Optimization Methods
Lecture 12

Solmaz S. Kia
Mechanical and Aerospace Engineering Dept.
University of California Irvine
solmaz@uci.edu

Consult: pages 276-297 (section 3.1 and 3.2) and sections 3.3, 3.3.1, 3.3.3 from Ref[1]

Lagrangian function $L : \mathbb{R}^{n+m} \mapsto \mathbb{R}$: $L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x)$

Proposition (Lagrange Multiplier Theorem-Necessary conditions)

Let $x^\star$ be a local minimum of f subject to $h(x) = 0$ and assume that the constraint gradients $\{\nabla h_1(x^\star), \cdots, \nabla h_m(x)\}$ are linearly independent. Then there exists a unique vectors $\lambda^\star = (\lambda_1^\star, \cdots, \lambda_m^\star)$ called Lagrange multiplier vector, s.t.

$$\nabla_x L(x^\star, \lambda^\star) = 0.$$

If in addition f and h are twice continuously differentiable we have

$$y^\top \nabla_{xx} L(x^\star, \lambda^\star)\, y \geqslant 0, \quad \forall y \in V(x^\star)$$

where $V(x^\star)$ is the space of first order feasible variations, i.e.,

$$V(x^\star) = \{d \in \mathbb{R}^n \mid \nabla h_i(x^\star)^\top d = 0\}.$$

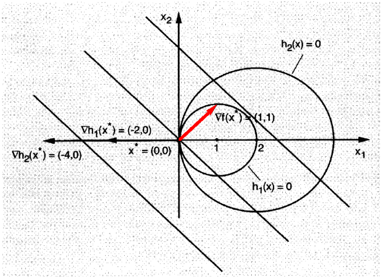
$$h(x^\star) = 0 \Leftrightarrow \nabla_\lambda L(x^\star, \lambda^\star) = 0.$$

A Problem with no Lagrange Multipliers: regularity of optimal point

- **Regular point of a set of constraints**: A feasible vector x for which the constraint gradients $\{\nabla h_1(x), \cdots, \nabla h_m(x)\}$ are linearly independent.
- For a local minimum that is not regular, there may not exist Lagrange multipliers.

$$\text{minimize } f(x) = x_1 + x_2, \quad \text{s.t.}$$
$$h_1(x) = (x_1 - 1)^2 + x_2^2 - 1 = 0, \quad h_2(x) = (x_1 - 2)^2 + x_2^2 - 4 = 0.$$

- $x^\star$ is not regular. Therefore, this problem cannot be solved using Lagrange multiplier theorem.
- $\nabla f(x^\star)$ cannot be written as linear combination of $\nabla h_1(x^\star)$ and $\nabla h_2(x^\star)$



Proofs for the Lagrange multiplier theorem

- **The elimination approach**:
  - We view the constraints as a system of m equations with n unknowns (recall that $m < n$).
  - We express m of the variables in terms of the remaining $n - m$ to reduce the problem to an unconstrained problem
  - Apply the corresponding first and second order necessary conditions for unconstrained minima: the Lagrange multiplier theorem follows.
  - We use the implicit function theorem here.
- **The penalty approach**:
  - We disregard the constraints, while adding to the cost a high *penalty* for violating them.
  - By Writing the necessary conditions for the “penalized" unconstrained problems, and by passing to the limit as the penalty increases we obtain the Lagrange multiplier theorem.

The regularity condition is crucial for the proof

Penalty approach for proof of necessary conditions for optimality

(C-OPT):

$$\begin{cases} \text{minimize } f(x) \text{ s.t.} \\ \quad h(x) = 0 \end{cases}$$

Penalty approach:

$$\begin{cases} x_k = \underset{x}{\operatorname{argmin}} F_k(x) = f(x) + \frac{k}{2}\|h(x)\|^2 + \frac{\alpha}{2}\|x - x^\star\|^2, \quad \text{s.t.} \\ \quad x \in S = \{x \in \mathbb{R}^n \,|\, \|x - x^\star\| \leqslant \epsilon\} \end{cases}$$

- $\frac{k}{2}\|h(x)\|^2$: imposes a penalty for violating the constraint $h(x) = 0$.
- $\frac{\alpha}{2}\|x - x^\star\|^2$: introduced for technical related reasons (to ensure $x^\star$ is a strict local minimum of function $f(x) + \frac{\alpha}{2}\|x - x^\star\|^2$ subject to $h(x) = 0$.
- $\epsilon > 0$ is chosen to be small and also such that for all $x \in S \cap \{x \in \mathbb{R}^n | h(x) = 0\}$ we have $f(x) \geqslant f(x^\star)$
- Weierstrass theorem guarantees that x_k exists for all $k \in \mathbb{R}_{\geqslant 0}$.

