Gabarito Programação não linear- introdução à otimização

Kelvin (Pol)

Exercício 1.

Sejam $a, b \in \mathbb{R}$ dois números reais positivos. Considere a função de Rosenbrock $f(x) = a(x_2 - x_1^2)^2 + b(1-x_1)^2$. Encontre o (único) ponto estacionário de f e verifique se é minimizador local. Prove que $\nabla^2 f(x)$ é singular se e somente se $x_2 - x_1^2 = \frac{b}{2a}$.

Solução:

Note-se que f é um polinômio,em particular, ele é de classe C^2. Assim,

$$\nabla f(x) = \begin{bmatrix} -4ax_1(x_2 - x_1^2) - 2b(1-x_1) \\ 2a(x_2 - x_1^2) \end{bmatrix},$$

logo

$$\nabla f(x) = 0 \Longrightarrow \begin{bmatrix} -4ax_1(x_2 - x_1^2) - 2b(1-x_1) \\ 2a(x_2 - x_1^2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \underbrace{\Longrightarrow}_{x_1 = x_2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Como o sistema $\nabla f(x) = 0$ tem uma única solução em $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{x}^*$, então f tem um único ponto estacionário.

Agora estudemos la natureza do Hessiano de f, de fato

$$\nabla^2 f(x) = \begin{bmatrix} 12ax_1^2 - 4ax_2 + 2b & -4ax_1 \\ -4ax_1 & 2a \end{bmatrix},$$

avaliando no ponto estacionário $\mathbf{x}^*$, temos

$$\nabla^2 f(\mathbf{x}^*) = \begin{bmatrix} 8a + 2b & -4a \\ -4a & 2a \end{bmatrix}.$$

Note que, $\det \nabla^2 f(\mathbf{x}^*) = 16a^2 + 2ab - 16a^2 = 4ab$, como $a > 0$ e $b > 0$ então $\det \nabla^2 f(\mathbf{x}^*) > 0$, além disso, a determinate da sub-matriz $\begin{bmatrix} 8a + 2b \end{bmatrix}$ é positiva. Assim, pelo critério de Sylvester, temos que a matriz $\nabla^2 f(\mathbf{x}^*)$ é positiva definida o que implica que o ponto $\mathbf{x}^*$ é um minimizador local de f.

Por outro lado, a matriz $\nabla^2 f(x)$ é singular se e somente se

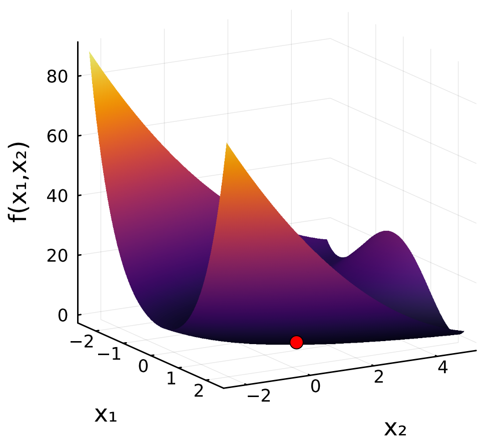
$$\det \nabla^2 f(x) = 0.$$

Calculando explicitamente:

$$\begin{aligned} \det \nabla^2 f(x) &= (12ax_1^2 - 4ax_2 + 2b)(2a) - 16a^2x_1^2 \\ &= 24a^2x_1^2 - 8a^2x_2 + 4ab - 16a^2x_1^2 \\ &= 8a^2x_1^2 - 8a^2x_2 + 4ab. \end{aligned}$$

Logo,

$$\begin{aligned} \det \nabla^2 f(x) = 0 &\Longleftrightarrow 8a^2x_1^2 - 8a^2x_2 + 4ab = 0 \\ &\Longleftrightarrow 8a^2(x_1^2 - x_2) = -4ab \\ &\Longleftrightarrow x_2 - x_1^2 = \frac{4ab}{8a^2} = \frac{b}{2a}. ☺ \end{aligned}$$



Observação

Observação (Critério para Determinar a Natureza do Hessiano):
Seja $A \in \mathbb{R}^{n\times n}$ uma matriz simétrica (i.e., $A = A^T$). Consideremos o conjunto de autovalores de A dado por

$$\lambda(A) := \{\lambda \in \mathbb{R} : \lambda \text{ é autovalor de } A\}.$$

Nota: Os autovalores de uma matriz simétrica são sempre reais.

1. Critério dos Autovalores:

- Se $\lambda > 0\ \forall \lambda \in \lambda(A)$, então A é *positiva definida* ($A \succ 0$);
- Se $\lambda \geq 0\ \forall \lambda \in \lambda(A)$, então A é *positiva semi-definida* ($A \succeq 0$);
- Se $\lambda < 0\ \forall \lambda \in \lambda(A)$, então A é *negativa definida* ($A \prec 0$);
- Se $\lambda \leq 0\ \forall \lambda \in \lambda(A)$, então A é *negativa semi-definida* ($A \preceq 0$);
- Se $\exists \lambda_i > 0$ e $\lambda_j < 0$ com $i \neq j$ para algum $i, j \in \{1, \ldots, n\}$, então A é *indefinida*.

2. Determinante e Traço:
O determinante e o traço de A estão relacionados aos autovalores da seguinte forma:

$$\det A = \prod_{i=1}^{n} \lambda_i \quad \text{e} \quad \operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i.$$

Consequência: Se $\det A = 0$, então pelo menos um autovalor é zero.

