Aula 1

Definições:

Problema de Programação matemática

$$\text{Min } f(x)$$
$$\text{s.a } \quad x \in S$$

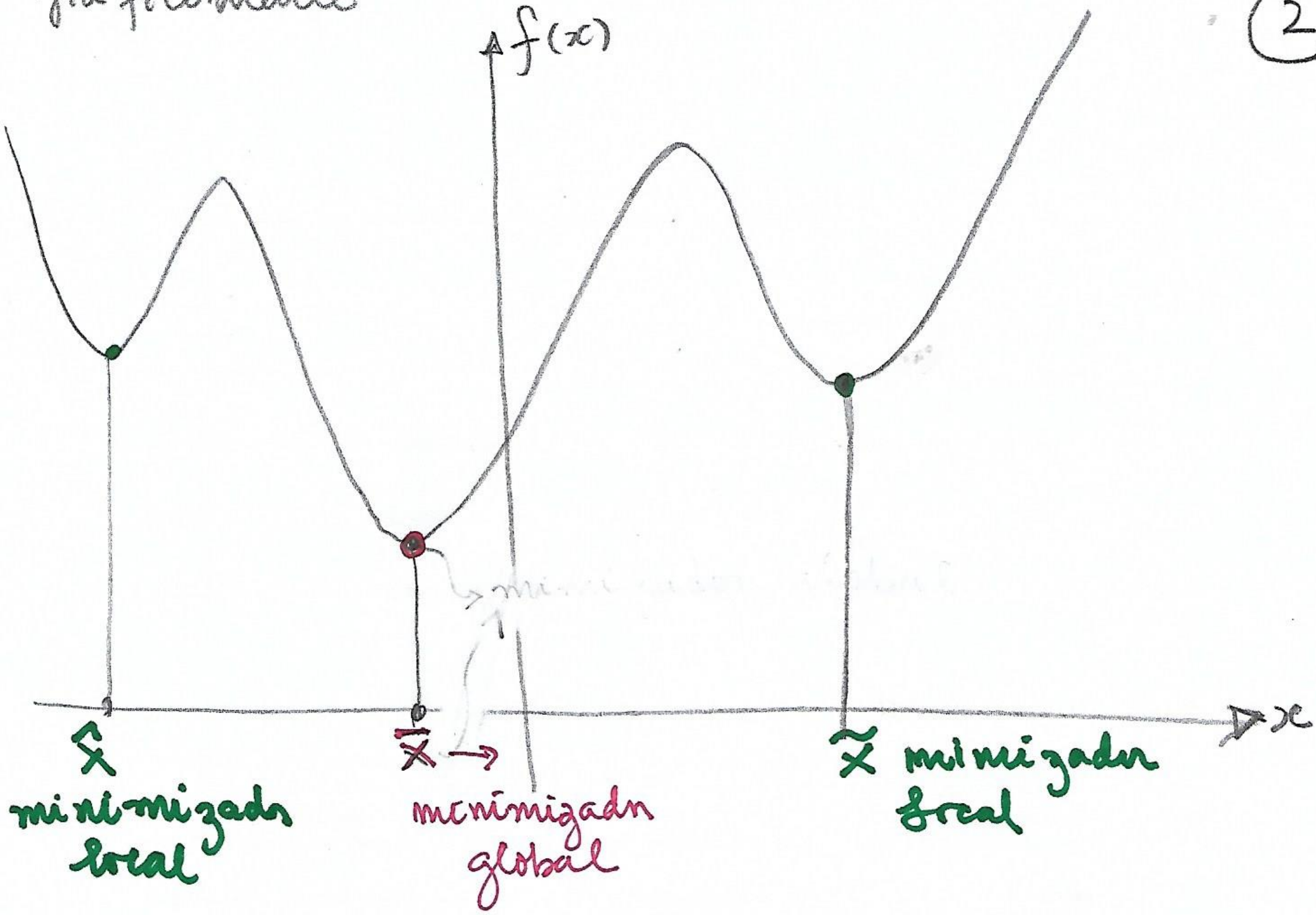
onde $f: \mathbb{R}^n \to \mathbb{R}$ e $S \subset \mathbb{R}^n$. S é chamado de conjunto factível

dois tipos de Soluções

1) Um ponto $x^* \in S$ é um minimizador local de f em $S \iff \exists\, \varepsilon > 0$ tal que $f(x) \geq f(x^*)$ $\forall\, x \in S$ tal que $\|x - x^*\| < \varepsilon$
Se $f(x) > f(x^*)$ para todo $x \in S$ tal que $x \neq x^*$ e $\|x - x^*\| < \varepsilon$ diremos que se trata de um minimizador local estrito em S.

