(Aula 2) – Intro. otimização contínua

$\bar{x}$ min. local de f em D:

$$f(\bar{x}) \leq f(x), \quad \forall x \in D \cap B(\bar{x}, \varepsilon)$$

Condições de otimalidade

- Problemas sem restrições: $D = \mathbb{R}^n$

$$(P) \quad \min. \ f(x) \quad \text{s.a} \quad x \in \mathbb{R}^n, \qquad f : \mathbb{R}^n \to \mathbb{R}$$

Notação: $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$
(gradiente) $\quad x \mapsto \nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}$

$\nabla^2 f(x)$ (Hessiana) $= \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \cdots & \frac{\partial^2 f(x)}{\partial x_n \partial x_1} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \cdots & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$

- f dif. em $B(x, \varepsilon)$ então $\nabla^2 f(x)$ é simétrica.

$o(t)$: vai para zero mais rápido que t

$$\lim_{t \to 0} \frac{o(t)}{t} = 0.$$

Usado por exemplo p/ descrever o "resto" de uma expansão de Taylor:

1ª ordem:
$$f(x + td) = f(x) + t \nabla f(x)^T d + o(t)$$

2ª ordem:
$$f(x + td) = f(x) + t \nabla f(x)^T d + \frac{t^2}{2} d^T \nabla^2 f(x) d + o(t^2)$$

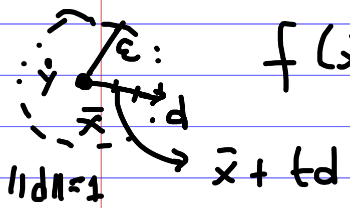
- Produto interno: $\langle \nabla f(x), d \rangle = \nabla f(x)^T d$

Hip: $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável

(CN1) **Teorema 1:** Seja $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável em $\bar{x} \in \mathbb{R}^n$. Se $\bar{x}$ é mín. local de f em $\mathbb{R}^n$, então $\boxed{\nabla f(\bar{x})^T d \geq 0, \ \forall d \in \mathbb{R}^n.}$

$(\hookrightarrow \Rightarrow \nabla f(\bar{x}) = 0)$

Dem. Seja $d \in \mathbb{R}^n$ arbitrária, mas fixa. Se $\bar{x}$ é mín. local de f em $\mathbb{R}^n$, então $\exists \, \varepsilon > 0$ tal que



$$f(\bar{x}) \leq f(\bar{x} + td), \ \forall t \in (0, \varepsilon]. \qquad (\mathrm{I})$$

Por Taylor (1ª ordem)

$$f(\bar{x} + td) = f(\bar{x}) + t \nabla f(\bar{x})^T d + o(t) \qquad (\mathrm{II})$$

De (I) e (II)

$$0 \leq t \nabla f(\bar{x})^T d + o(t)$$

$$\overset{t>0}{\Rightarrow} \quad 0 \leq \nabla f(\bar{x})^T d + \frac{o(t)}{t}$$

$$\Rightarrow \quad 0 \leq \lim_{t \to 0} \left(\nabla f(\bar{x})^T d + \frac{o(t)}{t} \right)$$

$$= \nabla f(\bar{x})^T d + \underbrace{\lim_{t \to 0} \frac{o(t)}{t}}_{=0}$$

$$= \nabla f(\bar{x})^T d$$

Como d era fixa porém arbitrária, temos que

$$0 \leq \nabla f(\bar{x})^T d, \quad \forall d \in \mathbb{R}^n.$$

#

OBS.: $d = -\nabla f(\bar{x}) \Rightarrow 0 \leq \nabla f(\bar{x})^T (-\nabla f(\bar{x}))$

$$= -\|\nabla f(\bar{x})\|^2$$

$$\Rightarrow \|\nabla f(\bar{x})\| \leq 0$$

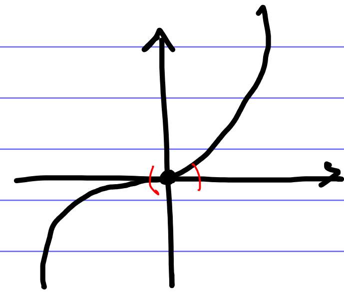
$$\Rightarrow \nabla f(\bar{x}) = 0 \ //$$

CN1: $\nabla f(\bar{x}) = 0$ é apenas necessária

e.g: $f(x) = x^3$, $\nabla f(x) = f'(x) = 3x^2$

$\bar{x} = 0$, $f'(0) = 0$

mas $\bar{x}$ não é min. local



OBS: pontos que satisfazem CN1 são chamados "estacionários".