3. Critério de Sylvester:
Seja $A^{[k]}$ a submatriz principal de ordem k de A, isto é, a submatriz formada pelas primeiras k linhas e k colunas de A. Então,

$$A \succ 0 \iff \det A^{[k]} > 0 \quad \forall k \in \{1, \ldots, n\}.$$

4. Multiplicação por Escalar:
Se $B = cA$ onde $c \in \mathbb{R}$ e $A \in \mathbb{R}^{n\times n}$, então:

- Se $c > 0$, então B tem a mesma natureza que A;
- Se $c < 0$, então B tem a natureza oposta à de A;
- Se $c = 0$, então B é a matriz nula (positiva e negativa semi-definida).

5. Interpretação em Otimização:
Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função de classe C^2 e x^* um ponto estacionário de f (i.e., $\nabla f(x^*) = 0$). A natureza do Hessiano $\nabla^2 f(x^*)$ determina o tipo de ponto crítico:

- Se $\nabla^2 f(x^*) \succ 0$, então x^* é um *minimizador local estrito* de f;
- Se $\nabla^2 f(x^*) \prec 0$, então x^* é um *maximizador local estrito* de f;
- Se $\nabla^2 f(x^*)$ é indefinida, então x^* é um *ponto de sela* de f.

Exercício 2.

Seja $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x) = \sin x_1 \sin x_2 + e^{x_1^2+x_2^2}$. Mostre que $\overline{x} = 0$ é ponto estacionário de f. Diga se é minimizador, maximizador ou sela.

Solução:

Note-se que f é de classe C^2 (pois é soma de funções de classe C^2). Assim, calculemos o gradiente de f:

$$\nabla f(x) = \begin{bmatrix} \cos x_1 \sin x_2 + 2x_1 e^{x_1^2+x_2^2} \\ \sin x_1 \cos x_2 + 2x_2 e^{x_1^2+x_2^2} \end{bmatrix}.$$

Avaliando em $\overline{x} = 0$, temos

$$\nabla f(0) = \begin{bmatrix} \cos(0)\sin(0) + 2(0)e^0 \\ \sin(0)\cos(0) + 2(0)e^0 \end{bmatrix} = \begin{bmatrix} 1 \cdot 0 + 0 \\ 0 \cdot 1 + 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Logo, $\overline{x} = 0$ é ponto estacionário de f.
Agora estudemos a natureza do Hessiano de f. De fato, calculando as derivadas parciais segundas:

$$\frac{\partial^2 f}{\partial x_1^2} = -\sin x_1 \sin x_2 + 2e^{x_1^2+x_2^2} + 4x_1^2 e^{x_1^2+x_2^2},$$

$$\frac{\partial^2 f}{\partial x_2^2} = -\sin x_1 \sin x_2 + 2e^{x_1^2+x_2^2} + 4x_2^2 e^{x_1^2+x_2^2},$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(x) = \cos x_1 \cos x_2 + 4x_1 x_2 e^{x_1^2+x_2^2} \underbrace{=}_{\text{teo. Schwarz}} \frac{\partial^2 f}{\partial x_2 \partial x_2}(x).$$

Assim,

$$\nabla^2 f(x) = \begin{bmatrix} -\sin x_1 \sin x_2 + 2e^{x_1^2+x_2^2} + 4x_1^2 e^{x_1^2+x_2^2} & \cos x_1 \cos x_2 + 4x_1 x_2 e^{x_1^2+x_2^2} \\ \cos x_1 \cos x_2 + 4x_1 x_2 e^{x_1^2+x_2^2} & -\sin x_1 \sin x_2 + 2e^{x_1^2+x_2^2} + 4x_2^2 e^{x_1^2+x_2^2} \end{bmatrix}.$$

Avaliando no ponto estacionário $\overline{x} = 0$, temos

$$\nabla^2 f(0) = \begin{bmatrix} -0 \cdot 0 + 2e^0 + 0 & 1 \cdot 1 + 0 \\ 1 \cdot 1 + 0 & -0 \cdot 0 + 2e^0 + 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

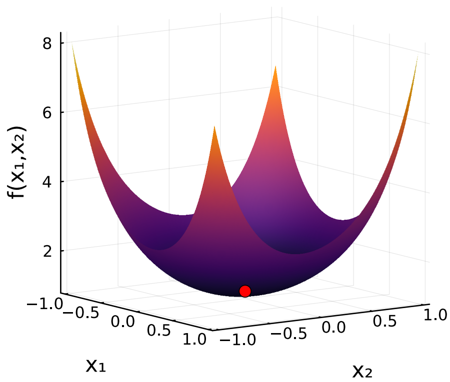
Note que $\det \nabla^2 f(0) = 4 - 1 = 3 > 0$, além disso, a determinante da sub-matriz $\begin{bmatrix} 2 \end{bmatrix}$ é positiva. Assim, pelo critério de Sylvester, temos que a matriz $\nabla^2 f(0)$ é positiva definida.

Alternativamente, podemos usar o critério dos autovalores sem calculá-los explicitamente. De fato, pela Observação anterior (item 2), sabemos que

$$\det \nabla^2 f(0) = \lambda_1 \cdot \lambda_2 = 3 > 0 \quad \text{e} \quad \operatorname{tr}(\nabla^2 f(0)) = \lambda_1 + \lambda_2 = 2 + 2 = 4 > 0.$$

Como o produto dos autovalores é positivo, temos duas possibilidades: ambos positivos ou ambos negativos. Porém, como a soma dos autovalores (traço) é estritamente positiva, necessariamente ambos autovalores devem ser estritamente positivos. Logo, pela Observação anterior (item 1), a matriz $\nabla^2 f(0)$ é positiva definida ($\nabla^2 f(0) \succ 0$).

