Otimização – COS360

Programação Não Linear Irrestrita

Convexidade

Prof Luidi Simonetti
luidisimonetti@poli.ufrj.br

Profª Laura Bahiense
laura@cos.ufrj.br

Engenharia de Sistemas e Computação PESC/Coppe

COPPE UFRJ
Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia

Agenda

- Conjuntos convexos
- Funções convexas
- Relembrando matrizes definidas positivas e semidefinidas positivas

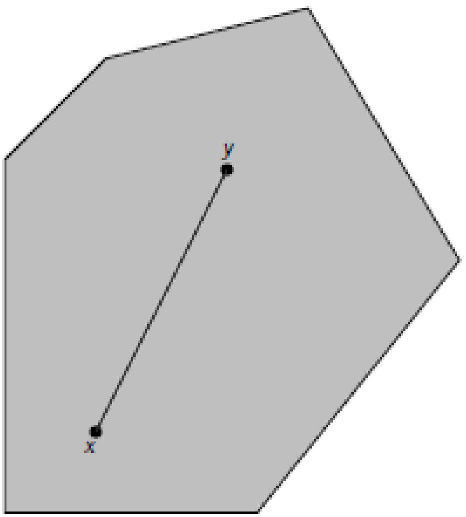
Conjunto convexo

Definição

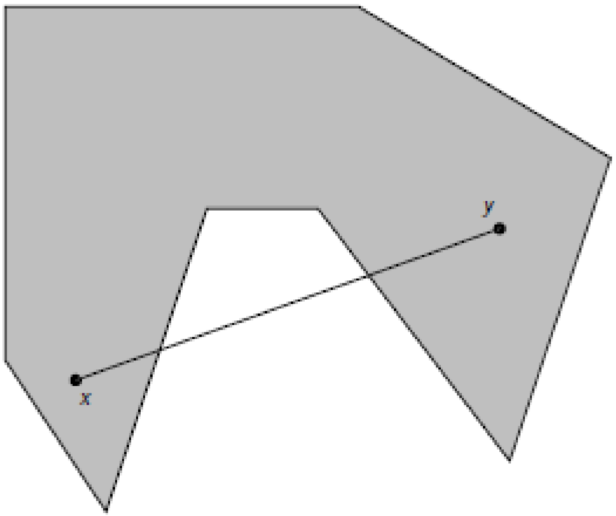
Um conjunto $C \subset \mathbb{R}^n$ é dito convexo quando dados $x, y \in C$, o segmento

$$[x,y] = \{(1-t)x + ty \mid t \in [0,1]\}$$

estiver inteiramente contido em C.



Conjunto convexo



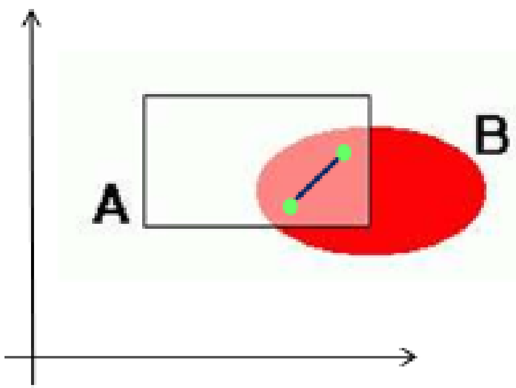
Conjunto não convexo

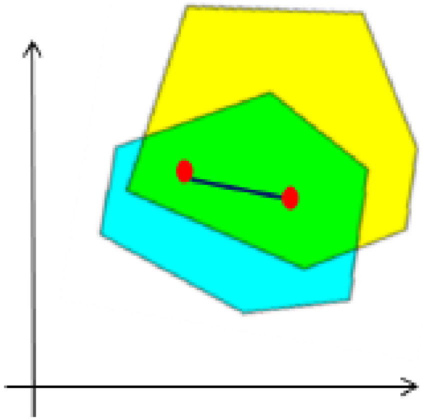
Interseção e união de conjuntos convexos

Interseção × União

Sejam C_i, $i = 1, \ldots, m$ conjuntos convexos.

- O conjunto interseção $C = \bigcap_{i=1}^{m} C_i$ também é convexo.
- Por outro lado, a união de convexos não é convexa.





Exemplo

Exemplo

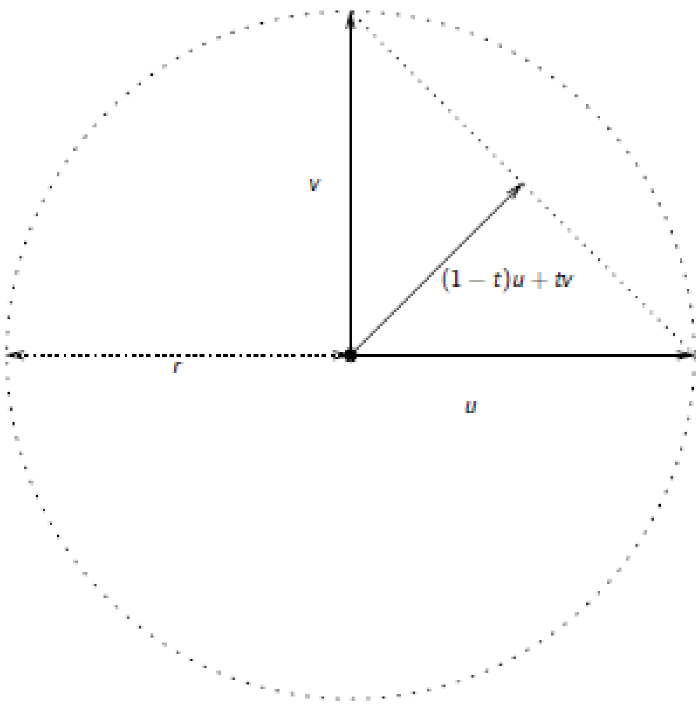
O conjunto solução de um sistema de equações lineares é convexo.

