^0) \\
&= \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1/8 & 0 \\ 0 & -1/2 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1/8 \\ -1 \end{pmatrix} \\
&= \begin{pmatrix} 9/8 \\ 0 \end{pmatrix}.
\end{aligned}$$

Thus $\widetilde{\boldsymbol{x}}^1 = (9/8, 0)^\top$ with objective value being $f(\widetilde{\boldsymbol{x}}^1) = -81/16$.

2. Notice that

$$f(\widetilde{\boldsymbol{x}}^1) < f(\boldsymbol{x}^1),$$

so Newton's method produced a better solution than the descent method (which is unsurprising as Newton's method involves the Hessian wheras the descent method only involves the gradient).

3. Yes: Observe we can rewrite $f(\boldsymbol{x})$ as

$$\begin{aligned}
f(\boldsymbol{x}) &= 4x_1^2 - 9x_1 + x_2^2 \\
&= \left(\frac{8x_1 - 9}{4}\right)^2 + x_2^2 - \frac{81}{16} \\
&= \widetilde{x}_1^2 + \widetilde{x}_2^2 - \frac{81}{16} \\
&= f(\widetilde{\boldsymbol{x}}),
\end{aligned}$$

where we used the change of variables $\widetilde{x}_1 = 2x_1 - 9/4$ and $\widetilde{x}_2 = x_2$. Now observe that since $f(\widetilde{\boldsymbol{x}})$ is a sum of squares plus a constant, we have $f(\widetilde{\boldsymbol{x}}) > -81/16$ for all $\widetilde{\boldsymbol{x}} \neq (0,0)$ and we have $f(\widetilde{\boldsymbol{x}}) = -81/16$ if and only if $\widetilde{\boldsymbol{x}} = (0,0)$. Thus $f(\widetilde{\boldsymbol{x}})$ has a global minimum at $\widetilde{\boldsymbol{x}} = (0,0)^\top$, that is $\widetilde{x}_1 = 0$ and $\widetilde{x}_2 = 0$ with optimal objective value being $-81/16$. In the $\boldsymbol{x}$-coordinates, this is given by

$$x_1 = 2\widetilde{x}_1 + 9/8 = 9/8 \quad \text{and} \quad x_2 = \widetilde{x}_2 = 0,$$

thus $f(\boldsymbol{x})$ has a global minimum at $\boldsymbol{x} = (9/8, 0) = \widetilde{\boldsymbol{x}}^1$, which is exactly the point we found using Newton's method. (we could have also concluded this by observing that $\nabla f(\boldsymbol{x}) = 0$ if and only if $\boldsymbol{x} = \widetilde{\boldsymbol{x}}^1$ and then combining this with the fact that the Hessian is positive definite everywhere).

Problem 5

For this problem, let $f, g_1, \dots, g_m \colon \mathbb{R}^n \to \mathbb{R}$. We consider the following NLP:

$$\begin{array}{lll} \text{minimize} & f(\boldsymbol{x}) & \\ \text{subject to} & g_i(\boldsymbol{x}) \leq 0 & \qquad\qquad 1 \leq i \leq m \end{array}$$

(this is how the book sets it up and hence what this problem is modelled off of).

Exercise 9. Answer whether each of the following statements is true or false:

1. The Lagrange multiplier corresponding to an inactive constraint is always zero.
2. The Lagrange multiplier corresponding to an active constraint is never zero.
3. A positive semidefinite matrix is never singular
4. A positive definite matrix is never singular.
5. If g_i are differentiable for $i = 1, \dots, m$ and $\overline{\boldsymbol{x}}$ is a local minimum, there exists a vector $\overline{\boldsymbol{\mu}}$ such that $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\mu}})$ satisfies the KKT conditions.
6. If there exists a vector $\overline{\boldsymbol{\mu}}$ such that $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\mu}})$ satisfies the KKT conditions and if f and g_i are differentiable convex functions for $i = 1, \dots, m$, then $\overline{\boldsymbol{x}}$ is a global minimum.

Solution 9. 1. True. By complementary slackness, we must have $\mu_i g_i(\boldsymbol{x}) = 0$. If g_i is inactive at $\boldsymbol{x}$, then this implies $g_i(\boldsymbol{x}) \neq 0$ which implies $\mu_i = 0$.

2. False. Complementary slackness only tells us that $\mu_i g_i(\boldsymbol{x}) = 0$. If g_i is active at $\boldsymbol{x}$, then this means $g_i(\boldsymbol{x}) = 0$, but this doesn't necessarily imply that $\mu_i \neq 0$.

3. False. Indeed, a Hermitian matrix A is positive semidefinite if and only its eigenvalues $\lambda_1, \dots, \lambda_n$ are all greater than or equal to zero. In particular, if one of its eigenvalues is zero, then the matrix will be singular (since $\det A = \lambda_1 \cdots \lambda_n = 0$).

4. True. Indeed, a Hermitian matrix A is positive definite if and only its eigenvalues are all strictly greater than zero. In particular, we have $\det A = \lambda_1 \cdots \lambda_n > 0$, and thus A is not singular.

5. False. We aslo need regularity conditions to hold (LCQ, LICQ, etc...) in order to apply the KKT FONC. If in addition a regularity condition is satisisfied, then this would be true.

6. True. In this case, we can use the KKT SOSC to conclude that $\overline{\boldsymbol{x}}$ is a global minimum.