Let $\bar{x}$ be a limit point of $\{x_k\}$:
Some observations:

- $F_k(x_k) \leqslant F_k(x^\star) = f(x^\star)$:
  - $f(x_k) + \frac{\alpha}{2}\|x_k - x^\star\|^2 \leqslant f(x^\star) \Rightarrow f(\bar{x}) + \frac{\alpha}{2}\|\bar{x} - x^\star\|^2 \leqslant f(x^\star)$
- $\lim_{k\to\infty} \|h(x_k)\| = 0$ (since $f(x_k)$ is bounded in S):
  - every limit point $\bar{x}$ of $\{x_k\}$ satisfies $h(\bar{x}) = 0$
- $\bar{x} \in S$ and it feasible, we have $f(x^\star) \leqslant f(\bar{x}) \|\bar{x} - x^\star\| = 0 \Rightarrow \bar{x} = x^\star$
- $\bar{x} = x^\star$ is an interior point of S (the constraint is not active), therefore $x^\star$ is a local minimizer of unconstrained optimization problem $x^\star = \underset{x}{\operatorname{argmin}} F_k(x)$ when $k \to \infty$.

Penalty approach for proof of necessary conditions for optimality

$$x^\star = \underset{x}{\operatorname{argmin}} F_k(x) \text{ when } k \to \infty$$

FONC:

$$\nabla F(x_k) = \nabla f(x_k) + k\nabla h(x_k)h(x_k) + \alpha(x_k - x^\star) = 0, \quad k \to \infty$$

- Under the assumption that $x^\star$ is a regular point, i.e., $\nabla h(x^\star)$ is full column rank: $(\nabla h(x^\star)^\top \nabla h(x^\star))$ is invertible
- $\lim_{k\to\infty} kh_i(x_k) = \lambda_i^\star$
- $\lambda^\star = -(\nabla h(x^\star)^\top \nabla h(x^\star))^{-1} \nabla h(x^\star)^\top \nabla f(x^\star)$

Then we get $f(x^\star) + \lambda^\star \nabla h(x^\star) = 0$

SONC:

$$\nabla^2 F(x_k) = \nabla^2 f(x_k) + k\nabla h(x_k)\nabla h(x_k)^\top + k\sum_{i=1}^{m} h_i(x_k)\nabla^2 h_i(x_k) + \alpha I$$

$\vdots$

$$y^\top \left(\nabla^2 f(x^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla^2 h_i(x^\star)\right) y \geqslant 0, \quad y \in V(x^\star)$$

Proposition (Second Order Sufficiency Conditions for Optimality)

Assume that f and h are twice continuously differentiable, and let $x^\star \in \mathbb{R}^n$ and $\lambda^\star \in \mathbb{R}^m$ satisfy

$$\nabla_x L(x^\star, \lambda^\star) = 0, \quad \nabla_\lambda L(x^\star, \lambda^\star) = 0,$$
$$y^\top \nabla_{xx} L(x^\star, \lambda^\star)\, y > 0, \quad \forall\, y \neq 0 \text{ with } \nabla h(x^\star)^\top y = 0.$$

Then $x^\star$ is a strict local minimum of f subject to $h(x) = 0$. In fact, there exists scalars $\gamma > 0$ and $\epsilon > 0$ such that

$$f(x) \geqslant f(x^\star) + \frac{\gamma}{2}\|x - x^\star\|, \quad \forall x \text{ with } h(x) = 0 \text{ and } \|x - x^\star\| < \epsilon.$$