- $C = \{x \in \mathbb{R}^n \mid Ax = b\}$.
- Se $Ax = b$ e $Ay = b$. Fazendo $x' = (1-t)x + ty$.
- $A[(1-t)x + ty] = Ax - tAx + tAy = b - tb + tb = b$.
- Logo o ponto x' também pertence ao conjunto C.

Norma 2 e convexidade

- Sejam $u, v \in \mathbb{R}^n$ com $u \neq v$. Se $\|u\|_2 = \|v\|_2 = r$, então

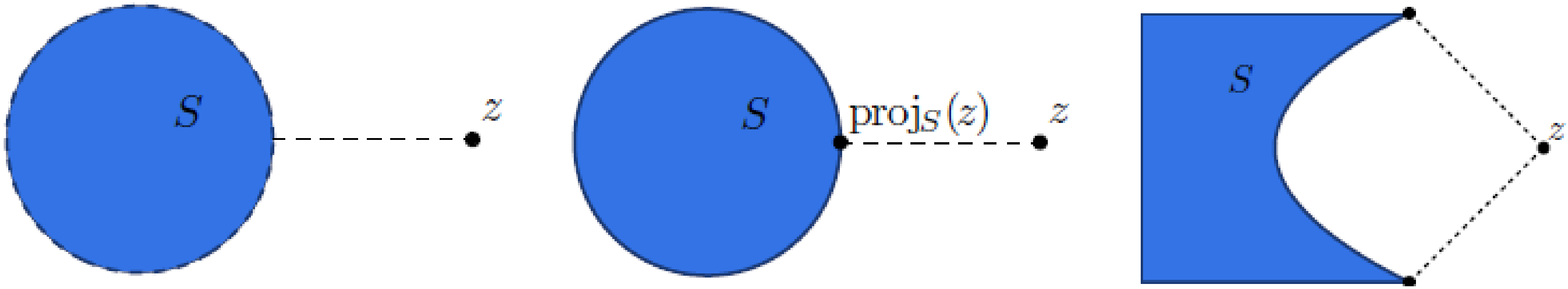
$$\|(1-t)u + tv\|_2 < r, \text{ para todo } t \in (0,1).$$



- Basta aplicar a desigualdade triangular:
  $||(1-t)u + tv||_2 \leq (1-t)\big||u|\big|_2 + t\big||v|\big|_2 = r$

Projeções

- Dado um conjunto $S \subset \mathbb{R}^n$ e um ponto $z \in \mathbb{R}^n$, o problema de encontrar o ponto de S mais próximo de z pode não ter solução, ter solução única ou ter soluções múltiplas.



S aberto
Sem solução

S fechado e convexo
Solução única

Múltiplas soluções

- **Lema** Seja $S \subset \mathbb{R}^n$ um conjunto fechado não vazio.
  Dado $z \in \mathbb{R}^n$, existe $\bar{z} \in S$ tal que $\|z - \bar{z}\| \leq \|z - x\|$, para todo $x \in S$.
  - Se o conjunto S for convexo, o lema acima muda para existe um único $\bar{z}$.

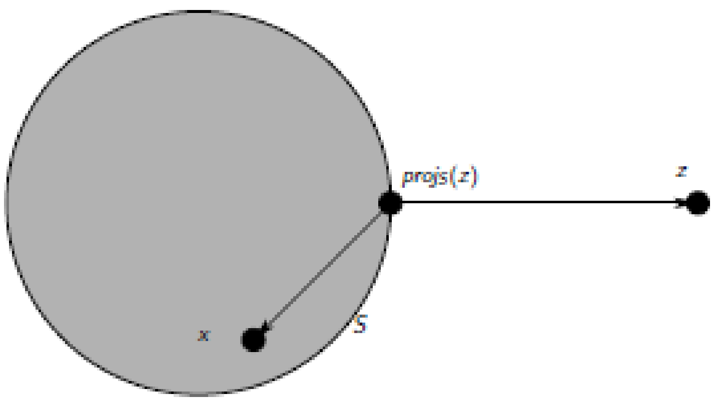
Projeção - caracterização

Teorema

Sejam $S \subset \mathbb{R}^n$ um conjunto não vazio, convexo e fechado, $z \in \mathbb{R}^n$ e $\bar{z} = \text{proj}_S(z)$. Então

$$(z - \bar{z})^T (x - \bar{z}) \leq 0,$$

para todo $x \in S$.



- Lembrando que $A.B = \|A\| \|B\| cos\theta$

- Se um ponto $\bar{z}$ satisfaz o teorema acima, isso é suficiente para caracterizar que $\bar{z}$ é a projeção de z em S.