2) Um ponto $x^* \in S$ é um minimizador global de f em $S \iff f(x) \geq f(x^*)$ para todo $x \in S$. Se $f(x) > f(x^*)$ para todo $x \in S$ com $x \neq x^*$ diremos que se trata de um minimizador global estrito em S

graficamente



Teorema de existencia: S um conjunto fechado e limitado e f uma função contínua então f admite um minimizador global em S

(Bolzano e Weierstrass)

Condições de Otimalidad Para minimização sem restrições

Para que servem: Para identificar candidatos a mínimos

Necessarias

Se x^* é um minimizador $\Rightarrow$ Se cumpre a condição

Suficiente

Se se cumpre a condição $\Rightarrow$ x^* é um minimizador

Necessarias $\quad \text{Min } f(x)$

1) $x \in \mathbb{R}^n$ (não serve de nada)

3) x é um minimizador (obvia)

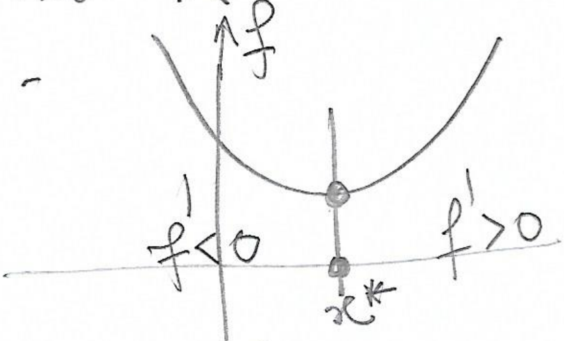
Procuramos condições Matematicas para identificar um mínimo

$f: \mathbb{R} \to \mathbb{R}$	$f: \mathbb{R}^2 \to \mathbb{R}$	$f: \mathbb{R}^n \to \mathbb{R}$
Em $\mathbb{R}$ conhecidos e faceis	desenho e entender graficamente Possíveis graficos $\mathbb{R}^3$	Generalização

interpretar

generalizar

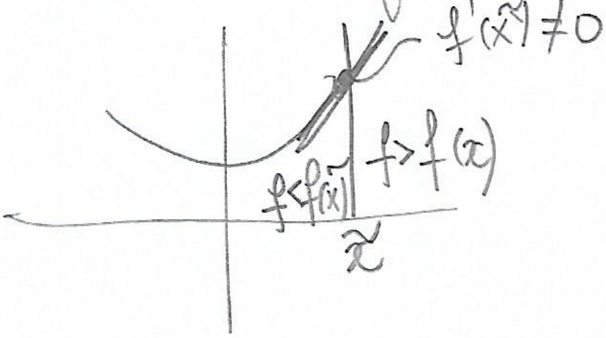
em $\mathbb{R}$



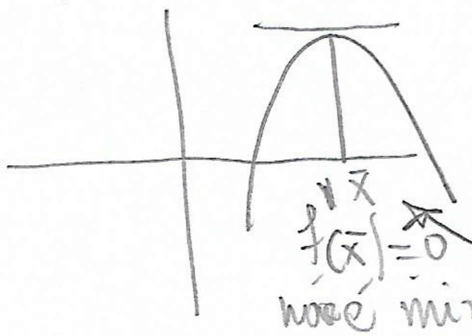
⇒ Condição necessária primeira ordem (1ª derivada)

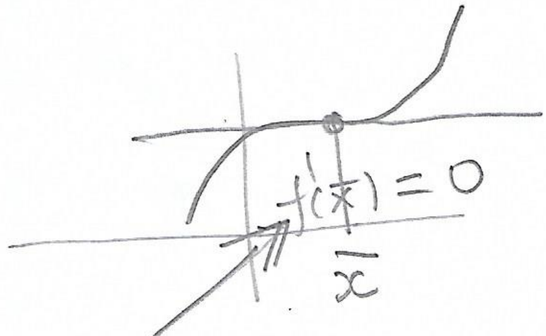
$$f'(x^*) = 0 \quad \boxed{\text{candidato a minimizador}}$$