Constrained optimization

$$x^\star = \underset{x\in\mathbb{R}^n}{\text{argmin}}\ f(x) \quad \text{s.t.}$$
$$h_i(x) = 0, \quad i \in \{1,\cdots,m\}$$
$$g_i(x) \leqslant 0, \quad i \in \{1,\cdots,r\}$$

or

$$x^\star = \underset{x\in\mathbb{R}^n}{\text{argmin}}\ f(x) \quad \text{s.t.}$$
$$h(x) = 0,$$
$$g(x) \leqslant 0,$$

f,h,g: continuously differentiable function of x
e.g., $f, h, g \in C^1$ continuously differentiable
e.g., $f, h, g \in C^2$ both f and its first derivative are continuously differentiable

First Order Necessary Condition for Optimality: $x^\star$ is a local minimizer then

$$\nabla f(x^\star)^\top \Delta x \geqslant 0, \qquad \text{for} \;\; \Delta x \in V(x^\star)$$

- Set of first order feasible variations at x

$$V(x) = \{d \in \mathbb{R}^n \mid \nabla h_i(x)^\top d = 0,\; \nabla g_j(x)^\top d \leqslant 0, \quad j \in A(x^\star)\}$$

- Active inequality constraints at x

$$A(x) = \{j \in \{1,\cdots,r\} \mid g_j(x) = 0\}$$

A feasible vector x is said to be regular of the equality constraint gradients $\nabla h_i(x)$, $i = 1,\cdots,m$, and the active inequality constraint gradients $\nabla g_j(x)$, $j \in A(x)$, are linearly independent.

Necessary Conditions for Optimality: equality and inequality conditions

$$x^\star = \underset{x\in\mathbb{R}^n}{\text{argmin}}\ f(x) \quad \text{s.t.}$$
$$h_i(x) = 0, \quad i \in \{1, \cdots, m\}$$
$$g_j(x) \leqslant 0, \quad j \in \{1, \cdots, r\}$$

or

$$x^\star = \underset{x\in\mathbb{R}^n}{\text{argmin}}\ f(x) \quad \text{s.t.}$$
$$h(x) = 0,$$
$$g(x) \leqslant 0,$$

- A simple approach relies on the theory for equality constraints:
  - Inactive constraints at $x^\star$ do not matter, they can be ignored in the statement of optimality conditions
  - Active inequality constraints can be treated to a large extent as equality constraints

$x^\star$ is also a local minimum of

$$x^\star = \underset{x\in\mathbb{R}^n}{\text{argmin}}\ f(x) \quad \text{s.t.}$$
$$h_i(x) = 0, \quad i \in \{1, \cdots, m\}$$
$$g_j(x) = 0, \quad \forall j \in A(x^\star)$$

If $x^\star$ is regular for this equivalent optimization problem, then there exists Lagrange multipliers $\lambda_1^\star, \cdots, \lambda_{,}^\star$, and $\mu_j^\star$, $j \in A(x^\star)$:

$$\nabla f(x^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla h_i(x^\star) + \sum_{j\in A(x^\star)} \mu_j^\star \nabla g_j(x^\star) = 0.$$

But we need to require that $\mu_j^\star \geqslant 0$ for $j \in A(x^\star)$.

This approach is limited by regularity condition!

Necessary Conditions for Optimality: equality and inequality conditions

Lagrangian function $L : \mathbb{R}^{n+m} \mapsto \mathbb{R}$: $L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{i=1}^{m} \mu_j g_j(x)$

Proposition (Karush-Huhn-Tucker Necessary conditions)

Let $x^\star$ be a local minimum of

$$x^\star = \underset{x\in\mathbb{R}^n}{\text{argmin}}\ f(x) \quad \text{s.t.}$$
$$h_1(x) = 0, \cdots, h_m(x) = 0$$
$$g_1(x) \leqslant 0, \cdots, g_r(x) \leqslant 0$$

where f, h_i and g_j are continuously differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}$. Assume the $x^\star$ is regular. Then there exists unique Lagrange multiplier vectors $\lambda^\star = (\lambda_1^\star, \cdots, \lambda_m^\star)$ and $\mu^\star = (\mu_1^\star, \cdots, \mu_r^\star)$, s.t.