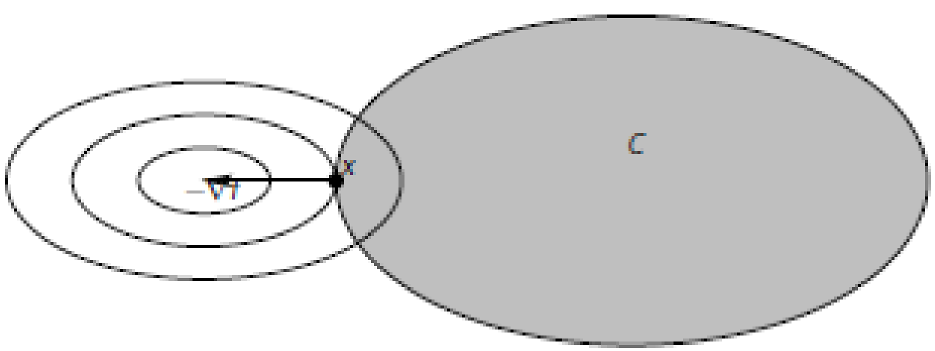
Minimização em conjuntos convexos fechados

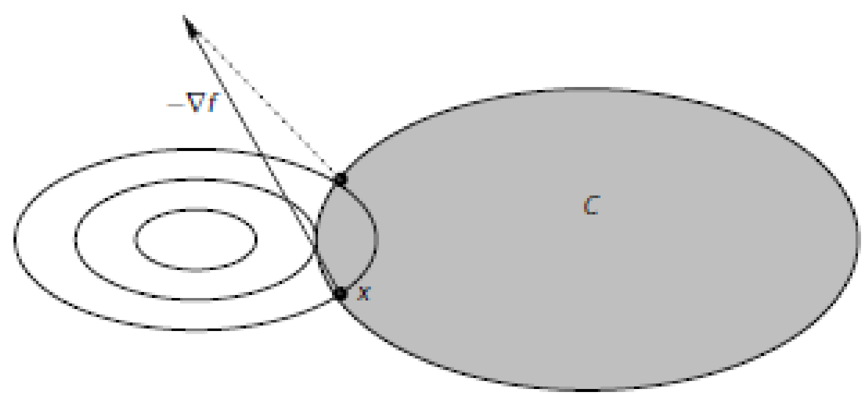
Teorema

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável e $C \subset \mathbb{R}^n$ convexo e fechado. Se $x^* \in C$ é minimizador local de f em C, então

$$\mathrm{proj}_C(x^* - \alpha \nabla f(x^*)) = x^*,$$

para todo $\alpha \geq 0$.





Funções convexas

- As funções convexas possuem ótimas propriedades no contexto de Otimização.

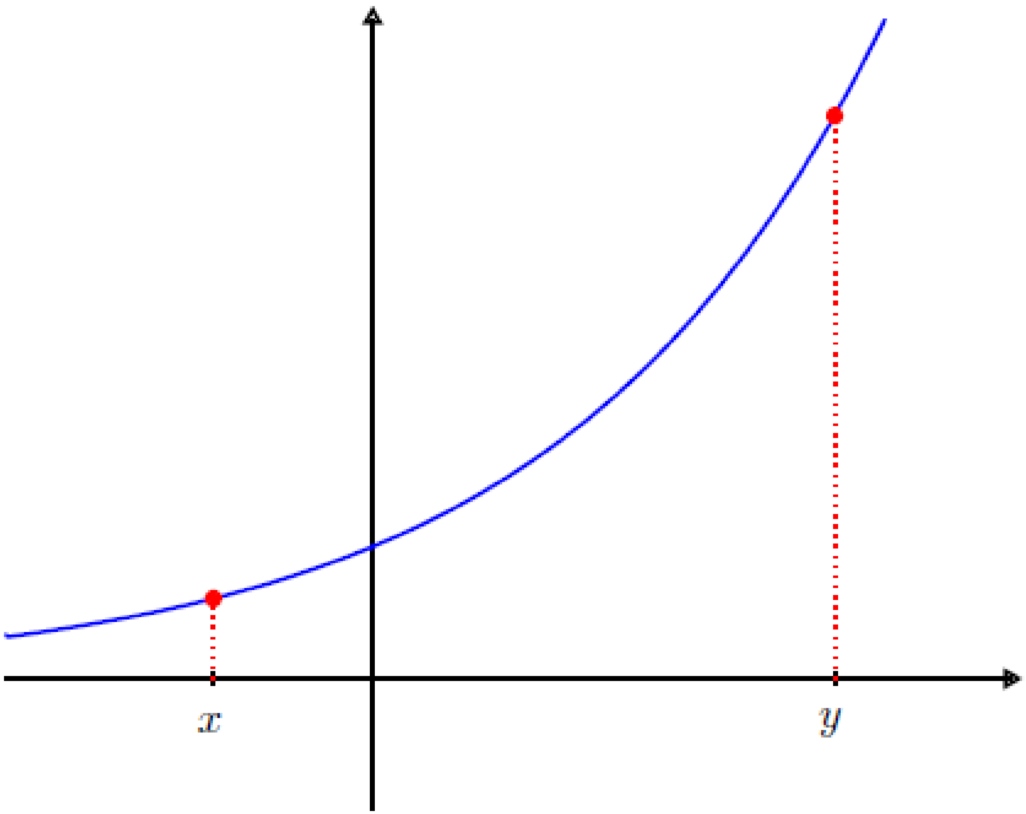
- A presença de convexidade garante, p.ex., que os minimizadores locais são também globais, ou, equivalentemente, que pontos estacionários são minimizadores.

