Otimização condicionada e não condicionada

Prof. Helson Gomes de Souza

1 Introdução

De maneira simplificada, a otimização pode ser entendida como um conjunto de técnicas matemáticas que permitem obter os valores extremos de uma dada função. Estas técnicas formam a base da teoria econômica e permitem teorizar de maneira formal os problemas dos agentes de uma dada economia por meio da modelagem de uma função objetivo e da otimização deste objetivo segundo as preferências de cada agente. A otimização pode ser condicionada (restrita - quando existe algum tipo de restrição sobre a função objetivo) ou não condicionada (irrestrita - quando a otimização é feita considerando apenas a função objetivo, sem quaisquer tipos de restrições sobre o problema).

2 Definições

Considere $f(x_i)$ com $i = 1, 2, ..., n$ como sendo uma função composta por múltiplos argumentos definida em um conjunto convexo S. Um problema de otimização pode ser escrito como:

$$max \quad f(x_i) \quad s.a \quad x \in S \tag{1}$$

No caso de um problema de maximização não condicionada, e

$$min \quad f(x_i) \quad s.a \quad x \in S \tag{2}$$

No caso de um problema de minimização não condicionada. Neste caso, $f(x_i)$ é a função objetivo do problema e busca-se encontrar os valores de cada argumento que resultam em um valor máximo(mínimo) para f. Se $i = 1$ e $f(x) = f(x_1)$, então diz-se que se trata de um problema de otimização simples ou otimização com um único argumento. Seja $x_i^\star$ o valor ótimo do argumento x_i, caso $x_i^\star$ esteja no interior de S, então diz-se que este problema tem uma solução interior.

Um vetor contendo valores únicos de cada argumento que provavelmente otimizem a função objetivo é conhecido como ponto crítico. Se $x_i^\star \in S$, então $x_i^\star$ será um ponto crítico global caso ocorra;

1. Se $f(x_i^\star) \geq f(x_i) \forall x_i \in S$, então $x_i^\star$ é um ponto crítico de máximo global.

2. Se $f(x_i^\star) > f(x_i) \forall x_i \in S$, então $x_i^\star$ é um ponto crítico de máximo global estrito.

3. Se $f(x_i^\star) \leq f(x_i) \forall x_i \in S$, então $x_i^\star$ é um ponto crítico de mínimo global.

4. Se $f(x_i^\star) < f(x_i) \forall x_i \in S$, então $x_i^\star$ é um ponto crítico de mínimo global estrito.

Seja $s \in S$ um subconjunto de S, Se $x_i^\star \in s$, então $x_i^\star$ será um ponto crítico local (ou relativo) caso ocorra:

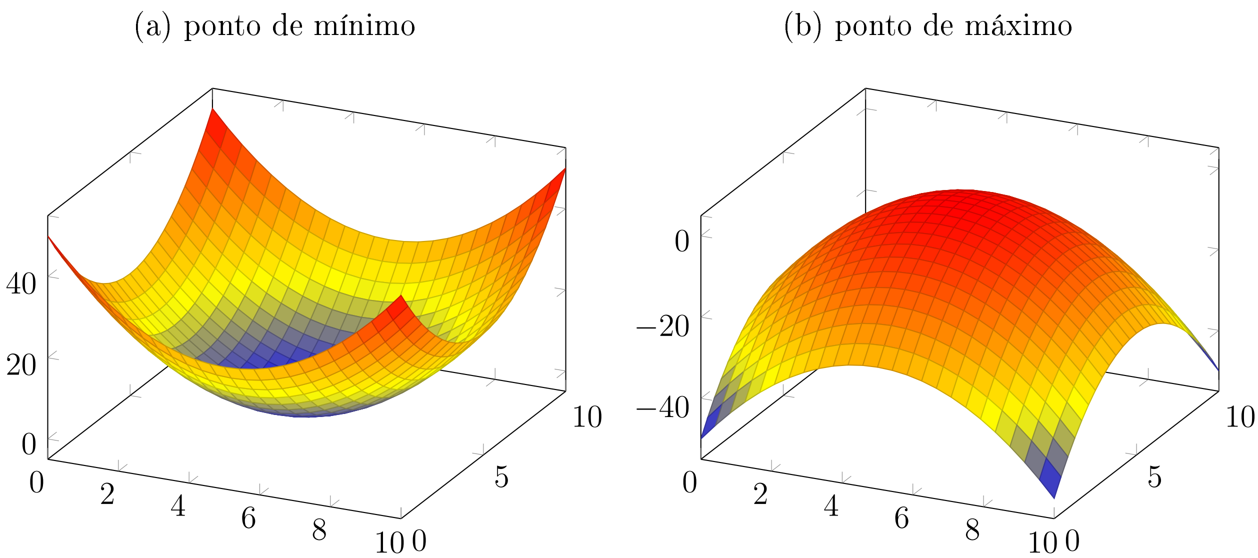
1. Se $f(x_i^\star) \geq f(x_i) \forall x_i \in s$, então $x_i^\star$ é um ponto crítico de máximo local.

