Convexidade

Departamento de Matemática
Universidade do Minho

Outline

Conjuntos Convexos

O conceito de convexidade é fundamental na otimização. Muitos problemas práticos possuem esta propriedade, o que geralmente os torna mais fáceis de resolver tanto na teoria como na prática. O termo "convexo" pode ser aplicado tanto a conjuntos como a funções.

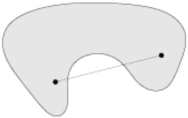
Definição 1 (Conjunto Convexo)

Um conjunto $\mathcal{D} \subseteq \mathbb{R}^d$ é convexo se o segmento de reta entre quaisquer dois pontos de $\mathcal{D}$ está contido em $\mathcal{D}$, i.e., para todos $w, z \in \mathcal{D}$ e para todo $t \in [0, 1]$

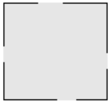
$$t\,w + (1-t)z \in \mathcal{D}$$

Exemplos:

convexo



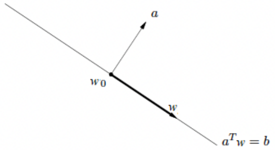
não convexo



não convexo

Exemplos de conjuntos convexos:

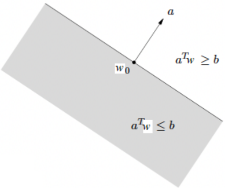
a) Hiperplano: $H = \{w \in \mathbb{R}^d : a^T w = b\}$, com $a \in \mathbb{R}^d \backslash \{0\}$;



Nota:

- O vetor $a \neq 0$ é o vetor normal ao hiperplano, i.e. normal a qualquer vetor contido no hiperplano $a^T(w - w_0) = 0$.
- O hiperplano pode ser interpretado como dividindo o espaço em dois semi-espaços (tal como a reta divide um plano em duas partes).

b) Semi-espaços: $S = \{w \in \mathbb{R}^d : a^T w \leq b\}$, com $a \in \mathbb{R}^d \backslash \{0\}$;



c) Bolas Euclidianas:

$$B(w_c, r) = \{w \in \mathbb{R}^d : \|w - w_c\| \leq r\}, \text{ com } r > 0 \text{ e } w_c \in \mathbb{R}^d;$$

d) Conjunto das matrizes simétricas semi-definidas positivas:

$$S = \{P \in \mathcal{M}_{\{d \times d\}} : w^T P w \geq 0, \forall w \in \mathbb{R}^d\}.$$

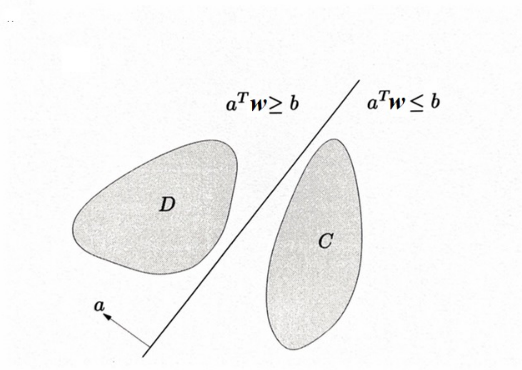
Exercício: Mostre que os hiperplanos, semi-espaços, as bolas euclideanas e o conjunto das matrizes semi-defindas positivas são conjuntos convexos.

Exercício: Seja $S = \{(w_1, w_2, w_3) : w_3 \geq {w_1}^2 + {w_2}^2\} \subset \mathbb{R}^3$. Represente o conjunto S e mostre que este é um conjunto convexo.

Definição 2 (Hiperplano Separador)

Sejam C_1 e C_2 conjuntos não vazios em $\mathbb{R}^d$. O hiperplano $H = \{w : a^T w = b\}$ é um **hiperplano separador** de C_1 e C_2 se $a^T w \geq b$ para todo $w \in C_1$ e $a^T w \leq b$ para todo $w \in C_2$.