Seja $C \subset \mathbb{R}^n$ um conjunto convexo. Uma função $f : \mathbb{R}^n \to \mathbb{R}$ é convexa em C quando, para todos $x, y \in C$ e $t \in [0,1]$,

$$f\big((1-t)x + ty\big) \leq (1-t)f(x) + tf(y).$$

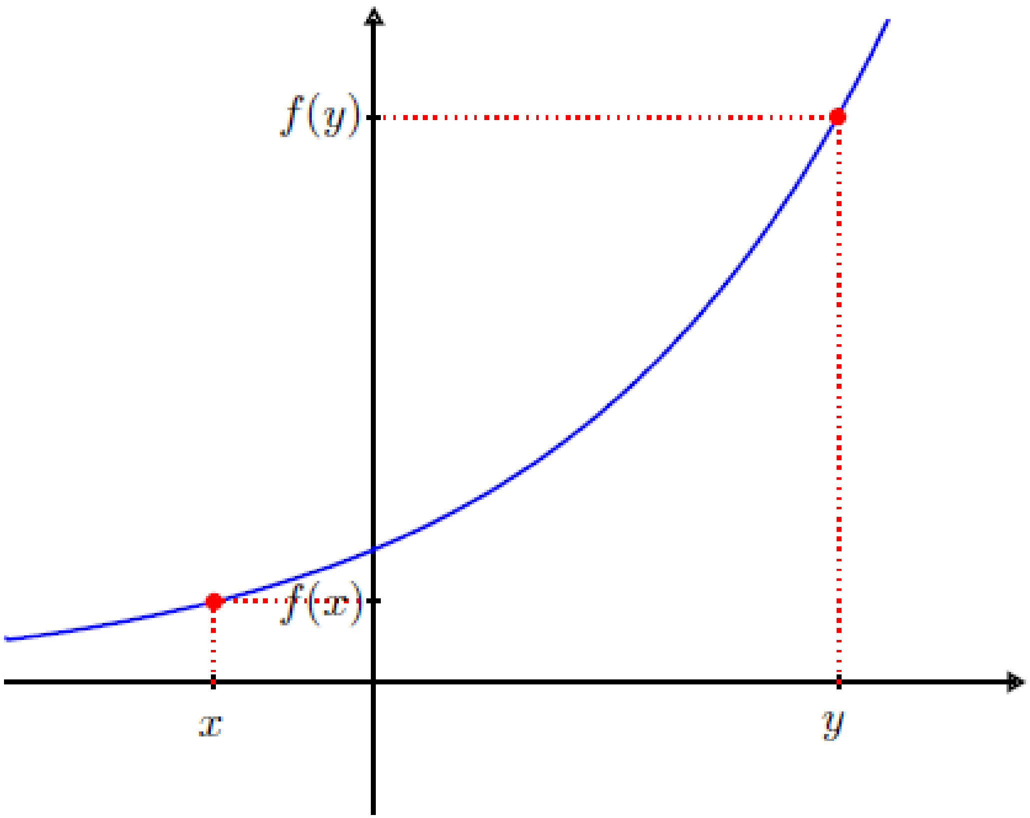
Função convexa

$$f\big((1-t)x + ty\big) \leq (1-t)f(x) + tf(y)$$



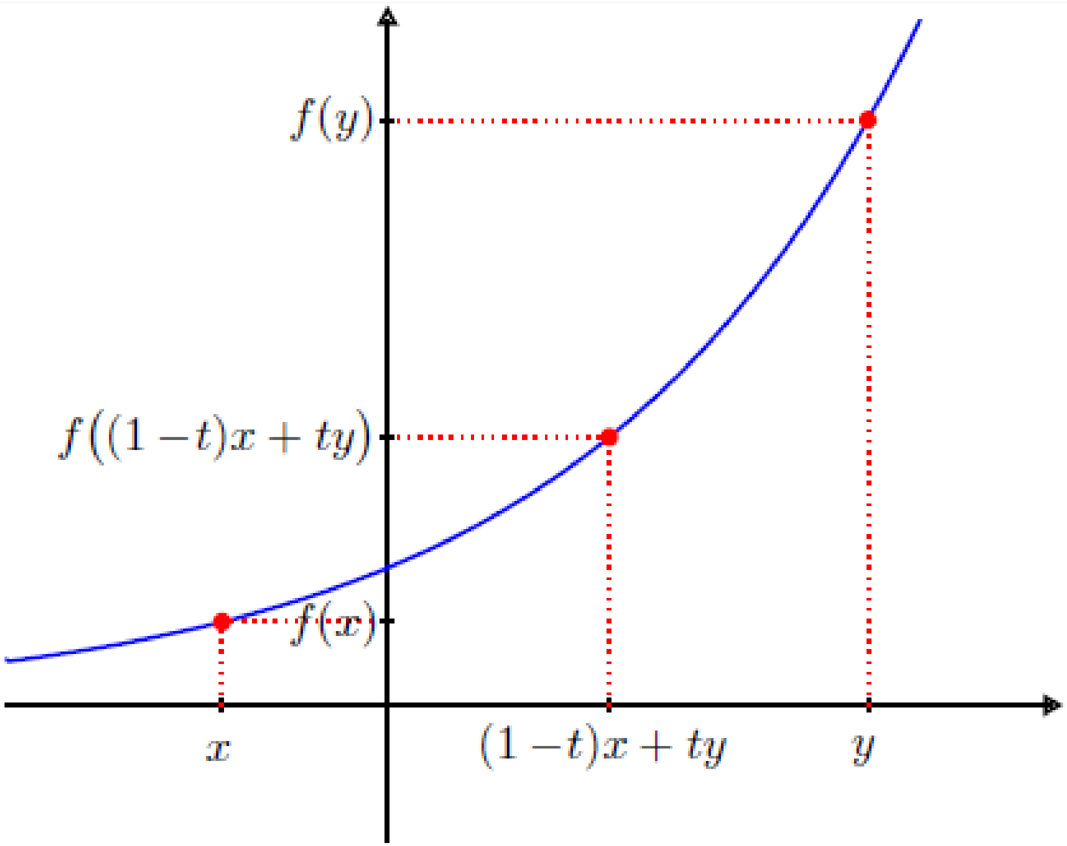
Função convexa