Portanto, o ponto $\overline{x} = 0$ é um minimizador local estrito de f. ☺



Observação

Observação: Neste exercício, foi possível encontrar o ponto estacionário $\overline{x} = 0$ analiticamente, resolvendo $\nabla f(x) = 0$. No entanto, na maioria dos problemas práticos de otimização, a equação $\nabla f(x) = 0$ pode ser um sistema não linear extremamente complexo ou até mesmo impossível de resolver de trivial.

É precisamente por isso que se desenvolvem *algoritmos iterativos de otimização*, tais como:

- **Método do Gradiente**: $x^{k+1} = x^k - t_k \nabla f(x^k)$ (descida mais íngreme)
- **Método de Newton**: $x^{k+1} = x^k - [\nabla^2 f(x^k)]^{-1} \nabla f(x^k)$ (convergência quadrática)
- **Métodos Quasi-Newton (BFGS)**: Aproximam $\nabla^2 f$ iterativamente, evitando o custo computacional da inversão do Hessiano

Esses métodos geram uma sequência $\{x^k\}$ que converge (sob certas condições) para um ponto estacionário, sem necessidade de resolver $\nabla f(x) = 0$ explicitamente. Além disso, a escolha do *tamanho de passo* t_k (através de busca exata ou condições de Armijo) é fundamental para garantir convergência e eficiência.

Exercício 3.

Prove que a função $f(x) = (x_2 - x_1^2)^2 + x_1^5$ tem um único ponto estacionário que não é minimizador nem maximizador local nem ponto de sela.

Solução:

É claro que f é de classe C^2 (pois é um polinômio). Calculemos o gradiente de f:

$$\nabla f(x) = \begin{bmatrix} -4x_1(x_2 - x_1^2) + 5x_1^4 \\ 2(x_2 - x_1^2) \end{bmatrix}.$$

Logo,

$$\nabla f(x) = 0 \iff \begin{cases} -4x_1(x_2 - x_1^2) + 5x_1^4 = 0 \\ 2(x_2 - x_1^2) = 0 \end{cases} \iff \begin{cases} -4x_1 \cdot 0 + 5x_1^4 = 0 \\ x_2 = x_1^2 \end{cases}$$

$$\iff \begin{cases} 5x_1^4 = 0 \\ x_2 = x_1^2 \end{cases} \iff \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}.$$

Como o sistema $\nabla f(x) = 0$ tem uma única solução em $\mathbf{x}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, então f tem um único ponto estacionário.

Assim,

$$\nabla^2 f(x) = \begin{bmatrix} -4(x_2 - x_1^2) + 8x_1^2 + 20x_1^3 & -4x_1 \\ -4x_1 & 2 \end{bmatrix}.$$

Avaliando no ponto estacionário $\mathbf{x}^* = 0$, temos

$$\nabla^2 f(0) = \begin{bmatrix} -4(0-0) + 0 + 0 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.$$

Note que $\det \nabla^2 f(0) = 0 \cdot 2 - 0 \cdot 0 = 0$. Pela Observação anterior (item 2), como o determinante é zero, pelo menos um autovalor é nulo. De fato, $\mathrm{tr}(\nabla^2 f(0)) = 0 + 2 = 2 > 0$, logo $\lambda_1 + \lambda_2 = 2$ e $\lambda_1 \cdot \lambda_2 = 0$, o que implica que um autovalor é zero e o outro é $\lambda = 2 > 0$.

Como $\nabla^2 f(0)$ possui um autovalor nulo, a matriz é singular e semi-definida, mas não definida. Portanto, o teste da segunda derivada é *inconclusivo* neste caso. Pela Observação anterior (item 5), não podemos classificar $\mathbf{x}^*$ como minimizador, maximizador ou ponto de sela usando apenas a informação do Hessiano.

Para determinar a natureza do ponto, analisemos o comportamento de f em torno de $\mathbf{x}^* = 0$. Consideremos pontos da forma (t, t^2) para t próximo de zero:

$$f(t, t^2) = (t^2 - t^2)^2 + t^5 = 0 + t^5 = t^5.$$

Note que:

- Se $t > 0$, então $f(t, t^2) = t^5 > 0 = f(0, 0)$;
- Se $t < 0$, então $f(t, t^2) = t^5 < 0 = f(0, 0)$.

Logo, em qualquer vizinhança de $\mathbf{x}^* = 0$, existem pontos onde f é maior que $f(0)$ e pontos onde f é menor que $f(0)$. Portanto, $\mathbf{x}^* = 0$ não é minimizador local (pois existem pontos próximos com valores menores), não é maximizador local (pois existem pontos próximos com valores maiores), e não é ponto de sela no sentido clássico (pois o Hessiano é singular). ☺

Exercício 4.

Seja $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x) = x_1^2 + x_2^2 - x_1 x_2^2$. Determine e faça um esboço do conjunto $\{x \in \mathbb{R}^2 \mid \nabla^2 f(x) \succ 0\}$.

Solução:

É claro que f é de classe C^2 (pois é um polinômio). Calculemos o Hessiano de f. De fato, temos:

$$\nabla f(x) = \begin{bmatrix} 2x_1 - x_2^2 \\ 2x_2 - 2x_1 x_2 \end{bmatrix}.$$

Calculando as derivadas parciais segundas:

$$\frac{\partial^2 f}{\partial x_1^2} = 2, \quad \frac{\partial^2 f}{\partial x_2^2} = 2 - 2x_1, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = -2x_2.$$

Assim,

$$\nabla^2 f(x) = \begin{bmatrix} 2 & -2x_2 \\ -2x_2 & 2 - 2x_1 \end{bmatrix}.$$

Para determinar o conjunto onde $\nabla^2 f(x) \succ 0$, usaremos o critério de Sylvester. Então, $\det \nabla^2 f(x)^{[1]} = 2 > 0$ (sempre satisfeito), $\det \nabla^2 f(x) > 0$.