Exemplo:

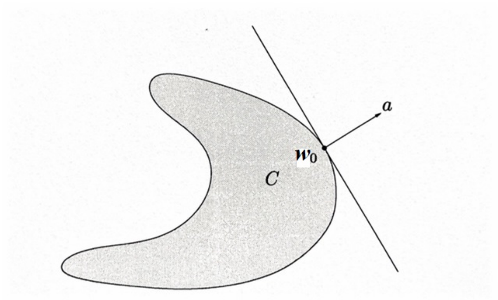


Definição 3 (Hiperplano de Suporte)

Se C é um conjunto não vazio $\mathbb{R}^d$, e seja $w_0 \in \partial C$. Um hiperplano $H = \{w : a^T(w - w_0) = 0\}$ é chamado um **hiperplano de suporte** de C em w_0 se $a^T(w - w_0) \leq 0 \;\; \forall w \in C$.

- ∂C representa a fronteira do conjunto C.

Exemplo:



Proposição 1 (Operações que preservam a convexidade de conjuntos)

i) *Sejam* C_i, $i = 1, \ldots, n$ *conjuntos convexos. Então o conjunto intersecção* $C = \bigcap_{i=1}^{n} C_i$ *é um conjunto convexo.*

ii) *A soma e a subtração de dois conjuntos convexos* C_1 *e* C_2 *é um conjunto convexo:* $C_1 \pm C_2 = \{w \pm z : w \in C_1, z \in C_2\}$

iii) *O conjunto* $\lambda C = \{\lambda w : \lambda \in \mathbb{R},\ w \in C\}$ *é convexo para qualquer conjunto convexo* C *e escalar* λ.

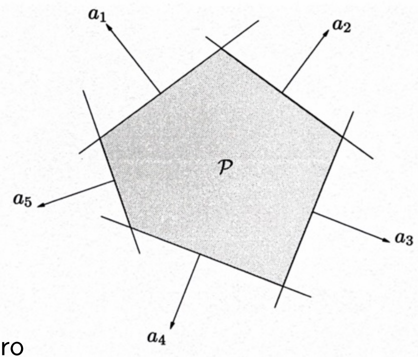
iv) *Sejam* $C_i \subset \mathbb{R}^{d_i}$ $i = 1, \ldots, k$ *conjuntos convexos. Então o produto* $C_1 \times \ldots \times C_k$ *é um conjunto convexo de* $\mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_k}$.

v) *Se* C *é convexo, então também são convexos* $int(C)$ *e* $cl(C)$.

Nota:

- $int(C)$ denota o interior do conjunto C
- $cl(C)$ denota o conjunto fechado de C

Exercício: Prove as alienas de i) a iv) da Proposição 1.



Exercício: O poliedro
$P = \Big\{ w \in \mathbb{R}^d : a_i^T w \leq b_i, \text{com } i = 1, \ldots, n, \; c_j^T w = d_j \text{, com}$
$j = 1, \ldots, p \Big\}, \; a_i, c_j \in \mathbb{R}^d \backslash \{0\}$ é um conjunto convexo? (Justifique.)

Considere agora um conjunto $C \subset \mathbb{R}^d$, um ponto $z \in \mathbb{R}^d$ e o problema de encontrar um ponto de C mais próximo de z. Este problema pode não ter solução e quando tem, não garantimos unicidade. No entanto, como veremos a seguir, se C é fechado, então existe solução. Se além disso, for convexo, a solução é única e será chamada de projeção de z sobre C, denotada por $P_C(z)$.

Teorema 4 (Projeção)

Seja $C \subset \mathbb{R}^d$ um conjunto não vazio, fechado e convexo. Para qualquer $z \in \mathbb{R}^d \setminus C$ tem-se:

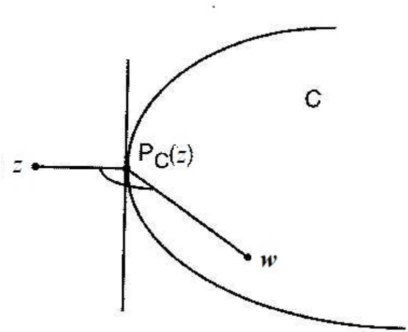
i) *existe um único ponto $w_0 \in C$ tal que*

$$w_0 = P_C(z) = \arg\min_{w \in C} \|z - w\|$$

ii)

$$w_0 = P_C(z) \Leftrightarrow (z - w_0)^T (w - w_0) \leq 0, \forall w \in C.$$

O Teorema 4 é ilustrado na figura abaixo.
(ii) estabelece que para todo vetor $w \in C$, os vetores $z - P_C(z)$ e $w - P_C(z)$ formam um ângulo maior ou igual a $\frac{\pi}{2}$.