Se não a verifica não é minimizador



Porque necessária





Condição necessária de segunda ordem
(segunda derivados)

$$f'(x^*) = 0 \quad e \quad f''(x^*) \geq 0$$

Se não se verifica não é minimizador

$f(x) = x^3$ em $x=0$ não é minimizador

$f'(0) = f''(0) = 0$

Condição suficiente de segunda ordem

$$f'(x^*) = 0 \quad e \quad f''(x^*) > 0$$

se $f'(x^*) = 0$ e $f''(x^*) > 0$ então x^* é um minimizador local único

$$f(y) = f(x^*) + \cancel{f'(x^*)(y-x^*)} + \frac{f''(\xi)}{2}(y-x^*)^2 \qquad \xi \in ((y, x^*))$$

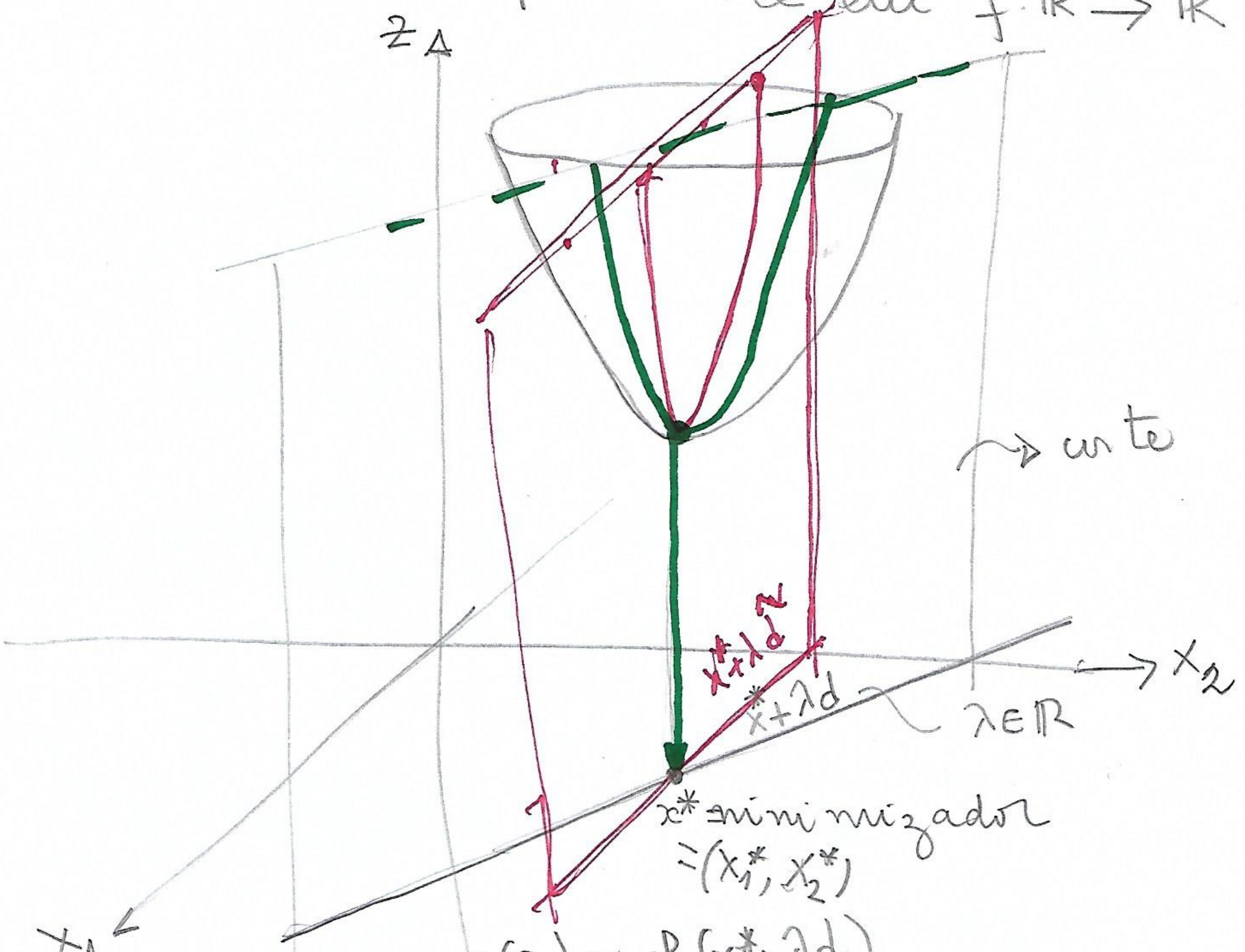
$\Rightarrow$ $\exists\ \delta > 0$ tal que para $\|z - x^*\| < \delta$