Calculemos o determinante:

$$\det \nabla^2 f(x) = 2(2 - 2x_1) - (-2x_2)^2 = 4 - 4x_1 - 4x_2^2 = 4(1 - x_1 - x_2^2).$$

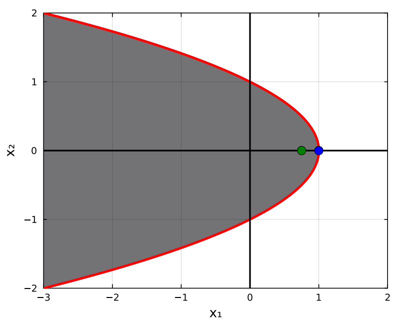
Logo,

$$\nabla^2 f(x) \succ 0 \iff \det \nabla^2 f(x) > 0 \iff 4(1 - x_1 - x_2^2) > 0 \iff 1 - x_1 - x_2^2 > 0$$
$$\iff x_1 < 1 - x_2^2.$$

Portanto, o conjunto onde o Hessiano é positivo definido é

$$\{x \in \mathbb{R}^2 \mid \nabla^2 f(x) \succ 0\} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 < 1 - x_2^2\}.$$

Geometricamente, este conjunto corresponde à região à *esquerda* da parábola $x_1 = 1 - x_2^2$ no plano (x_1, x_2). A fronteira é dada pela parábola com vértice em $(1, 0)$ e concavidade voltada para a esquerda (abrindo para $x_1 < 0$). ☺



Exercício 5.

Seja $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x) = \frac{1}{2}(x_1^2 - x_2)^2 + \frac{1}{2}(1 - x_1)^2$.

a) Mostre que f tem um único ponto estacionário e obtenha tal ponto.

b) Utilize as condições de segunda ordem para classificar este ponto.

c) Mostre que f é coerciva.

Solução:

É claro que f é de classe C^2 (pois é composição de funções de classe C^2).

(a) Calculemos o gradiente de f. De fato,

$$\nabla f(x) = \begin{bmatrix} 2x_1(x_1^2 - x_2) - (1 - x_1) \\ -(x_1^2 - x_2) \end{bmatrix} = \begin{bmatrix} 2x_1^3 - 2x_1x_2 - 1 + x_1 \\ -x_1^2 + x_2 \end{bmatrix}.$$

Logo,

$$\nabla f(x) = 0 \iff \begin{cases} 2x_1^3 - 2x_1x_2 + x_1 - 1 = 0 \\ x_2 = x_1^2 \end{cases}.$$

Substituindo $x_2 = x_1^2$ na primeira equação:

$$2x_1^3 - 2x_1(x_1^2) + x_1 - 1 = 0 \iff 2x_1^3 - 2x_1^3 + x_1 - 1 = 0 \iff x_1 - 1 = 0 \iff x_1 = 1.$$

Assim, $x_1 = 1$ e $x_2 = 1^2 = 1$. Logo, o único ponto estacionário é $\mathbf{x}^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(b) Agora estudemos a natureza do Hessiano de f.

$$\nabla^2 f(x) = \begin{bmatrix} 6x_1^2 - 2x_2 + 1 & -2x_1 \\ -2x_1 & 1 \end{bmatrix}.$$

Avaliando no ponto estacionário $\mathbf{x}^* = (1, 1)$, temos

$$\nabla^2 f(1, 1) = \begin{bmatrix} 6(1)^2 - 2(1) + 1 & -2(1) \\ -2(1) & 1 \end{bmatrix} = \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}.$$

Calculemos o determinante:

$$\det \nabla^2 f(1, 1) = 5 \cdot 1 - (-2)^2 = 5 - 4 = 1 > 0.$$

Além disso, a determinante da sub-matriz $\begin{bmatrix} 5 \end{bmatrix}$ é positiva. Assim, pelo critério de Sylvester, temos que a matriz $\nabla^2 f(1, 1)$ é positiva definida. Portanto, o ponto $\mathbf{x}^* = (1, 1)$ é um minimizador local estrito de f.

(c) Para mostrar que f é coerciva, devemos provar que $f(x) \to +\infty$ quando $\|x\| \to +\infty$.
Note que

$$f(x) = \frac{1}{2}(x_1^2 - x_2)^2 + \frac{1}{2}(1 - x_1)^2 \geq 0,$$

pois f é soma de quadrados. Além disso, observe que

$$(x_1^2 - x_2)^2 \geq 0 \quad \text{e} \quad (1 - x_1)^2 \geq 0.$$

Analisemos o comportamento de f quando $\|x\| \to +\infty$. Temos $\|x\|^2 = x_1^2 + x_2^2 \to +\infty$. Consideremos dois casos:
Caso 1: Se $|x_1| \to +\infty$, então $(1 - x_1)^2 \geq (|x_1| - 1)^2 \to +\infty$.
Caso 2: Se $|x_2| \to +\infty$ enquanto x_1 permanece limitado, então existe $M > 0$ tal que $|x_1| \leq M$. Logo, $x_1^2 \leq M^2$, e assim

$$(x_1^2 - x_2)^2 \geq (|x_2| - M^2)^2 \to +\infty \quad \text{quando } |x_2| \to +\infty.$$

Em ambos os casos, temos $f(x) \to +\infty$ quando $\|x\| \to +\infty$. Portanto, f é coerciva. ☺

Observação

Como f é coerciva (c) e de classe C^2 (em particular, contínua), e possui um único ponto estacionário $(1, 1)$ que é minimizador local estrito (item 2), segue que $(1, 1)$ é, de fato, um *minimizador global* de f. Isto é,

$$f(1, 1) \leq f(x) \quad \forall x \in \mathbb{R}^2.$$

Exercício 6.