$$\nabla_x L(x^\star, \lambda^\star, \mu^\star) = 0$$
$$\mu_j^\star \geqslant 0, \quad j = 1, \cdots, r$$
$$\mu_j^\star = 0, \quad \forall j \notin \underbrace{A(x^\star)}_{\text{active constraint set}}.$$

If in addition f g and h are twice continuously differentiable we have

$$y^\top \nabla_{xx} L(x^\star, \lambda^\star, \mu^\star)\, y \geqslant 0,$$

for all
$y \in V(x^\star) = \{y \in \mathbb{R}^n | h_i(x^\star)^\top y = 0, \quad \forall i = 1, \cdots, m, \quad \nabla g_j(x^\star)^\top y = 0, \quad j \in A(x^\star)\}.$

Penalty approach for proof of necessary conditions for optimality

(C-OPT):

$$\begin{cases}\text{minimize } f(x) \text{ s.t.} \\ h(x) = 0\end{cases}$$

Penalty approach:

$$\begin{cases} x_k = \underset{x}{\operatorname{argmin}}\, F_k(x) = f(x) + \frac{k}{2}\|h(x)\|^2 + \frac{k}{2}\sum_{j=1}^{r}(g_j^+(x))^2 + \frac{\alpha}{2}\|x - x^\star\|^2, \\ x \in S = \{x \in \mathbb{R}^n \,|\, \|x - x^\star\| \leqslant \epsilon\} \end{cases}$$

- $\frac{k}{2}\|h(x)\|^2$: imposes a penalty for violating the constraint $h(x) = 0$.
- $g_j^+(x) = \max\{0, g_j(x)\}$, $j = 1, \cdots, r$: penalizes violating the constraint $g_j(x) \leqslant 0$.
- $\frac{\alpha}{2}\|x - x^\star\|^2$: introduced for technical related reasons (to ensure $x^\star$ is a strict local minimum of function $f(x) + \frac{\alpha}{2}\|x - x^\star\|^2$ subject to $h(x) = 0$.
- $\epsilon > 0$ is chosen to be small and also such that for all $x \in S \cap \{x \in \mathbb{R}^n | h(x) = 0, g(x) \leqslant 0\}$ we have $f(x) \geqslant f(x^\star)$
- Weierstrass theorem guarantees that x_k exists for all $k \in \mathbb{R}_{\geqslant 0}$.

Analysis results are similar to the one for the equality constraint in the earlier slides.

- $x^\star$ being regular is essential for proof
- $$\lambda_i^\star = \lim_{t\to\infty} k h_i(x_k), \quad i = 1, \cdots, m,$$
  $$\mu_j^\star = \lim_{t\to\infty} k g_j^+(x_k), \quad i = j, \cdots, r.$$

  Since $g_j^+(x) \geqslant 0$, we obtain $\mu_j^\star \geqslant 0$ for all j.

Sufficiency Conditions for Optimality

Lagrangian function $L : \mathbb{R}^{n+m} \mapsto \mathbb{R}$: $L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{i=1}^{m} \mu_j g_j(x)$

Second Order Sufficiency Conditions

Assume that f, h_i and g_j are twice continuously differentiable f, and let $x^\star \in \mathbb{R}^n$, $\lambda^\star = (\lambda_1^\star, \cdots, \lambda_m^\star)$ and $\mu^\star = (\mu_1^\star, \cdots, \mu_r^\star)$ satisfy

$$\begin{aligned}
&\nabla_x L(x^\star, \lambda^\star, \mu^\star) = 0, \quad h(x^\star) = 0_m, \\
&\mu_j^\star \geqslant 0, \quad j = 1, \cdots, r, \\
&\mu_j^\star = 0, \quad \forall j \notin A(x^\star), \\
&y^\top \nabla_{xx} L(x^\star, \lambda^\star, \mu^\star) y > 0,
\end{aligned}$$

for all $y \in \mathbb{R}^n$ such that $h_i(x^\star)^\top y = 0, \quad \forall i = 1, \cdots, m, \quad \nabla g_j(x^\star)^\top y = 0, \quad j \in A(x^\star)$.
Assume also that

$$\mu_j^\star > 0, \qquad \forall j \in A(x^\star).$$

Then $x^\star$ is a strict local minimum of

$$\begin{aligned}
&\min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s.t.} \\
&h_1(x) = 0, \cdots, h_m(x) = 0 \\
&g_1(x) \leqslant 0, \cdots, g_r(x) \leqslant 0
\end{aligned}$$

Solution approach

One approach for using necessary conditions to solve inequality constrained problems is to consider separately all the possible combinations of constraints being active or inactive.