$$f\big((1-t)x+ty\big) \leq (1-t)f(x)+tf(y)$$



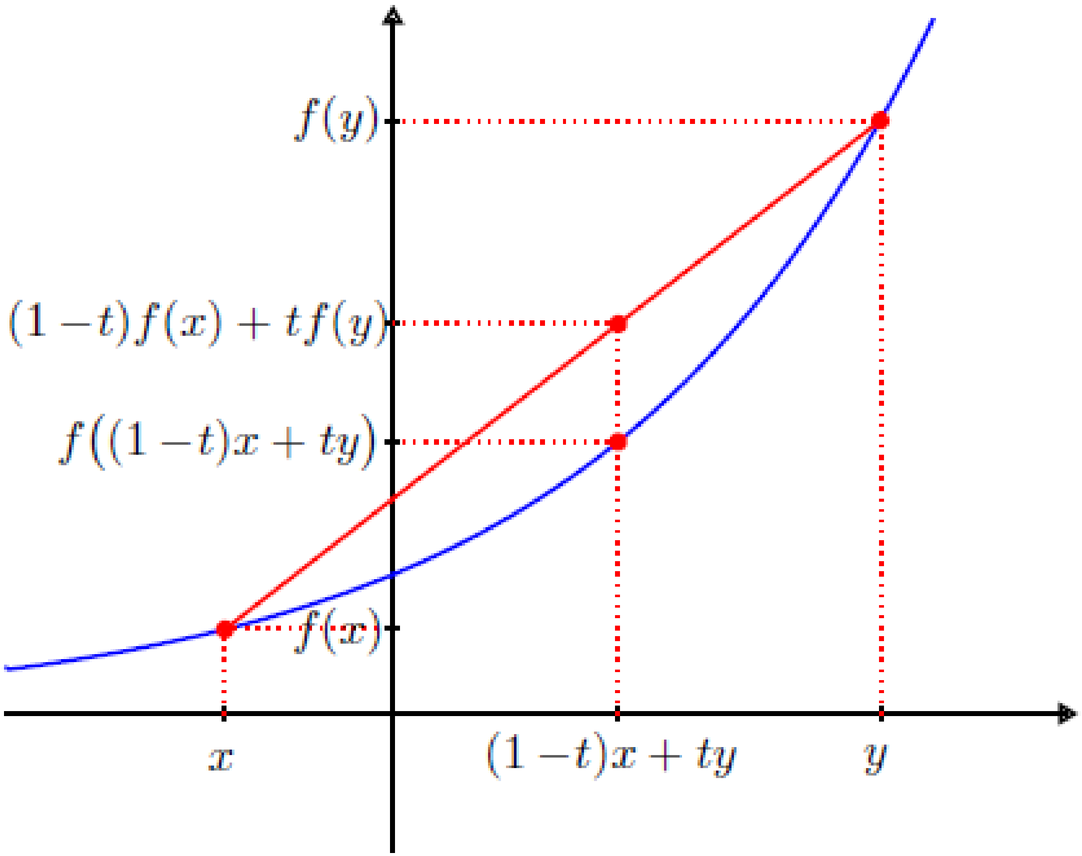
Função convexa

$$f\big((1-t)x+ty\big) \leq (1-t)f(x)+tf(y)$$



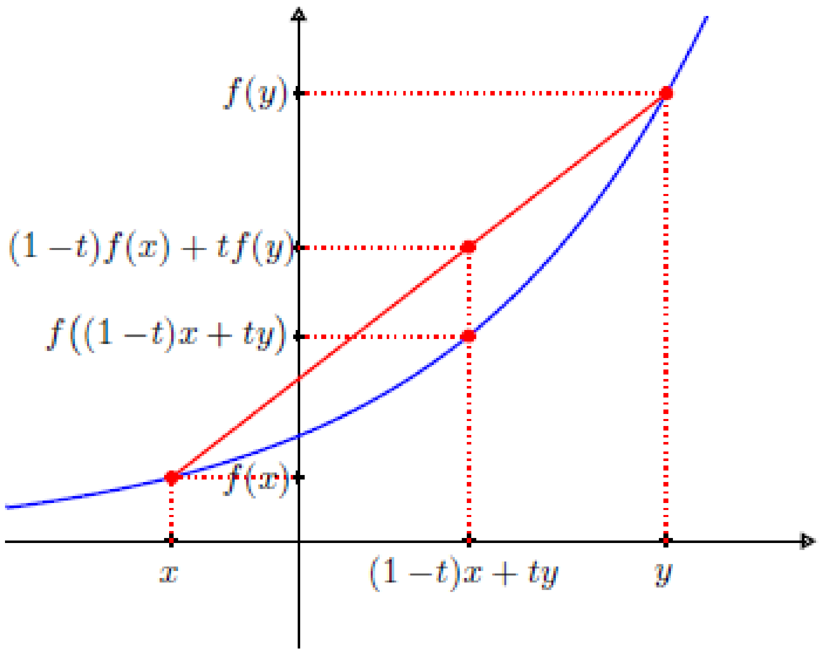
Função convexa

$$f\big((1-t)x+ty\big) \leq (1-t)f(x)+tf(y)$$

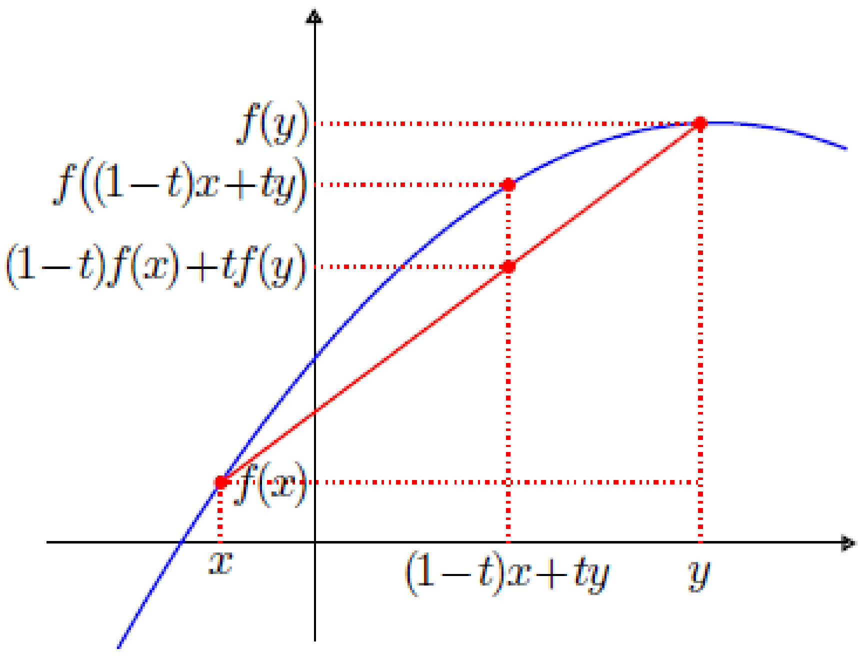


Função convexa × não convexa

$$f\big((1-t)x+ty\big) \le (1-t)f(x)+tf(y)$$



Função convexa

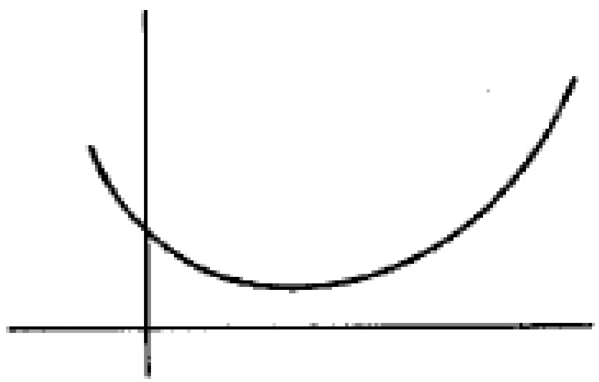