Dados $a, b > 0$, considere a função $f : (0, \infty) \to \mathbb{R}$ dada por $f(x) = ax^2 + \frac{b}{x}$.

(a) Determine o único ponto estacionário, x^*, de f.

(b) Prove que f é decrescente em $(0, x^*]$ e crescente em $[x^*, \infty)$.

(c) Prove que x^* é minimizador global de f.

(d) Represente geometricamente o gráfico de f e seu minimizador.

Solução:

É claro que f é de classe C^2 em $(0, \infty)$.

(a) Calculemos o ponto estacionário de f. De fato,

$$f'(x) = 2ax - \frac{b}{x^2}.$$

Logo,

$$f'(x) = 0 \iff 2ax - \frac{b}{x^2} = 0 \iff 2ax = \frac{b}{x^2} \iff 2ax^3 = b \iff x^3 = \frac{b}{2a} \iff x = \sqrt[3]{\frac{b}{2a}}.$$

Como $a, b > 0$, temos que $x^* = \sqrt[3]{\frac{b}{2a}} > 0$ é o único ponto estacionário de f em $(0, \infty)$.

(b) Analisemos o sinal de $f'(x)$ para determinar os intervalos de crescimento e decrescimento. De fato, podemos reescrever

$$f'(x) = 2ax - \frac{b}{x^2} = \frac{2ax^3 - b}{x^2}.$$

Como $x^2 > 0$ para todo $x \in (0, \infty)$, o sinal de $f'(x)$ é determinado pelo sinal de $2ax^3 - b$. Note que, do item (a), temos $x^* = \sqrt[3]{\frac{b}{2a}}$, ou seja, $2a(x^*)^3 = b$.

Assim, para $x \in (0, x^*)$, temos $x < x^*$, logo $x^3 < (x^*)^3 = \frac{b}{2a}$, o que implica $2ax^3 < b$, isto é, $2ax^3 - b < 0$. Portanto, $f'(x) < 0$ para todo $x \in (0, x^*)$, o que implica que f é estritamente decrescente em $(0, x^*]$.

Analogamente, para $x \in (x^*, \infty)$, temos $x > x^*$, logo $x^3 > (x^*)^3 = \frac{b}{2a}$, o que implica $2ax^3 > b$, isto é, $2ax^3 - b > 0$. Portanto, $f'(x) > 0$ para todo $x \in (x^*, \infty)$, o que implica que f é estritamente crescente em $[x^*, \infty)$.

(c) Mostremos que f é coerciva, isto é, $f(x) \to +\infty$ quando $|x| \to +\infty$ no domínio $(0, \infty)$. De fato, analisemos o comportamento nos extremos:

- Quando $x \to 0^+$: temos $ax^2 \to 0$ e $\frac{b}{x} \to +\infty$, logo $f(x) \to +\infty$.
- Quando $x \to +\infty$: temos $ax^2 \to +\infty$ e $\frac{b}{x} \to 0$, logo $f(x) \to +\infty$.

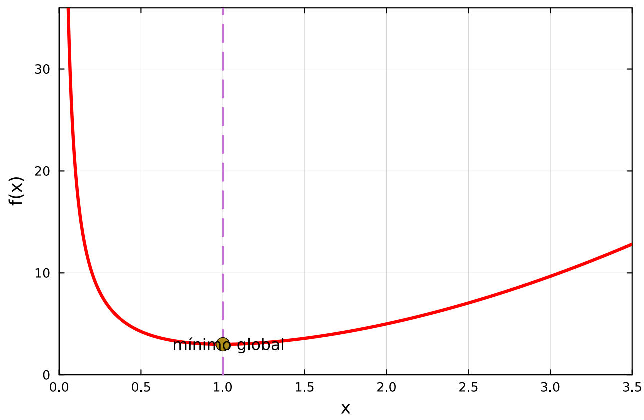
Portanto, f é coerciva em $(0, \infty)$. Como f é contínua, coerciva e possui um único ponto estacionário x^*, segue que x^* é minimizador global de f em $(0, \infty)$. ☺

(d) Para esboçar o gráfico, podemos fixar valores convenientes para os parâmetros a e b, pois eles apenas alteram a escala da função, sem modificar o seu comportamento qualitativo (decresce até x^* e cresce depois).

Como $x^* = \left(\frac{b}{2a}\right)^{1/3}$, escolhemos $b = 2a$ para obter $x^* = 1$. Por simplicidade, tomamos $a = 1$ e $b = 2$, e representamos o gráfico de

$$f(x) = x^2 + \frac{2}{x},$$

que atinge o mínimo global em $x = 1$.



Exercício 7.

Seja $f : \mathbb{R}^n \to \mathbb{R}$ dada por $f(x) = \frac{x^T x}{e^{x^T x}}$. Encontre todos os pontos estacionários de f e mostre que tais pontos são extremos globais.

Solução:

É claro que f é de classe C^2 em $\mathbb{R}^n$ (pois é composição e quociente de funções de classe C^2, com denominador sempre positivo).

Observação

Observação (Derivadas úteis): Seja $g(x) = x^T x = \|x\|^2$ o produto interno euclidiano. Então:

- $\nabla g(x) = 2x$,
- $\nabla^2 g(x) = 2I_n$ (onde I_n é a matriz identidade $n \times n$).

Pode mostrar o leitor ;)

Calculemos o gradiente de f usando a regra do quociente e a regra da cadeia. Seja $u(x) = x^T x$ e $v(x) = e^{x^T x} = e^{u(x)}$. Então $f(x) = \frac{u(x)}{v(x)}$.