2. Se $f(x_i^\star) > f(x_i) \forall x_i \in s$, então $x_i^\star$ é um ponto crítico de máximo local estrito.

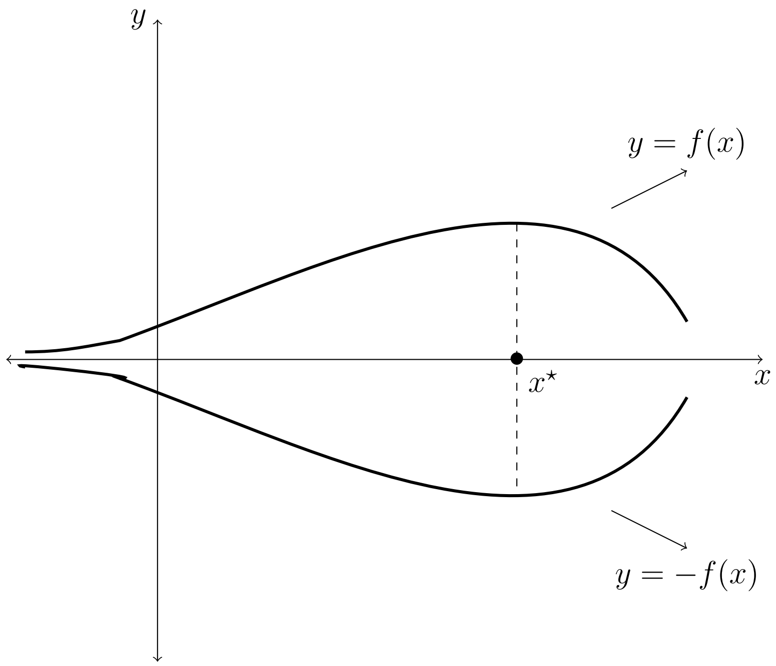
3. Se $f(x_i^\star) \leq f(x_i) \forall x_i \in s$, então $x_i^\star$ é um ponto crítico de mínimo local.

4. Se $f(x_i^\star) < f(x_i) \forall x_i \in s$, então $x_i^\star$ é um ponto crítico de mínimo local estrito.

Assim, $x_i^\star$ será um ponto crítico relativo (ou extremo relativo) de $f(x_i)$ se este valor representa um extremo na vizinhança de $x_i = x_i^\star$. A Figura abaixo traz uma demonstração simplificada destas definições. No painel (a) a função possui um ponto de mínimo local, enquanto no painel (b) a função apresenta um ponto de mínimo local.



Por definição, $x_i^\star$ maximiza $f(x_i)$ em S se e somente se $x_i^\star$ minimiza $-f(x_i)$ em S, isto é:



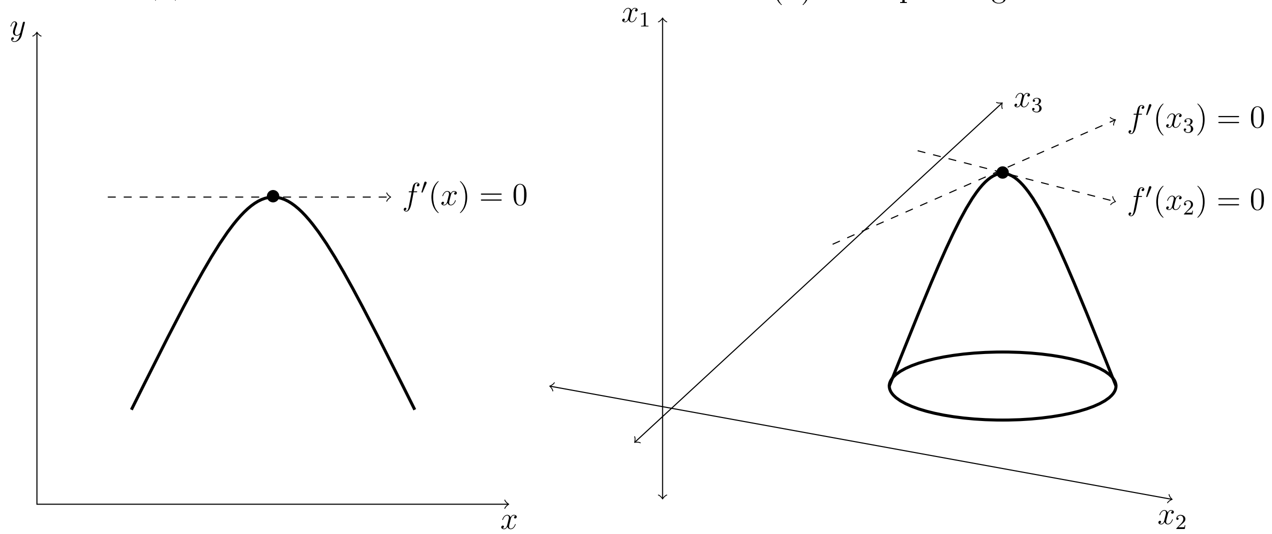
3 Otimização não condicionada

Se $f : S \to \mathbb{R}^n$ é contínua, então haverá um máximo e um mínimo global de f em $\mathbb{R}$. A condição necessária para a existência destes valores extremos estabelece que todas as derivadas parciais de primeira ordem de f em relação aos seus argumentos devem ser nulas, isto é:

$$\frac{\partial f(x_i)}{\partial x_1} = \frac{\partial f(x_i)}{\partial x_2} = ... \frac{\partial f(x_i)}{\partial x_n} = 0 \tag{3}$$

A lógica desta condição é demonstrar o ponto em que a inclinação da reta tangente à representação gráfica da função é nula, indicando um ponto de máximo ou mínimo da função conforme demonstrado na figura a seguir:

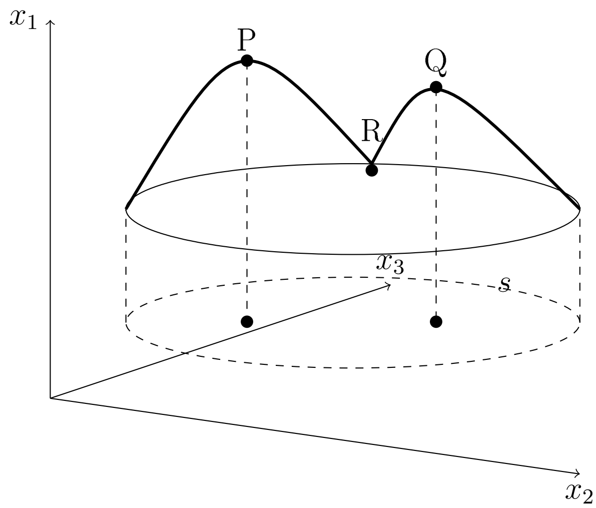
(a) Um argumento (b) Múltiplos argumentos



Se $x_i^\star$ é um ponto crítico de $f(x_i)$, então:

- $x_i^\star$ será um ponto de máximo local de $f(x_i)$ se o sinal de $f'(x_i^\star)$ muda de positivo para negativo da direita para a esquerda do ponto $x_i = x_i^\star$.
- $x_i^\star$ será um ponto de mínimo local de $f(x_i)$ se o sinal de $f'(x_i^\star)$ muda de negativo para positivo da direita para a esquerda do ponto $x_i = x_i^\star$.
- $x_i^\star$ não será máximo nem mínimo caso o sinal de $f'(x_i^\star)$ seja o mesmo à direita e à esquerda do ponto $x_i = x_i^\star$.

Se $f(x_i^\star)$ não é máximo nem mínimo, então diz-se que $x_i^\star$ é um ponto de cela. Um ponto de cela ocorre quando em um dado subconjunto $s \in S$ existe existe nas proximidades do vetor $x_i^\star$ um vetor x_i' tal que $f(x_i') < f(x_i^\star)$ e um outro vetor x_i'' tal que $f(x_i''$ tal que $f(x_i'') > f(x_i^\star) < f(x_i^\star)$. Esta condição é representada pelo ponto R na figura a seguir.