Função não convexa

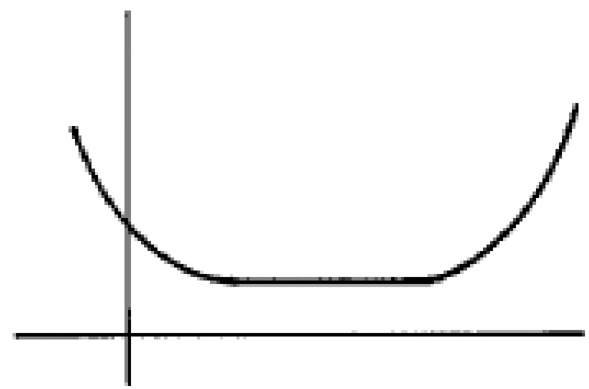
Funções convexa e estritamente convexa

$$f\big((1-t)x+ty\big) < (1-t)f(x)+tf(y)$$

$$f\big((1-t)x+ty\big) \leq (1-t)f(x)+tf(y)$$



ESTRITAMENTE CONVEXA



CONVEXA

Função convexa – exemplos

$$f((1-t)x+ty) \leq (1-t)f(x)+tf(y)$$

Exemplos de funções convexas $f : \mathbb{R} \to \mathbb{R}$

- $f(x) = x^2$.
- $f(x) = e^x$.