Pela regra do quociente:

$$\nabla f(x) = \frac{v(x)\nabla u(x) - u(x)\nabla v(x)}{[v(x)]^2}.$$

Calculemos cada termo:

- $\nabla u(x) = 2x$ (pela Observação),
- $\nabla v(x) = \nabla(e^{u(x)}) = e^{u(x)}\nabla u(x) = e^{x^T x} \cdot 2x$ (regra da cadeia).

Substituindo:

$$\nabla f(x) = \frac{e^{x^T x} \cdot 2x - x^T x \cdot e^{x^T x} \cdot 2x}{[e^{x^T x}]^2} = \frac{2xe^{x^T x}(1 - x^T x)}{e^{2x^T x}} = \frac{2x(1 - x^T x)}{e^{x^T x}}.$$

Logo,

$$\nabla f(x) = 0 \iff \frac{2x(1 - x^T x)}{e^{x^T x}} = 0 \iff 2x(1 - x^T x) = 0.$$

Como $e^{x^T x} > 0$ sempre, temos duas condições:

a) $x = 0$, ou

b) $1 - x^T x = 0 \iff x^T x = 1 \iff \|x\|^2 = 1$.

Portanto, os pontos estacionários são:

- $x = 0$ (a origem), e
- $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ (a esfera unitária em $\mathbb{R}^n$).

Agora classifiquemos estes pontos usando a segunda derivada. Para calcular o Hessiano de f, usaremos a seguinte identidade. Seja $h(x) = \frac{u(x)}{v(x)}$ onde u, v são funções diferenciáveis. Então, após alguns cálculos (aplicando regra do quociente e do produto), o Hessiano pode ser expresso como:

$$\nabla^2 f(x) = \frac{1}{e^{x^T x}} \left[2I_n(1 - x^T x) - 4xx^T(1 - x^T x) - 4xx^T\right].$$

Simplificando:

$$\nabla^2 f(x) = \frac{2}{e^{x^T x}} \left[I_n(1 - x^T x) - 2xx^T(1 - x^T x) - 2xx^T\right].$$

No ponto $x = 0$:

$$\nabla^2 f(0) = \frac{2}{e^0} \left[I_n(1 - 0) - 0 - 0\right] = 2I_n.$$

Como $2I_n$ é positiva definida (todos os autovalores são iguais a $2 > 0$), implica que $x = 0$ é um minimizador local estrito de f.

Além disso, como $f(x) \geq 0$ para todo $x \in \mathbb{R}^n$ (pois $x^T x \geq 0$ e $e^{x^T x} > 0$) e $f(0) = 0$, segue que $x = 0$ é um **minimizador global** de f.

Nos pontos $x \in S^{n-1}$ (onde $x^T x = 1$):

$$\nabla^2 f(x) = \frac{2}{e} \left[I_n(1 - 1) - 2xx^T(1 - 1) - 2xx^T\right] = \frac{2}{e} \left[-2xx^T\right] = -\frac{4}{e}xx^T.$$

Para analisar a natureza desta matriz, consideremos a forma quadrática. Para qualquer $v \in \mathbb{R}^n$, temos:

$$v^T \nabla^2 f(x) v = v^T \left(-\frac{4}{e}xx^T\right) v = -\frac{4}{e}(x^T v)^2 \leq 0.$$

Portanto, $\nabla^2 f(x)$ é negativa semi-definida, o que sugere que os pontos em S^{n-1} são candidatos a maximizadores locais. No entanto, o teste da segunda derivada é inconclusivo neste caso.

Para confirmar, avaliemos f nos pontos críticos:

- $f(0) = 0$,
- $f(x) = \frac{1}{e} > 0$ para todo $x \in S^{n-1}$.

Além disso, quando $\|x\| \to +\infty$, temos $f(x) = \frac{x^T x}{e^{x^T x}} \to 0$ (pois o denominador cresce exponencialmente). Como f é contínua, $f(x) \geq 0$ para todo $x \in \mathbb{R}^n$, e o valor máximo $\frac{1}{e}$ é atingido em S^{n-1}, concluímos que todo ponto $x \in S^{n-1}$ é um **maximizador global** de f. ☹

Exercício 8.

Considere a função $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x) = ax_1x_2 + x_1^3 + x_2^3 + e^{x_1^2 + x_2^2}$, onde $a > 2$ é uma constante. Mostre que f tem pelo menos dois pontos estacionários.

Solução:

É claro que f é de classe C^2 em $\mathbb{R}^2$. Calculemos o gradiente de f. De fato,

$$\nabla f(x) = \begin{bmatrix} ax_2 + 3x_1^2 + 2x_1 e^{x_1^2 + x_2^2} \\ ax_1 + 3x_2^2 + 2x_2 e^{x_1^2 + x_2^2} \end{bmatrix}.$$

Logo,

$$\nabla f(x) = 0 \iff \begin{cases} ax_2 + 3x_1^2 + 2x_1 e^{x_1^2+x_2^2} = 0 \\ ax_1 + 3x_2^2 + 2x_2 e^{x_1^2+x_2^2} = 0 \end{cases}.$$

Note que $x^* = 0$ é uma solução trivial do sistema, pois $\nabla f(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Portanto, $x^* = 0$ é um ponto estacionário de f.