Se $f'(x_i)$ é contínua e derivável e seja $x_i^\star$ um ponto crítico que satisfaz as condições necessárias de $f'(x_i^\star) = 0$, então a condição suficiente para a otimização estabelece que:

1. $x_i^\star$ é um máximo local se $f''(x_i) < 0$.
2. $x_i^\star$ é um mínimo local se $f''(x_i) > 0$.
3. $x_i^\star$ é um ponto de inflexão se $f''(x_i) = 0$.

Quando $i > 1$, esta condição pode ser verificada por meio do cálculo do determinante da matriz de derivadas parciais de segunda ordem da função objetivo. Para tanto, seja:

$$\begin{aligned} &f''(x_{1,1}) = \frac{\partial(\partial f(x_i)/\partial x_1)}{\partial x_1} \quad ... \quad f''(x_{1,n}) = \frac{\partial(\partial f(x_i)/\partial x_1)}{\partial x_n} \\ &\vdots \\ &f''(x_{n,1}) = \frac{\partial(\partial f(x_i)/\partial x_n)}{\partial x_1} \quad ... \quad f''(x_{n,n}) = \frac{\partial(\partial f(x_i)/\partial x_n)}{\partial x_n} \end{aligned} \tag{4}$$

Então o determinante da matriz de derivadas parciais de segunda ordem é:

$$D_k\left(x_i{}^\star\right) = \begin{vmatrix} f''_{11}\left(x_i{}^\star\right) & f''_{12}\left(x_i{}^\star\right) & \cdots & f''_{1k}\left(x_i{}^\star\right) \\ f''_{21}\left(x_i{}^\star\right) & f''_{22}\left(x_i{}^\star\right) & \cdots & f''_{2k}\left(x_i{}^\star\right) \\ \vdots & \vdots & \ddots & \vdots \\ f''_{k1}\left(x_i{}^\star\right) & f''_{k2}\left(x_i{}^\star\right) & \cdots & f''_{kk}\left(x_i{}^\star\right) \end{vmatrix} \tag{5}$$

Neste caso, pode ocorrer:

- Se $(-1)^k D_k(x_i^\star) > 0 \; \forall \; k = 1, 2, ..., n$, então $x_i^\star$ é um máximo local de $f(x_i)$ em S.
- Se $D_k(x_i^\star) > 0 \; \forall \; k = 1, 2, ..., n$, então $x_i^\star$ é um mínimo local de $f(x_i)$ em S.
- Se $D_n(x_i^\star) \neq 0$, e se nenhuma das condições anteriores é satisfeita, então $x_i^\star$ é um ponto de sela $f(x_i)$ em S.

Uma aplicação rotineira da otimização não condicionada na teoria econômica está nas definições das funções custo médio e custo marginal. Por definição e considerando a ótica do insumo, o custo médio do insumo x ($cme(x)$) é a razão entre o custo total ($ct(x)$) e montante total utilizado do insumo x . Já o custo marginal é a derivada do custo total em relação ao insumo x. Portanto, tem-se:

$$
\begin{aligned}
cme(x) &= \frac{ct(x)}{x} \\
\frac{\partial cme(x)}{\partial x} &= \frac{x(\partial ct(x)/\partial x) - ct(x)}{x^2} \\
\frac{\partial cme(x)}{\partial x} &= \frac{\partial ct(x)}{x} - \frac{ct(x)}{x^2} \\
\frac{\partial cme(x)}{\partial x} &= \left(\frac{1}{x}\right)[cmg(x) - cme(x)]
\end{aligned} \tag{6}
$$

De onde é possível deduzir que:

- A condição necessária para a otimização estabelece que o custo médio será mínimo caso $\partial cme(x)/\partial x = 0$, o que só é possível se $cmg(x) = cme(x)$.
- A condição suficiente estabelece que o custo médio só será mínimo caso todas as derivadas parciais de segunda ordem em relação aos argumentos da função objetivo sejam positivas, o que só é possível caso ocorra:

$$
\begin{aligned}
&\frac{\partial^2 cme(x)}{\partial x^2} = \left(\frac{-1}{x^2}\right)[cmg(x) - cme(x)] + \left(\frac{1}{x}\right)\left[\frac{\partial cmg(x)}{\partial x} - \frac{\partial cme(x)}{\partial x}\right] > 0 \\
&\left(\frac{1}{x^2}\right)[cmg(x) - cme(x)] < \left(\frac{1}{x}\right)\left[\frac{\partial cmg(x)}{\partial x} - \frac{\partial cme(x)}{\partial x}\right] \\
&\left(\frac{1}{x}\right)\underbrace{[cmg(x) - cme(x)]}_{=0} < \left[\frac{\partial cmg(x)}{\partial x} - \frac{\partial cme(x)}{\partial x}\right] \\
&\frac{\partial cmg(x)}{\partial x} - \frac{\partial cme(x)}{\partial x} > 0 \\
&\frac{\partial cme(x)}{\partial x} < \frac{\partial cmg(x)}{\partial x}
\end{aligned} \tag{7}
$$