Constrained optimization: numerical example

$$\text{minimize } f(x) = x_1 + x_2 \text{ subject to}$$
$$g(x) = (x_1 - 1)^2 + x_2^2 - 1 \leqslant 0$$

- H1: Constraint is active. To validate H1, we should have $\mu \geqslant 0$.

$$L(x, \mu) = x_1 + x_2 + \mu(x_1 - 1)^2 + x_2^2 \leqslant 1$$

FONC:

$$\left.\begin{array}{l} \nabla_{x_1} L(x, \mu) = 1 + 2\mu(x_1 - 1) = 0 \\ \nabla_{x_2} L(x, \mu) = 1 + 2\mu(x_2) = 0 \\ \nabla_{\mu} L(x, \mu) = (x_1 - 1)^2 + x_2^2 - 1 = 0 \end{array}\right\} \Rightarrow$$

$$\begin{cases} x_1 = 1, x_2 = 1, \mu = -\frac{1}{2} & \text{since } \mu < 0 \text{ this solution is not acceptable} \\ x_1^\star = 1, x_2^\star = -1, \mu^\star = \frac{1}{2} & \text{since } \mu^\star > 0 \text{ this solution is a candidate for local minimizer} \end{cases}$$

SONC:

$$y \nabla_{xx} L(x^\star, \mu^\star) y \geqslant 0 \text{ for } y \in V(x^\star) = \{y \in \mathbb{R}^2 | \nabla g(x^\star)^\top y = 0\} = \{y \in \mathbb{R}^2 | \begin{bmatrix} 0 & -2 \end{bmatrix} y = 0\}$$

Since $\nabla_{xx} L(x^\star, \mu^\star) = \begin{bmatrix} 2\mu^\star & 0 \\ 0 & 2\mu^\star \end{bmatrix} > 0$ ($\mu^\star = \frac{1}{2}$), then SONC condition is definitely satisfied. Also since the condition holds for strict > 0, then the second order sufficiency condition is satisfied and $x_1^\star = 1, x_2^\star = -1$ is a local minimizer.

- H2: Constraint is not active. To validate H2, we should check that the identified stationary points $x^\star$ satisfy $g(x^\star) < 0$.

$$\left.\begin{array}{l} \nabla_{x_1} f(x) = 1 = 0 \\ \nabla_{x_2} f(x) = 1 = 0 \end{array}\right\} \Rightarrow \text{there is no solution in this case}$$

Constrained optimization: numerical example

$$\text{minimize } f(x) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 \text{ subject to}$$
$$g_1(x) = x_1^2 + x_2^2 - 5 \leqslant 0$$
$$g_2(x) = 3x_1 + x_2 - 6 \leqslant 0$$

$$\nabla_x f(x) = \begin{bmatrix} 4x_1 + 2x_2 - 10 \\ 2x_1 + 2x_2 - 10 \end{bmatrix}, \quad \nabla_x g_1(x) = \begin{bmatrix} 2x_1 \\ 3 \end{bmatrix}, \quad \nabla_x g_2(x) = \begin{bmatrix} 2x_2 \\ 1 \end{bmatrix}$$

- H1: both constraints are inactive: $g_1 < 0$, $g_2 < 0$ and $\mu_1 = \mu_2 = 0$.
  FONC:
  $$\left.\begin{array}{l} \nabla_{x_1} f(x) = 4x_1 + 2x_2 - 10 = 0 \\ \nabla_{x_2} f(x) = 2x_1 + 2x_2 - 10 = 0 \end{array}\right\} \Rightarrow \; x_1 = 0, x_2 = 5$$
  $g_1(x_1 = 0, x_2 = 5) = 20 > 0$ and $g_2(x_1 = 0, x_2 = -1 < 0$. Since H1 is not correct, this case is not possible.