Proposição 2 (Separação de um conjunto convexo e um ponto)

Seja $C \subset \mathbb{R}^d$ *um conjunto não vazio, fechado e convexo e* $z \in \mathbb{R}^d \setminus C$. *Então existe um vetor não nulo* a *e um escalar* b *tal que* $a^T z > b$ *e* $a^T w \leq b,\ \forall w \in C$.

Demonstração.

O conjunto C é não vazio, fechado e convexo e $z \notin C$. Logo, pelo Teorema 4, existe um único ponto minimizante $w_0 \in C$ tal que

$$(z - w_0)^T (w - w_0) \leq 0, \forall w \in C.$$

Seja $a = (z - w_0)(\neq 0)$ e $b = a^T w_0$. A inequação anterior pode ser escrita como

$$a^T (w - w_0) \leq 0 \Leftrightarrow a^T w \leq b, \forall w \in C.$$

$\square$

$$
\begin{aligned}
a^T z - b &= (z - w_0)^T z - a^T w_0 \\
&= (z - w_0)^T z - (z - w_0)^T w_0 \\
&= (z - w_0)^T (z - w_0) = \|z - w_0\|^2 \\
&> 0.
\end{aligned}
$$

Logo $a^T z > b$.

$\square$

Proposição 3 (Hiperplano de Suporte)

Seja $C \subset \mathbb{R}^d$ *um conjunto não vazio e convexo e seja* $w_0 \in \partial C$.
Então existe um vetor não nulo a *tal que* $a^T(w - w_0) \leq 0, \forall w \in cl(C)$.

Teorema 5 (Hiperplano Separador)

Sejam $C_1, C_2 \subset \mathbb{R}^d$ dois conjuntos não vazios, fechados e convexos tal que $C_1 \cap C_2 = \emptyset$. Então, existe um vetor a e um escalar b tal que:

$$a^T w_1 \geq b, \forall w_1 \in C_1$$
$$a^T w_2 \leq b, \forall w_2 \in C_2.$$

Demonstração.

Consideremos o conjunto convexo C,

$$C = C_1 - C_2 = \{w \in \mathbb{R} : w = w_1 - w_2, w_1 \in C_1, w_2 \in C_2\}.$$

Uma vez que C_1 e C_2 são disjuntos, a origem não pertence a C. Pela Proposição 2, existe um $a \in \mathbb{R}^d$ e um escalar b tal que

$$a^T z > b$$
$$a^T w \leq b, \forall w \in C.$$

□

Considerando $z = 0$.

Tem-se $b < 0$ (pela 1ª condição) e $a^T w \leq b < 0, \forall w \in C$.
Portanto,

$$a^T w_1 < a^T w_2, \; \forall w_1 \in C_1, \forall w_2 \in C_2.$$

Logo, $\exists b \in \mathbb{R}$ tal que

$$a^T w_1 \leq b \leq a^T w_2, \; \forall w_1 \in C_1, \forall w_2 \in C_2$$

. $\square$

Definição 6 (Função Convexa)

Uma função $F : Dom\ F \to \mathbb{R}$ é convexa se o seu domínio $Dom\ F \subset \mathbb{R}^d$ é um conjunto convexo e se, $\forall w, z \in Dom\ F$,

$$F(tw + (1-t)z) \leq tF(w) + (1-t)F(z), \quad \forall t \in [0,1]. \tag{1}$$

Exemplo:

- Geometricamente, a desigualdade (1) significa que o segmento de reta que une os pontos $(w, F(w))$ e $(z, F(z))$, encontra-se acima ao gráfico de F.
- F é estritamente convexa se a desigualdade em (1) for estrita, sempre que $w \neq z$.
- Diz-se que F é côncava se $-F$ é convexa.

Exemplos de funções convexas em $\mathbb{R}$:

- $F(w) = |w|$
- $F(w) = w^2$
- A função descontínua

$$F(w) = \begin{cases} 1 & w = 0 \\ w^2 & w > 0 \end{cases}$$

Exemplos de funções convexas em $\mathbb{R}^d$:

- todas as normas de $\mathbb{R}^d$
- $F(w) = \max\{w_1, w_2, \ldots, w_d\}$
- função afim: $F(w) = a^T w + b$, $a \in \mathbb{R}^d, b \in \mathbb{R}$ (convexa e côncava)

Exercício: Mostre que as funções acima são convexas.