Agora analisemos a natureza deste ponto. Calculando o Hessiano de f, temos

$$\nabla^2 f(x) = \begin{bmatrix} 6x_1 + 2e^{x_1^2+x_2^2} + 4x_1^2 e^{x_1^2+x_2^2} & a + 4x_1x_2 e^{x_1^2+x_2^2} \\ a + 4x_1x_2 e^{x_1^2+x_2^2} & 6x_2 + 2e^{x_1^2+x_2^2} + 4x_2^2 e^{x_1^2+x_2^2} \end{bmatrix}.$$

Avaliando no ponto $x^* = 0$, obtemos

$$\nabla^2 f(0) = \begin{bmatrix} 2 & a \\ a & 2 \end{bmatrix}.$$

Analisemos a natureza desta matriz. Pela primeira Observação, temos

$$\det \nabla^2 f(0) = 4 - a^2 \quad \text{e} \quad \operatorname{tr}(\nabla^2 f(0)) = 4.$$

Por hipótese temos que

$$a > 2 \implies a^2 > 4 \implies 4 - a^2 < 0 \implies \det \nabla^2 f(0) < 0$$

. Além disso, $\operatorname{tr}(\nabla^2 f(0)) = \lambda_1 + \lambda_2 = 4 > 0$ e $\det \nabla^2 f(0) = \lambda_1 \cdot \lambda_2 < 0$.

Como o produto dos autovalores é negativo, necessariamente um autovalor é positivo e o outro é negativo. Logo, a matriz $\nabla^2 f(0)$ é indefinida, o que implica que $x^* = 0$ é um ponto de sela de f.

Para mostrar que existe ao menos outro ponto estacionário, provaremos que f é coerciva e, portanto, possui um minimizador global. De fato, consideremos $\|x\|_2^2 = x_1^2 + x_2^2$. Então, podemos reescrever

$$f(x) = ax_1x_2 + x_1^3 + x_2^3 + e^{\|x\|^2}.$$

Quando $\|x\| \to +\infty$, equivalentemente temos $\|x\|_2^2 \to \infty$ (pois a função $t \mapsto t^2$ é estritamente crescente e continua), analisemos o comportamento de cada termo.

Perceba que

$$0 \le (|x_1| - |x_2|)^2 = |x_1|^2 - 2|x_1||x_2| + |x_2|^2 \implies 2|x_1||x_2| \le |x_1|^2 + |x_2|^2 = x_1^2 + x_2^2$$
$$\implies |x_1||x_2| \le \frac{x_1^2 + x_2^2}{2} \underbrace{\implies}_{a>0} a|x_1||x_2| \le \frac{a}{2}(x_1^2 + x_2^2) = \frac{a}{2}\|x\|_2^2$$

então quando $\|x\|_2^2 \to \infty$, obtemos que o termo $|ax_1x_2| \le \frac{a}{2}\|x\|_2^2 \to \infty$.

Por outro lado, note que

$$|x_1^3| + |x_2^3| = |x_1|^3 + |x_2|^3 \le (|x_1| + |x_2|)^3 \le (2\|x\|)^3 = 8\|x\|^3$$

quando $\|x\|_2^2 \to \infty$, temos o termo $|x_1^3| + |x_2^3| \le 8\|x\|_2^3 \to \infty$.

Finalmente

$$f(x) \ge e^{\|x\|^2} - \frac{a}{2}\|x\|^2 - 8\|x\|^3.$$

Como a função exponencial $e^{\|x\|^2}$ cresce muito mais rapidamente que os termos polinomiais $\|x\|^2$ e $\|x\|^3$, segue que $f(x) \to +\infty$ quando $\|x\| \to +\infty$. Portanto, f é coerciva. Como f é contínua e coerciva, f tem um mínimo global em algum ponto $\tilde{x} \in \mathbb{R}^2$. Necessariamente, $\tilde{x}$ é um ponto estacionário de f (pois $\nabla f(\tilde{x}) = 0$). Além disso, como $x^* = 0$ é um ponto de sela, temos que $\tilde{x} \neq 0$.

Portanto, f possui pelo menos dois pontos estacionários, $x^* = 0$ (ponto de sela) e $\tilde{x}$ (minimizador global). ☺

Exercício 9.

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ contínua, $x^* \in \mathbb{R}^n$ e $f^* = f(x^*)$. Suponha que todo x, tal que $f(x) = f^*$, é um minimizador local de f. Mostre que x^* é um minimizador global de f.

Exercício 10.

Considere $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x) = a \operatorname{sen} x_1 \operatorname{sen} x_2 + e^{x_1^2+x_2^2}$, onde $a \in \mathbb{R}$ e $\bar{x} = 0 \in \mathbb{R}^2$.

a) Mostre que se $a \in (-2, 2)$, então $\bar{x}$ é minimizador local de f.

b) Prove que se $|a| > 2$, então $\bar{x}$ é ponto de sela de f.

c) Verifique se para algum valor de a, o ponto $\bar{x}$ pode ser maximizador local de f.

Solução:

É claro que f é de classe C^2 em $\mathbb{R}^2$.

(a) Denotemos $x_1^2 + x_2^2 = u$. Assim,

$$\nabla f(x) = \begin{bmatrix} a\cos x_1 \sin x_2 + 2x_1 e^u \\ a\cos x_2 \sin x_1 + 2x_2 e^u \end{bmatrix},$$

perceba que $\bar{x} = 0$ é um ponto estacionário pois $\nabla f(\bar{x}) = 0$. Agora estudemos a natureza do hessiano nesse ponto $\bar{x}$,

$$\nabla^2 f(x) = \begin{bmatrix} -a\sin x_1 \sin x_2 + 2e^u + 4x_1^2 e^u & a\cos x_1 \cos x_2 + 4x_1x_2 e^u \\ a\cos x_1 \cos x_2 + 4x_1x_2 e^u & -a\sin x_2 \sin x_1 + 2e^u + 4x_2^2 e^u \end{bmatrix}.$$

Avaliando no ponto $\bar{x} = 0$, obtemos

$$\nabla^2 f(\bar{x}) = \begin{bmatrix} 2 & a \\ a & 2 \end{bmatrix},$$

calculamos o determinante: $\det \nabla^2 f(\bar{x}) = 4 - a^2$.