- H2: both constraints are active: $g_1 = 0$, $g_2 = 0$ and $\mu_1, \mu_2 \geqslant 0$.
  $$L(x, \mu) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 + \mu_1(x_1^2 + x_2^2 - 5) + \mu_2(3x_1 + x_2 - 6)$$
  FONC:
  $$\left.\begin{array}{l} \nabla_{x_1} L(x, \mu) = 4x_1 + 2x_2 - 10 + 2\mu_1 x_1 + 3\mu_2 = 0 \\ \nabla_{x_2} L(x, \mu) = 2x_1 + 2x_2 - 10 + 2\mu_2 x_2 + \mu_2 = 0 \\ \nabla_{\mu_1} L(x, \mu) = x_1^2 + x_2^2 - 5 = 0 \\ \nabla_{\mu_1} L(x, \mu) = 3x_1 + x_2 - 6 = 0 \end{array}\right\} \Rightarrow$$
  $$\left\{\begin{array}{ll} x = \begin{bmatrix} 2.1742 \\ -0.5225 \end{bmatrix}, \mu = \begin{bmatrix} -2.37 \\ 4.22 \end{bmatrix} & \text{since } \mu_1 < 0 \text{ this solution is not acceptable.} \\ x = \begin{bmatrix} 1.4258 \\ 1.7228 \end{bmatrix}, \mu = \begin{bmatrix} 1.37 \\ -1.02 \end{bmatrix} & \text{since } \mu_2 < 0 \text{ this solution is not acceptable.} \end{array}\right.$$

Constrained optimization: numerical example

- H3: g_1 is inactive ($g_1 < 0,\ \mu_1 = 0$), and g_2 is active ($\mu_2 \geqslant 0$).

$$L(x,\mu) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 + \mu_2(3x_1 + x_2 - 6)$$

FONC:

$$\left.\begin{array}{l} \nabla_{x_1} L(x,\mu) = 4x_1 + 2x_2 - 10 + 3\mu_2 = 0 \\ \nabla_{x_2} L(x,\mu) = 2x_1 + 2x_2 - 10 + \mu_2 = 0 \\ \nabla_{\mu_1} L(x,\mu) = 3x_1 + x_2 - 6 = 0 \end{array}\right\} \Rightarrow x = \begin{bmatrix} 0.4 \\ 0.8 \end{bmatrix},\ \mu_2 = -0.4.$$

since $\mu_2 < 0$ this solution is not acceptable.

- H4: g_2 is inactive ($g_2 < 0,\ \mu_2 = 0$), and g_1 is inactive ($\mu_1 \geqslant 0$).

$$L(x,\mu) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 + \mu_1(x_1^2 + x_2^2 - 5)$$

FONC:

$$\left.\begin{array}{l} \nabla_{x_1} L(x,\mu) = 4x_1 + 2x_2 - 10 + 2\mu_1 x_1 = 0 \\ \nabla_{x_2} L(x,\mu) = 2x_1 + 2x_2 - 10 + 2\mu_1 x_2 = 0 \\ \nabla_{\mu_1} L(x,\mu) = x_1^2 + x_2^2 - 5 = 0 \end{array}\right\} \Rightarrow x^\star = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ \mu_1^\star = 1.$$

since $\mu_1 \geqslant 0$ this solution is qualified as KKT solution.

Now we need to validate H4: $g_2(x_1 = 1, x_2 = 2) = -1 < 0$, therefore H4 is correct.
SONC:

$$y \nabla_{xx} L(x^\star, \mu^\star) y \geqslant 0 \text{ for } y \in V(x^\star) = \{y \in \mathbb{R}^2 | \nabla g_1(x^\star)^\top y = 0\} = \{y \in \mathbb{R}^2 | \begin{bmatrix} 2 & 4 \end{bmatrix} y = 0\}$$

Since $\nabla_{xx} L(x^\star, \mu^\star) = \begin{bmatrix} 4 + 2\mu_1^\star & 2 \\ 2 & 2 + 2\mu_1^\star \end{bmatrix} > 0$ ($\mu^\star = 1$), then SONC condition is definitely satisfied. Also since the condition holds for strict > 0, then the second order sufficiency condition is satisfied and $x_1^\star = 1, x_2^\star = 2$ is a local minimizer.