(CN2) Teorema 2: Seja $f: \mathbb{R}^n \to \mathbb{R}$ duas vezes difer// em $\bar{x} \in \mathbb{R}$.

Se $\bar{x}$ é min. local de f, então

$\nabla f(\bar{x}) = 0$ e

$\langle \nabla^2 f(\bar{x}) d, d \rangle = d^T \nabla^2 f(\bar{x}) d \geq 0, \quad \forall d \in \mathbb{R}^n.$

Dem. $0 \leq f(\bar{x} + td) - f(\bar{x})$, d fixo, $t \in (0, \varepsilon]$

$$= t \underbrace{\nabla f(\bar{x})^T d}_{\substack{=0 \\ (CN1)}} + \frac{t^2}{2} d^T \nabla^2 f(\bar{x}) d + o(t^2)$$

$$\Rightarrow \quad 0 \leq \frac{1}{2} d^T \nabla^2 f(\bar{x}) d + \frac{o(t^2)}{t^2} \xrightarrow{t \to 0} \frac{1}{2} d^T \nabla^2 f(\bar{x}) d$$

$$\Rightarrow \quad d^T \nabla^2 f(\bar{x}) d \geq 0$$

#

• $\nabla^2 f(\bar{x})$ é positiva semidefinida.

ATENÇÃO: CN2 não é suficiente.

e.g., $f(x) = x^3$, $\bar{x} = 0$, $f'(x) = 3x^2$, $f''(x) = 6x$

$f'(0) = 0$, $f''(0) = 0$.

Teorema 3 (cond. suficiente) Se $f: \mathbb{R}^n \to \mathbb{R}$ é duas vezes dif. ~~em~~ $\bar{x} \in \mathbb{R}^n$ e

$\nabla f(\bar{x}) = 0$ e $d^T \nabla^2 f(\bar{x}) d > 0$, $\forall d \in \mathbb{R}^n \setminus \{0\}$

então $\bar{x}$ é min. local estrito de f.

Dem.
$$\boxed{\text{Afirmação: } \inf_{d \neq 0} \frac{d^T \nabla^2 f(\bar{x}) d}{\|d\|^2} := \beta > 0} . \quad (\text{III})$$

Suponha que $\bar{x}$ não é min. local. Então existe $\{x^k\} \subset \mathbb{R}^n$, $x^k \to \bar{x}$

tal que $f(x^k) < f(\bar{x})$, p/ k suficiente grande.

Para k suf. grande.

$$0 > f(x^k) - f(\bar{x})$$

Taylor 2ª ordem em torno de $\bar{x}$

$$= \underbrace{\nabla f(\bar{x})^T (x_k - \bar{x})}_{=0} + \frac{1}{2}(x^k - \bar{x})^T \nabla^2 f(\bar{x})(x^k - \bar{x}) + o(\|x^k - \bar{x}\|^2)$$

$$\Rightarrow 0 > \frac{1}{2} \frac{(x^k - \bar{x})^T}{\|x^k - \bar{x}\|} \nabla^2 f(\bar{x}) \frac{(x^k - \bar{x})}{\|x^k - \bar{x}\|} + \frac{o(\|x^k - \bar{x}\|^2)}{\|x^k - \bar{x}\|^2}$$

tomando limite q $k \to +\infty$:
(se necessário em subsequência convergente)

$$0 \geq \frac{1}{2} d^T \nabla^2 f(\bar{x}) d \geq \frac{1}{2}\beta \Rightarrow \beta \leq 0$$

$(\|d\| = 1)$ o que contradiz (III).

Portanto, $\bar{x}$ é min. local de f em $\mathbb{R}^n$.

Exercício: Usando (III) prove que $\exists \varepsilon > 0$ e $\gamma > 0$, tal que *

(IV) $f(x) - f(\bar{x}) \geq \gamma \|x - \bar{x}\|^2$, $\forall x \in B(\bar{x}, \varepsilon)$.