$f''(z) > 0 \Rightarrow$

$$f(y) > f(x^*) + \underbrace{\frac{f''(\xi)}{2}}_{>0}\underbrace{(y-x^*)^2}_{>0 \text{ se } y \neq x^*} > f(x^*)$$

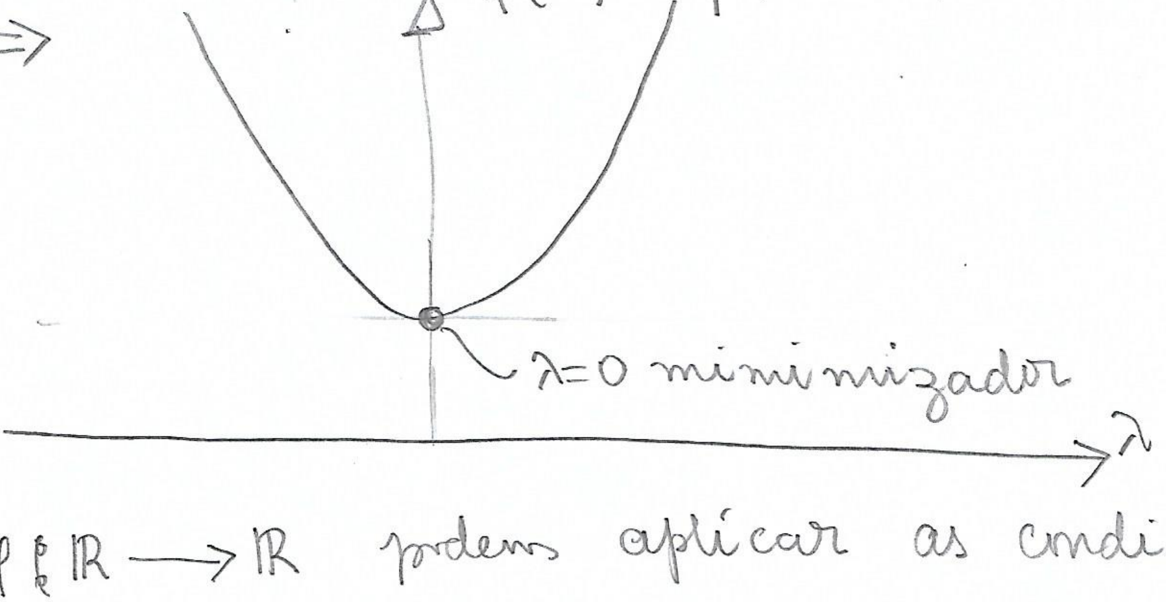
$\Rightarrow$ $f(y) > f(x^*)$ se $|y - x^*| < \delta$ e $y \neq x^*$