Suponhamos que $a \in (-2, 2)$, então

$$|a| < 2 \Longrightarrow |a|^2 < 4 \Longrightarrow a^2 < 4 \Longrightarrow 0 < 4 - a^2 = \det \nabla^2 f(\bar{x}) \Longrightarrow \det \nabla^2 f(\bar{x}) > 0,$$

além disso, a determinante da sub-matriz [2] é positiva. Assim, pelo critério de Sylvester, temos que $\nabla^2 f(\bar{x})$ é positiva definida. Logo $\bar{x} = 0$ é um minimizador local de f quando $a \in (-2, 2)$.

(b) Suponhamos agora que $|a| > 2$, analogamente do item anterior obtemos $\det \nabla^2 f(\bar{x}) < 0$ y como $\det \nabla^2 f(\bar{x}) = \lambda_1 \cdot \lambda_2$ (onde λ_i são autovalores de $\det \nabla^2 f(\bar{x})$), então $\lambda_1 \cdot \lambda_2 < 0$, além disso $\operatorname{tr} \nabla^2 f(\bar{x}) = 4 = \lambda_1 + \lambda_2 > 0$ o que implica que pelo menos algum autovalor é positiva e outro é negativo, sem perda de generalidade consideremos $\lambda_1 > 0$ e $\lambda_2 < 0$, logo a matriz $\nabla^2 f(\bar{x})$ é indefinida. Portanto $\bar{x}$ é um ponto de sela de f.

(c) Para que $\bar{x}$ seja um ponto máximo local de f, precisamos que $\nabla^2 f(\bar{x})$ seja negativa definida ($\nabla^2 f(\bar{x}) \prec 0$), para que todos os autovalores sejam negativos $\lambda_1, \lambda_2 < 0$. Mas $\operatorname{tr} \nabla^2 f(\bar{x}) = 4 > 0$, então não é possível que a soma de dois números negativos sejam 4. Portanto, não é possível obter um máximo local para algum valor de a no ponto $\bar{x}$.☺

Exercício 11.

Mostre que a função $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x) = x_1^2 + x_2 + e^{x_1^2 + x_2^2}$$

possui um único ponto estacionário e que tal ponto é um minimizador global.

<u>**Solução:**</u>

Observemos inicialmente que f é de classe $C^2(\mathbb{R}^2)$.

Denotemos $u = x_1^2 + x_2^2$. Assim,

$$\nabla f(x) = \begin{pmatrix} 2x_1 + 2x_1 e^u \\ 1 + 2x_2 e^u \end{pmatrix}.$$

Os pontos estacionários satisfazem o sistema

$$\begin{cases} 2x_1(1 + e^u) = 0, \\ 1 + 2x_2 e^u = 0. \end{cases}$$

Da primeira equação, como $1 + e^u > 0$ para todo $u \in \mathbb{R}$, obtemos necessariamente

$$x_1 = 0.$$

Substituindo na segunda equação, temos $u = x_2^2$ e

$$1 + 2x_2 e^{x_2^2} = 0.$$

Definimos então a função auxiliar

$$g(t) = te^{t^2} + \frac{1}{2}, \quad t \in \mathbb{R}.$$

Note que a equação acima é equivalente a $g(t) = 0$.

Calculando a derivada,

$$g'(t) = e^{t^2}(1 + 2t^2) > 0 \quad \forall t \in \mathbb{R},$$

logo g é estritamente crescente e contínua. Além disso,

$$\lim_{t \to -\infty} g(t) = -\infty, \qquad \lim_{t \to +\infty} g(t) = +\infty.$$

Pelo Teorema do Valor Intermediário, existe um único $t_0 \in \mathbb{R}$ tal que

$$g(t_0) = 0.$$

Portanto, existe um único ponto estacionário de f, dado por

$$x^* = (0, t_0).$$

Por outro lado, o Hessiano de f é

$$\nabla^2 f(x) = \begin{pmatrix} 2 + 2e^u + 4x_1^2 e^u & 4x_1 x_2 e^u \\ 4x_1 x_2 e^u & 2e^u + 4x_2^2 e^u \end{pmatrix}.$$

Avaliando em $x^* = (0, t_0)$, obtemos

$$\nabla^2 f(x^*) = \begin{pmatrix} 2 + 2e^{t_0^2} & 0 \\ 0 & 2e^{t_0^2}(1 + 2t_0^2) \end{pmatrix}.$$

Os menores principais são

$$2 + 2e^{t_0^2} > 0, \qquad \det \nabla^2 f(x^*) = (2 + 2e^{t_0^2})\, 2e^{t_0^2}(1 + 2t_0^2) > 0.$$

Pelo critério de Sylvester, $\nabla^2 f(x^*)$ é positiva definida. Logo,

$$x^* \text{ é um minimizador local de } f.$$

Finalmente, mostremos que f é coerciva. Observemos que

$$f(x) = x_1^2 + x_2 + e^{x_1^2 + x_2^2} \geq x_1^2 + \left(x_2 + e^{x_2^2}\right),$$

pois $e^{x_1^2+x_2^2} \geq e^{x_2^2}$ para todo $x \in \mathbb{R}^2$.

Definamos $h(t) = t + e^{t^2}$. Como

$$\lim_{|t|\to\infty} h(t) = +\infty,$$

segue que $x_2 + e^{x_2^2} \to +\infty$ quando $|x_2| \to \infty$. Além disso, $x_1^2 \to +\infty$ quando $|x_1| \to \infty$.

Portanto,

$$\lim_{\|x\|\to\infty} f(x) = +\infty,$$

isto é, f é coerciva.☹