Ou seja, o custo médio só será mínimo caso a curva de custo médio seja menos inclinada do que a curva de custo marginal no ponto em que o custo médio é igual ao custo marginal. Como neste ponto a inclinação da curva de custo médio é nula, então no ponto em que o custo médio é igual ao custo marginal necessariamente a curva de custo marginal deve ser positivamente inclinada para que o custo médio seja mínimo.

4 Otimização condicionada

Este tipo de otimização é mais comum nos problemas econômicos e ocorrem quando a função objetivo está condicionada a uma restrição que limita o seu domínio. Seja $f(x_i)$ uma função objetivo definida em um conjunto convexo S com $i = 1, ..., n$ argumentos, sejam $g_j(x_i)$ as $j = 1, ..., m$ restrições do problema, então um problema de otimização condicionada pode ser escrito como:

$$
max \quad f(x_i) \qquad s.a \quad \begin{cases} g_1(x_1) & = b_1 \\ \vdots & \vdots \\ g_n(x_n) & = b_n \end{cases} \tag{8}
$$

No caso de um problema de maximização para todo $m < n$ e

$$min \quad f(x_i) \qquad s.a \qquad \begin{cases} g_1(x_1) & = b_1 \\ \vdots & \vdots \\ g_n(x_n) & = b_n \end{cases} \tag{9}$$

No caso de um problema de minimização para todo $m < n$. Se $f(x_i)$ é contínua em todo o seu domínio, então existe um vetor $\lambda^\star = (\lambda_1^\star, ..., \lambda_n^\star)$ que é solução da seguinte função de Lagrange:

$$L(x_i, \lambda_j) = f(x_i) + \sum_{j=1}^{m} \lambda_j \left[b_j - g_j(x_i)\right] \tag{10}$$

De tal modo que as seguintes condições necessárias devem ser satisfeitas:

$$\frac{\partial L(x_i^\star, \lambda_j^\star)}{\partial x_i} = \frac{\partial L(x_i^\star)}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x_i^\star)}{\partial x_i} \quad \forall \ i = 1, ..., n \tag{11}$$

As condições suficientes para a otimização desta classe de problema advém da condição de concavidade/convexidade da função de lagrange, se modo que, se a função de lagrange avaliada no ponto crítico é côncava(convexa), então o ponto crítico estabelecido nas condições necessárias para a otimização é um máximo(mínimo). Para avaliar estas condições é preciso analisar o padrão dos determinantes dos menores principais líderes da matriz heciana horlada de derivadas parciais de segunda ordem do problema, dado por:

$$H_r = \begin{vmatrix} 0 & \dots & 0 & \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_r} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_r} \\ \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_1} & \mathcal{L}''_{11} & \dots & \mathcal{L}''_{1r} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_r} & \dots & \frac{\partial g_m}{\partial x_r} & \mathcal{L}''_{r1} & \dots & \mathcal{L}''_{rr} \end{vmatrix} \tag{12}$$

Com $r = 1, ..., n$, de modo que os seguintes casos podem ocorrer:

- Se $(-1)^m H_r(x_i^\star) > 0 \; \forall \; r = m+1, ..., n$, então $x_i^\star$ é um ponto de mínimo local do problema.
- Se $(-1)^r H_r(x_i^\star) > 0 \; \forall \; r = m+1, ..., n$, então $x_i^\star$ é um ponto de máximo local do problema.
- Se nenhuma destas condições é satisfeita, então $x_i^\star$ é um ponto de sela do problema.