Vamos a entender que acontece em $f: \mathbb{R}^2 \to \mathbb{R}$



$$\varphi(\lambda) = f(x^* + \lambda d)$$

$\Rightarrow$



$\varphi: \mathbb{R} \to \mathbb{R}$ podemos aplicar as condições de otimalidade de funções em $\mathbb{R}$

$$\varphi(\lambda) = f(\overbrace{(x_1^*, x_2^*)}^{x^*} + \lambda \overbrace{(d_1, d_2)}^{d}) = f(\underbrace{x_1^* + \lambda d_1}, \underbrace{x_2^* + \lambda d_2})$$

regra da

Regra da Cadeia

$$\varphi'(\lambda) = \frac{\partial f}{\partial x_1}(x^* + \lambda d) \cdot d_1 + \frac{\partial f}{\partial x_2}(x^* + \lambda d) \cdot d_2 = \underbrace{\left\langle \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right), (d_1, d_2) \right\rangle}_{\text{Produto interno}} =$$

$$= \nabla f(x^* + \lambda d)^T \cdot d$$

mas

$$\varphi'(0) = 0 \Rightarrow \nabla f(x^*)^T d = 0$$

isto acontece para toda $d \in \mathbb{R}^2 \Rightarrow$

$$\boxed{\nabla f(x^*) = 0}$$

Condição de otimalidade necessaria de 1ra. ordem.

Condições de otimalidade Necessarias de segunda ordem

temos $\varphi'(0) = 0$ e $\varphi''(0) \geq 0$

calculamos $\varphi''(\lambda)$

temos

$$\varphi'(\lambda) = \frac{\partial f}{\partial x_1}(x^* + \lambda d) \, \overset{\text{cte}}{d_1} + \frac{\partial f}{\partial x_2}(x^* + \lambda d) \, \overset{\text{cte}}{d_2}$$

$$\varphi''(\lambda) = \left[\frac{\partial^2 f}{\partial x_1^2}(x^* + \lambda d) \, d_1 + \frac{\partial^2 f}{\partial x_1 \partial x_2}(x^* + \lambda d) \, d_2 \right] d_1 +$$

$$\left[\frac{\partial^2 f}{\partial x_2 \partial x_1}(x^* + \lambda d) \, d_1 + \frac{\partial^2 f}{\partial x_2^2}(x^* + \lambda d) \, d_2 \right] d_2 =$$

$$\varphi''(\lambda) = \frac{\partial^2 f}{\partial x_1^2}(x^* + \lambda d)\, d_1^2 + \frac{\partial^2 f(x^*+\lambda d)}{\partial x_1 \partial x_2} d_1 . d_2 + \frac{\partial^2 f(x^*+\lambda d)}{\partial x_2 \partial x_1} d_1 d_2 +$$

$$+ \frac{\partial^2 f}{\partial x_2^2}(x^* + \lambda d)\, d_2^2 =$$

$$= \underbrace{(d_1 , d_2)}_{d^T} \underbrace{\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}}_{H(x^* + \lambda d)} \underbrace{\begin{pmatrix} d_1 \\ d_2 \end{pmatrix}}_{d}$$

$$\Rightarrow$$

$$\boxed{\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} \quad \text{Cauchy-Schwarz}}$$

$$\Rightarrow \varphi''(0) = d^T H(x^*)\, d \geq 0$$

$\Rightarrow$ vale para todo d

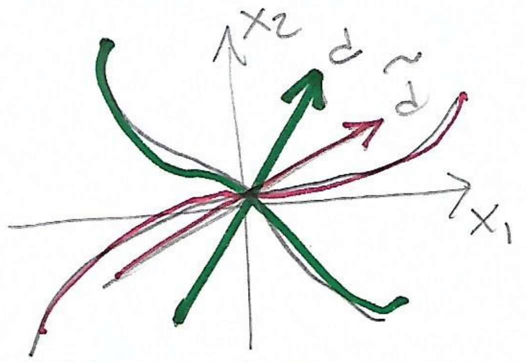
$\Rightarrow$ $H(x^*)$ é semidefinida positiva

<u>Condição suficiente</u> de Segunda ordem

Não alcança com $\boxed{\varphi'(0) = 0 \text{ e } \varphi''(0) > 0}$

abordar o tema por retas

Porque estamos tomando retas e não curvas



Vamos a provar a condição suficiente de segunda ordem.