(isto é conhecido como "crescimento quadrático")

$\to f(x) - f(\bar{x}) > 0$, $\forall x \in B(\bar{x}, \varepsilon) \setminus \{\bar{x}\}$

Problemas de otimização com restrições

$$(P) \quad \min. \ f(x) \quad \text{s.a} \quad x \in D$$

$\emptyset \neq D \subsetneq \mathbb{R}^n$.

Def.: Chamamos de cone um conjunto $K \subseteq \mathbb{R}^n$ tal que

$$d \in K \Rightarrow td \in K, \ \forall t \geq 0.$$

Obs.: se $K \neq \emptyset$, $0 \in K$.

$K \neq \emptyset \Rightarrow K$ ilimitado

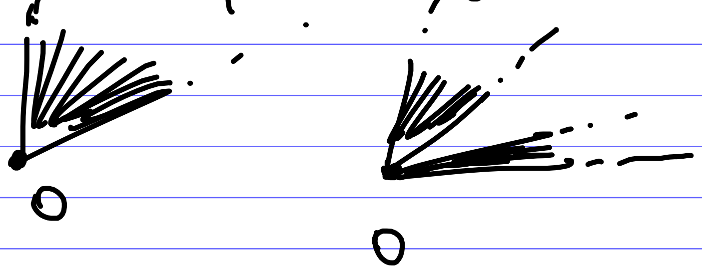
Exemplos:

(i) $\mathbb{R}^n$

(ii) $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0, i = 1, \dots, n\}$

(iii) $\{(x, \eta) \in \mathbb{R}^n \times \mathbb{R}_+ \mid \|x\|_2 \leq \eta\}$

(iv) S^n_+: matrizes simétricas positivas semidefinidas

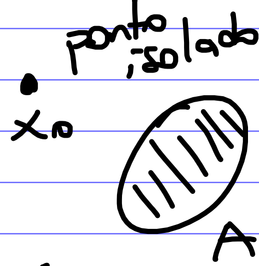


Def.: Dizemos que $d \in \mathbb{R}^n$ é direção viável para D em $\bar{x} \in D$ se $\exists \varepsilon > 0$ tal que

$$\bar{x} + td \in D, \ \forall t \in [0, \varepsilon),$$

Notação: $V_D(\bar{x})$ conjunto de todas as direções viáveis p/ D em $\bar{x} \in D$.

OBS.



$V_D(x_0) = \emptyset$

$D = \{x_0\} \cup A$

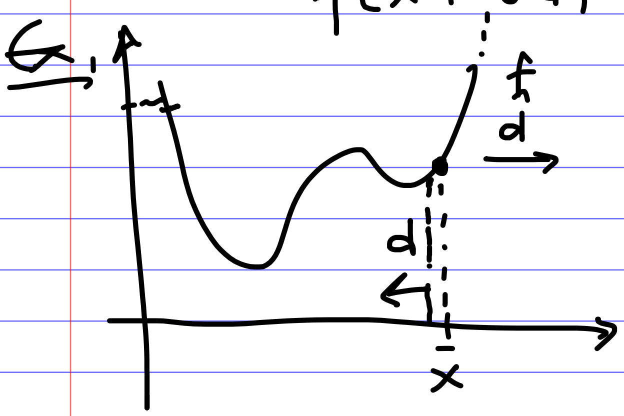
OBS: $V_D(\bar{x})$ é um cone.

Def.: Chamamos $d \in \mathbb{R}^n$ de direção de descida (para f) a partir de $\bar{x} \in \mathbb{R}^n$ quando $\exists \varepsilon > 0$ tal que

$$f(\bar{x} + td) < f(\bar{x}), \quad \forall t \in (0, \varepsilon].$$

tamanho de passo

Ex.:



Notação: $D_f(\bar{x})$: conjunto de todas as direções de descida.

Exercício: (i) $D_f(\bar{x}) \cup \{0\}$ é um cone

(ii) Se $D_f(\bar{x}) \neq \emptyset$, então $\overline{D_f(\bar{x})}$ é um cone.

Lema 1. Seja $f:\mathbb{R}^n \to \mathbb{R}$ diferenciável em $\bar{x} \in \mathbb{R}^n$. Então

(i) $d \in D_f(\bar{x}) \Rightarrow \nabla f(\bar{x})^T d \leq 0$.

(ii) $\nabla f(\bar{x})^T d < 0 \Rightarrow d \in D_f(\bar{x})$.