Função convexa – x^2

$$f((1-t)x + ty) \le (1-t)f(x) + tf(y)$$

- Mostre que $f(x) = x^2$ é convexa.
- $$\begin{aligned} f((1-t)x + ty) &= f(x + t(y-x)) = (x + t(y-x))^2 \\ &= x^2 + 2tx(y-x) + t^2(y-x)^2. \end{aligned}$$
- Como $t \le 1$,

$$\begin{aligned} x^2 + 2tx(y-x) + t^2(y-x)^2 &\le x^2 + 2tx(y-x) + t(y-x)^2 \\ &= x^2 + t(y^2 - x^2). \end{aligned}$$

- $(1-t)f(x) + tf(y) = (1-t)x^2 + ty^2 = x^2 + t(y^2 - x^2)$.
- Logo $f(x)$ é convexa.

Função convexa – e^x

$$f((1-t)x+ty) \le (1-t)f(x)+tf(y)$$

- Mostre que $g(x) = e^x$ é convexa.

- Sabendo que $e^d \ge 1+d$, para todo $d \in \mathbb{R}$, e fazendo $z = (1-t)x + ty$.
- $e^{x-z} = e^x \frac{1}{e^z} \ge 1 + x - z \;\to\; e^x \ge e^z + e^z(x-z)$.
- Chegamos a resultado similar para $e^{y-z} \;\to\; e^y \ge e^z + e^z(y-z)$.
- $(1-t)e^x + te^y \ge (1-t)(e^z + e^z(x-z)) + t(e^z + e^z(y-z))$
- $= e^z + e^z(x-z) - te^z - te^z(x-z) + te^z + te^z(y-z)$
- $= e^z + e^z(x-z) - te^z(x-z) + te^z(y-z)$
- $= e^z(1 + x - z - tx + ty)$
- $= e^z(1 + x - (1-t)x - ty - tx + ty) = e^z = e^{(1-t)x+ty}$.
- Logo $g(x)$ é convexa.

Função convexa: minimizador local é global

Teorema

Sejam $C \subset \mathbb{R}^n$ convexo e $f : C \to \mathbb{R}$ uma função convexa.
Se $x^* \in C$ é minimizador local de f, então
x^* é minimizador global de f.

- **Prova:**
  - Seja $\delta > 0$ tal que $f(x^*) \leq f(x)$, para todo $x \in C$ onde $\|x - x^*\| \leq \delta$.
  - Dado $y \in C$, $\|y - x^*\| > \delta$, tome $t > 0$ de modo que $t\|y - x^*\| < \delta$.
  - O ponto $x' = (1 - t)x^* + ty$ satisfaz $\|x' - x^*\| = t\|y - x^*\| \leq \delta$.
  - Portanto $x' \in C$ e $\|x' - x^*\| \leq \delta$.
  - Deste modo temos

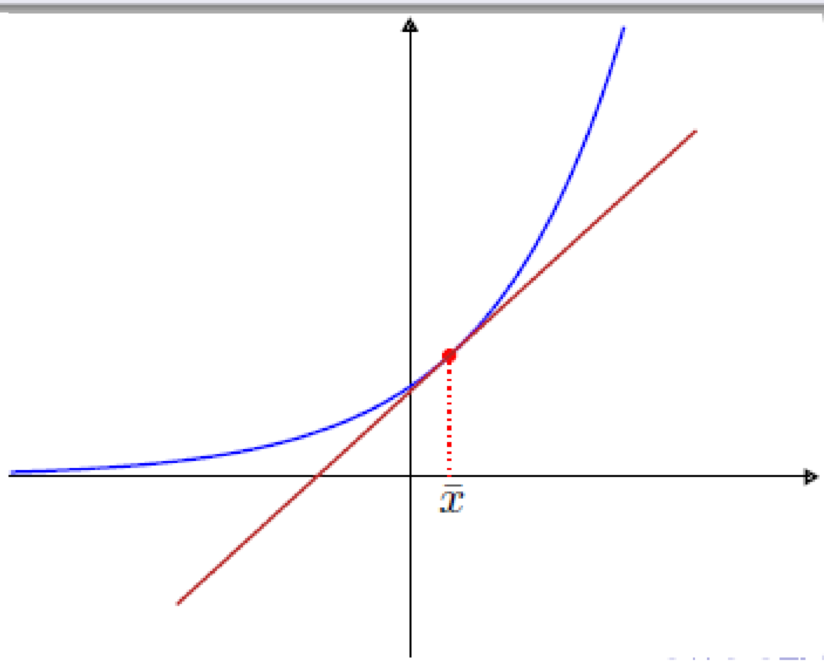
$$f(x^*) \leq f(x') \leq (1 - t)f(x^*) + tf(y) \;\; \to f(x^*) \leq f(y). \blacksquare$$

Função convexa diferenciável está sempre acima de sua aproximação linear

Teorema

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável e $C \subset \mathbb{R}^n$ convexo. A função f é convexa em C se, e somente se, para todos $x, y \in C$,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x).$$