Se $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*) = H(x^*) > 0$

então x^* é um minimizador de $f(x)$

Prova

definimos a seguinte função $\Gamma : \mathbb{R}^2 \to \mathbb{R}$

$$\Gamma(h) = h^T \nabla^2 f(x^*) h$$

Seja $B = \{ h \in \mathbb{R}^n / \|h\| = 1 \}$ Bola unitária
Fechada e limitada

$$\Rightarrow \quad \Gamma(h) > a > 0 \quad \forall\, h \in B$$

Agora temos que

$$f(x^* + d) = f(x^*) + \cancelto{0}{\nabla f(x^*)^T} d + \frac{1}{2} d^T \nabla^2 f(x^*) d + o(\|d\|^2)$$

$$\Rightarrow \quad f(x^* + d) - f(x^*) = \frac{1}{2} \|d\|^2 \left(\frac{d^T}{\|d\|} \nabla^2 f(x^*) \frac{d}{\|d\|} \right) + o(\|d\|^2)$$

$$> \frac{1}{2} a \|d\|^2 + o(\|d\|^2) = \|d\|^2 \left(\frac{a}{2} + \frac{o(\|d\|^2)}{\|d\|^2} \right)$$

$\Rightarrow$ para $\exists\, \varepsilon > 0 \quad \|d\| < \varepsilon \Rightarrow \frac{a}{2} + \frac{o(\|d\|^2)}{\|d\|^2} > 0$

$\Rightarrow \quad f(x^* + d) - f(x^*) > 0$ para todo

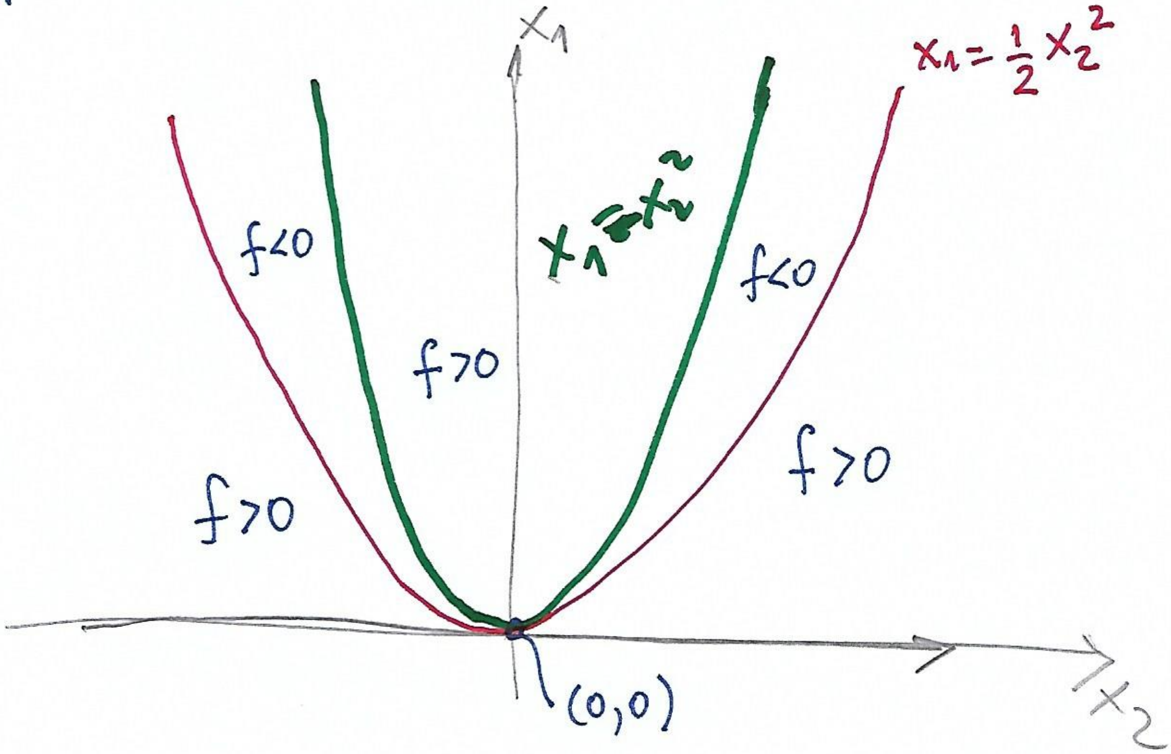
$x \in B(x^*; \varepsilon) = x \neq x^*$

Exercicio:

Exemplo:

Analizar a função no ponto (0.0)

$$f(x) = (x_1 - x_2^2)\left(x_1 - \frac{1}{2}x_2^2\right)$$



$$\frac{\partial f}{\partial x_1} = \left(x_1 - \frac{1}{2}x_2^2\right) + (x_1 - x_2^2) = 2x_1 - \frac{3}{2}x_2^2$$

$$\frac{\partial f}{\partial x_1}(0,0) = 0$$

$$\frac{\partial f}{\partial x_2} = -2x_2\left(x_1 - \frac{1}{2}x_2^2\right) + -x_2(x_1 - x_2^2) =$$

$$= -3x_1x_2 + 2x_2^3$$

$$\frac{\partial f}{\partial x_2}(0,0) = 0$$

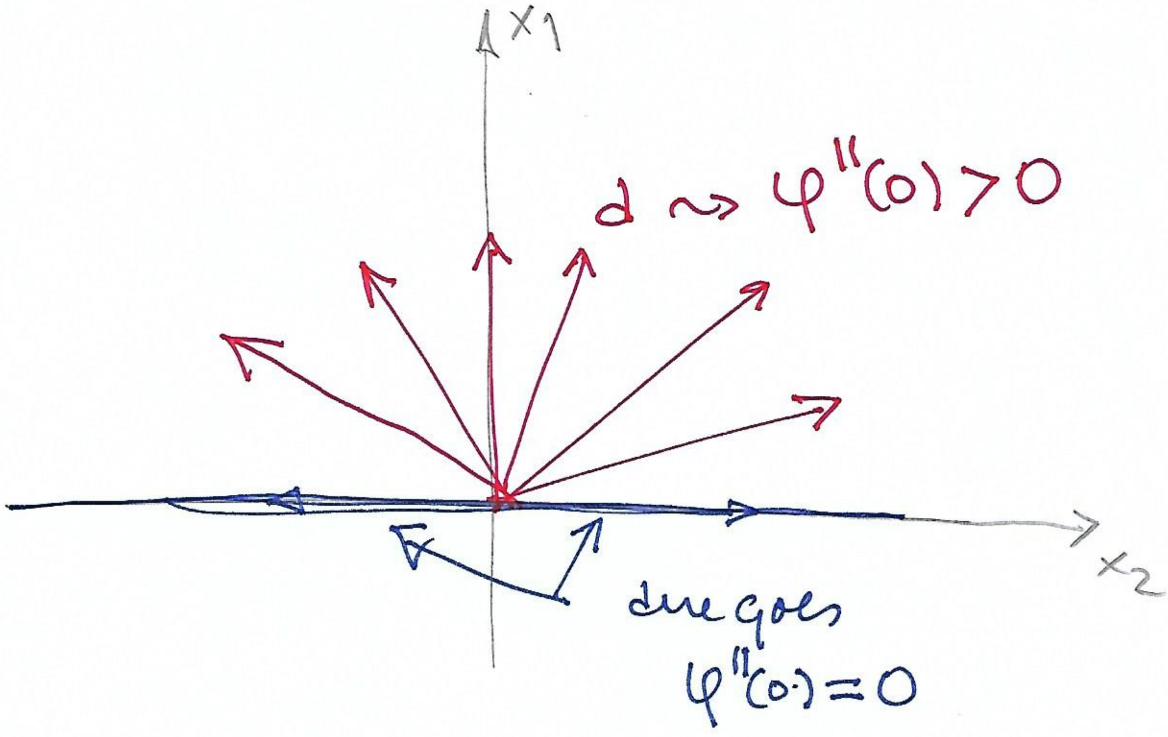
$$\Rightarrow \nabla f(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Segunda ordem

$$H(x_1,x_2) = \begin{bmatrix} 2 & -3x_2 \\ -3x_2 & -3x_1+6x_2^2 \end{bmatrix}$$

$$H(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \quad (d_1, d_2)\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 2d_1^2 \geq 0$$



cumpre a condição necessaria de primeiro e segunda ordem em minimizador ao logo de quelquer direção, ao longo de qualquer direção salvo d=(0,d_2) e a derivada é estrita e em qualquer direção \lambda = 0 e estrito unico!!!. Mas em qualquer curva quadratica entre as quadráticas que realizam o produto da função a função decrece estritamente.