Dem. (i) $0 > f(\bar{x}+td) - f(\bar{x})$

$$= t\nabla f(\bar{x})^T d + o(t)$$

$$\Rightarrow \quad 0 \geq \nabla f(\bar{x})^T d.$$

(ii) exercício.

#

se $\exists\, d \in D_f(\bar{x})$, $d \in V_D(\bar{x})$

$\bar{x} + td \in D$ p/ t suf. peq.

$f(\underbrace{\bar{x}+td}_{\in D \cap B(\bar{x},\varepsilon)}) < f(\bar{x})$ p/ t suf. peq.

$\Rightarrow \bar{x}$ não é min. local.

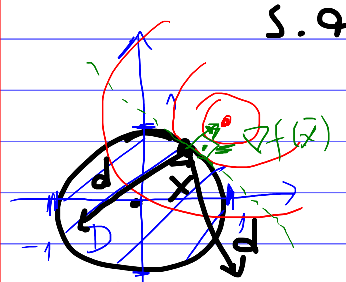
Teorema 4.: Se $\bar{x} \in D$ é min. local de f em D, então $V_D(\bar{x}) \cap D_f(\bar{x}) = \emptyset$.

Corolário 1. $\nabla f(\bar{x})^T d \geq 0$, $\forall d \in V_D(\bar{x})$.
(cond. necessária)

Ex. 2. $\min \ (x_1-1)^2 + (x_2-1)^2$

s.a $\underbrace{x_1^2 + x_2^2 \leq 1}$

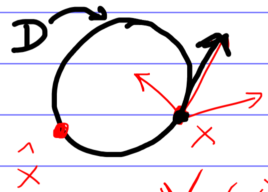
$D = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$



$-\nabla f(\bar{x})^T d \leq 0 \Leftrightarrow \nabla f(\bar{x})^T d \geq 0$

OBS: Há casos em que $V_D(\bar{x}) = \emptyset$ e Teorema 4 e Corolário 1 "deixam de ser úteis".

Ex. 3.
$$\min \ (x_1-1)^2 + (x_2-1)^2$$
$$\text{s.a} \ \ x_1^2 + x_2^2 = 1$$



$x + td \notin D$, $\forall d \neq 0$, $\forall t > 0$

$V_D(x) = \emptyset$, $\forall x \in D$.

Def. Dados $x \in \mathbb{R}^n$, definimos a distância de x ao conjunto D como:
$$\text{dist}(x, D) = \inf_{y \in D} \|y - x\|_2 .$$

Def. Dizemos que $d \in \mathbb{R}^n$ é uma direção tangente a D no ponto $\bar{x} \in D$ quando
$$\text{dist}(\bar{x} + td, D) = o(t), \quad t \in \mathbb{R}_+.$$

Notação: $T_D(\bar{x})$: todas as direções tangentes a D em $\bar{x}$.

- $T_D(\bar{x})$ é um cone não-vazio
- $T_D(\bar{x})$ é chamado cone tangente

OBS: (i) $V_D(\bar{x}) \subset T_D(\bar{x})$

(ii) $\text{int}(D) \neq \emptyset$, $\bar{x} \in \text{int}(D)$
$$\Rightarrow V_D(\bar{x}) = T_D(\bar{x}) = \mathbb{R}^n$$

Proposição 1:

$$T_D(\bar{x}) = \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \forall \{t_k\} \subset \mathbb{R}_+,\ t_k \to 0, \\ \exists \{d_k\},\ d_k \to d, \\ \bar{x} + t_k d_k \in D,\ \forall k \in \mathbb{N} \end{array} \right\}$$

Dem. Seja $d \in \mathbb{R}^n$ tal que $\forall \{t_k\} \subset \mathbb{R}_+$
($\Leftarrow$) $t_k \to 0$, $\exists \{d_k\}$, $d_k \to d$ com $\bar{x} + t_k d_k \in D$.

$$\begin{aligned} \operatorname{dist}(\bar{x} + t_k d, D) &\leq \| \bar{x} + t_k d - (\bar{x} + t_k d_k) \| \\ &= \| t_k (d - d_k) \| = t_k \| d - d_k \| \\ &= o(t_k) \end{aligned}$$

d é direção tangente.

($\Rightarrow$) exercício. #