Função convexa diferenciável todo ponto estacionário é minimizador global

Teorema

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável e $C \subset \mathbb{R}^n$ convexo. A função f é convexa em C se, e somente se, para todos $x, y \in C$,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x).$$

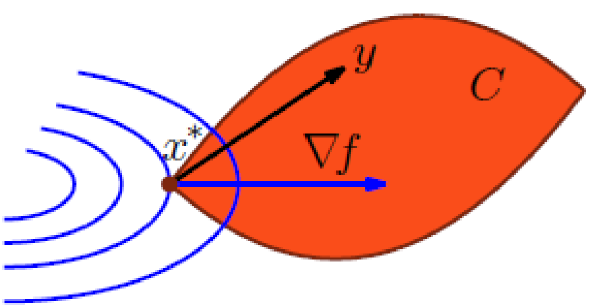
Corolário

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função convexa, diferenciável e $C \subset \mathbb{R}^n$ convexo. Se $\nabla f(x^*)^T (y - x^*) \geq 0$, para todo $y \in C$, então x^* é um minimizador global de f em C. Em particular, todo ponto estacionário é minimizador global.

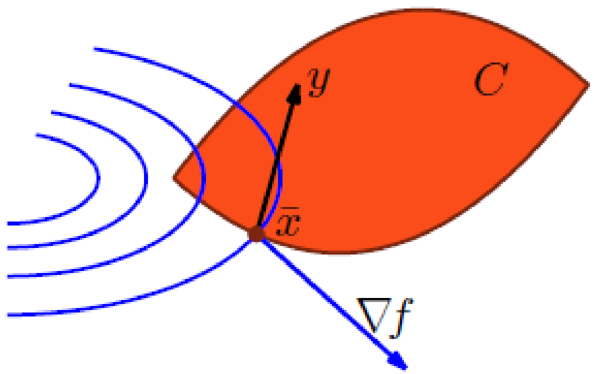
Função convexa diferenciável todo ponto estacionário é minimizador global

Corolário

Sejam $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ uma função convexa, diferenciável e $C \subset \mathbb{R}^n$ convexo. Se $\nabla f(x^*)^T(y - x^*) \geq 0$, para todo $y \in C$, então x^* é um minimizador global de f em C. Em particular, todo ponto estacionário é minimizador global.



Hipóteses satisfeitas



Hipóteses não satisfeitas

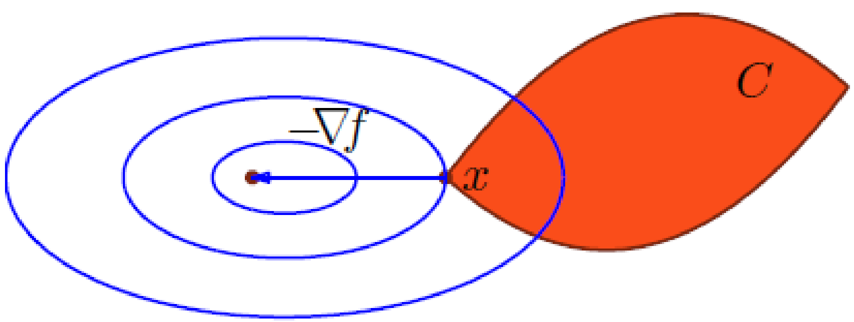
Função convexa diferenciável outra condição suficiente

Teorema

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função convexa diferenciável e $C \subset \mathbb{R}^n$ convexo e fechado. Se

$$\text{proj}_C\left(x^* - \nabla f(x^*)\right) = x^*,$$

então $x^* \in C$ é minimizador global de f em C.



Função convexa diferenciável classe C^2 critério para caracterizar convexidade

Teorema

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função de classe C^2 e $C \subset \mathbb{R}^n$ convexo.

(i) Se $\nabla^2 f(x) \geq 0$, para todo $x \in C$, então f é convexa em C.

(ii) Se f é convexa em C e $\mathrm{int}C \neq \emptyset$, então $\nabla^2 f(x) \geq 0$, para todo $x \in C$.

Relembrando os conceitos de matriz definida positiva e semidefinida positiva

- Seja a matriz $A_{m\times n}$, a matriz $A^T A$ é simétrica e semidefinida positiva, ou seja $x^T A^T A\, x \geq 0, \forall\, x \in \mathbb{R}^n$.

- A matriz $A^T A$ é definida positiva se e somente se $posto(A) = n$.

- Se $n = m$, então a matriz $A^T A$ é definida positiva se e somente se A é não singular.

- Uma matriz quadrada é definida positiva se e somente se todos os seus autovalores são positivos (> 0).

- Uma matriz quadrada é semidefinida positiva se e somente se todos os seus autovalores são não negativos (≥ 0).

- A inversa de uma matriz definida positiva simétrica é definida positiva simétrica.

Sugestão de bibliografia complementar

- Bertsekas, D.P.; Nedic, A.; Ozdaglar, A.E. Convex Analysis and Optimization, Athena Scientific, 2003.
- Hiriart-Urruty J-B.; Lemaréchal, C. Convex Analysis and Minimization Algorithms I, Springer-Verlag, 1993.
- Friedlander, A. Elementos de programação não linear, Unicamp, 1994.