A otimização condicionada é rotineiramente utilizada na microeconomia, por exemplo, na teoria do consumidor, onde o consumidor representativo busca maximizar a utilidade do consumo sujeito ao seu poder aquisitivo. Para representar, suponha uma economia simplificada com apenas dois bens, x_1 e x_2, em que p_1 e p_2 são os preços de x_1 e x_2, respectivamente. Deixe $u(x_1, x_2)$ representar a utilidade do consumo e R o poder aquisitivo do consumidor. Considere que a utilidade pode ser fielmente representada por uma função do tipo Cobb-Douglas com α representando a participação do consumo de x_1 no orçamento total do consumidor. O problema é:

$$max \quad u(x_1, x_2) = x_1^\alpha, x_2^{1-\alpha} \qquad s.a \qquad R = p_1 x_1 + p_2 x - 2 \tag{13}$$

A função de Lagrange é:

$$\mathcal{L} = x_1^\alpha x_2^{1-\alpha} + \lambda[R - p_1 x_1 - p_2 x - 2] \tag{14}$$

As condições necessárias para a maximização são:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_1} &= \alpha x_1^{\alpha-1} x_2^{1-\alpha} - \lambda p_1 = 0 \\
\frac{\partial \mathcal{L}}{\partial x_2} &= (1-\alpha) x_1^{\alpha} x_2^{-\alpha} - \lambda p_2 = 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= R - p_1 x_1 - p_2 x_2 = 0
\end{aligned} \tag{15}$$

A resolução deste sistema resulta no seguinte ponto crítico:

$$(x_1^\star, x_2^\star) = \left(\frac{\alpha R}{p_1}, \frac{(1-\alpha)R}{p_2} \right) \tag{16}$$

As derivadas parciais de segunda ordem são:

$$\begin{aligned}
L_{1,1} &= \alpha \underbrace{(\alpha-1)}_{<0} x_1^{\alpha-2} x_2^{1-\alpha} \Rightarrow \quad < 0 \\
L_{1,2} &= \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \Rightarrow \quad > 0 \\
L_{2,1} &= \alpha(1-\alpha) x_1^{\alpha-1} x_2^{-\alpha} \Rightarrow \quad > 0 \\
L_{2,2} &= -\alpha(1-\alpha) x_1^{\alpha} x_2^{-\alpha-1} \Rightarrow \quad < 0
\end{aligned} \tag{17}$$

As derivadas parciais de primeira ordem das restrições são:

$$\begin{aligned}
g_{1,1} &= \frac{\partial g_1}{x_1} = p_1 \Rightarrow \quad > 0 \\
g_{1,2} &= \frac{\partial g_1}{x_2} = p_2 \Rightarrow \quad > 0
\end{aligned} \tag{18}$$

A hessiana orlada do problema é:

$$H = \begin{bmatrix} 0 & g_{1,1} & g_{1,2} \\ g_{1,1} & L_{1,1} & L_{1,2} \\ g_{1,2} & L_{2,1} & L_{2,2} \end{bmatrix} = \begin{bmatrix} 0 & g_{1,1} > 0 & g_{1,2} > 0 \\ g_{1,1} > 0 & L_{1,1} < 0 & L_{1,2} > 0 \\ g_{1,2} > 0 & L_{2,1} > 0 & L_{2,2} < 0 \end{bmatrix} \tag{19}$$

Para verificar se o ponto crítico é um ponto de mínimo é preciso testar se $(-1)^m H_r(x_i^\star) > 0 \; \forall \; r = m+1, ..., n$. Como $m = 1$, então o valor mínimo de r para o teste é $r = m + 1 = 2$, de modo que $|H_2(x_i^\star)|$ engloba o determinante da matriz hessiana orlada completa. O determinante da matriz hessiana horlada para $r = 2$ é:

$$|H_2(x_i^\star)| = \begin{vmatrix} 0 & > & > \\ > & < & > \\ > & > & < \end{vmatrix}$$

$$\begin{aligned}
&|H_2(x_i^\star)| => * > * > +0* < * < + > * > * > -(0* > * > + > * < * > + > * > * <) \\
&|H_2(x_i^\star)| => +0+ > -(0+ < + <) \\
&|H_2(x_i^\star)| => - < \\
&|H_2(x_i^\star)| ==> + > \\
&|H_2(x_i^\star)| > 0
\end{aligned} \tag{20}$$

Portanto:

$$(-1)^r H_r(x_i^\star) = (-1)^2 |H_2(x_i^\star)| = 1 * (>) \quad \Rightarrow \quad > 0 \tag{21}$$

O que indica que a solução $(x_1^\star, x_2^\star)$ é um ponto de máximo local da função de utilidade do consumidor.