Nota: Uma norma é uma função que satisfaz as condições:

- $F(\alpha w) = |\alpha| F(w)$, $\forall \alpha \in \mathbb{R}$
- $F(w + z) \leq F(w) + F(z)$
- $F(w) \geq 0, \forall\ w$; $F(w) = 0 \Rightarrow w = 0$.

Proposição 4

Seja C um conjunto não vazio, fechado e convexo em $\mathbb{R}^d$. A função distância $d : \mathbb{R}^d \to \mathbb{R}$, definida por

$$d_C(z) = \inf\{\|z - c\| : c \in C\}$$

é convexa.

Demonstração.

Vamos provar por contradição, vamos assumir que existe $w_1, w_2 \in \mathbb{R}^d$ e um $\alpha \in [0, 1]$ tal que

$$d(tw_1 + (1-t)w_2, C) > td(w_1, C) + (1-t)d(w_2, C).$$

Então, deve existir $z_1, z_2 \in C$ tal que

$$d(tw_1 + (1-t)w_2, C) > t\|z_1 - w_1\| + (1-t)\|z_2 - w_2\|,$$

$\square$

que implica

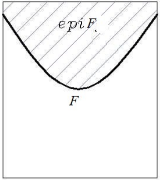
$$\|tz_1 + (1-t)z_2 - tw_1 - (1-t)w_2\| > t\|w_1 - z_1\| + (1-t)\|w_2 - z_2\|.$$

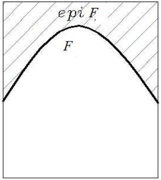
Contradiz a desigualdade triangular na definição de norma. $\square$

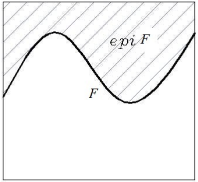
Definição 7

Seja *Dom F* um conjunto não vazio $\mathbb{R}^d$, e seja $F : Dom\ F \longrightarrow \mathbb{R}$. O **epigrafo** de F, denota-se por epi F, é o subconjunto de $\mathbb{R}^{d+1}$ definido por

$$\{(w,z) : w \in Dom\ F, z \in \mathbb{R}, z \geq F(w)\}$$







Proposição 5

Seja Dom F um conjunto não vazio e convexo em $\mathbb{R}^d$, e seja $F : \text{Dom } F \longrightarrow \mathbb{R}$. Então, F é convexa se e só se epi F é um conjunto convexo.

Teorema 8

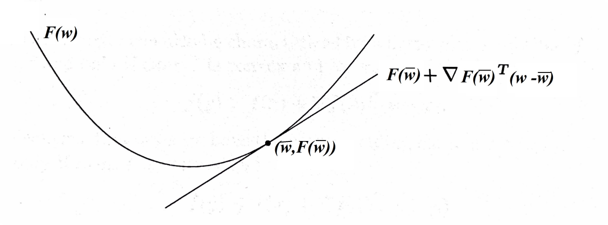
Uma função $F : \mathbb{R}^d \longrightarrow \mathbb{R}$ é convexa se e só se para todo $w \in \text{Dom } F$ e para todo $s \in \mathbb{R}^d$, a função $g(\alpha) = F(w + \alpha s)$ é convexa (no seu domínio $\{\alpha : w + \alpha s \in \text{Dom } F\}$).

Nota: Esta propriedade é útil porque nos permite reduzir o problema de verificação da convexidade de uma multivariada função para verificar a convexidade de uma função univariada, para a qual podemos usar critérios muito mais simples.

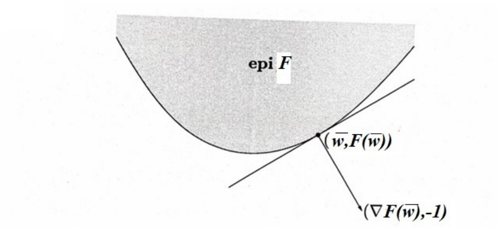
Proposição 6

Seja Dom F um subconjunto $\mathbb{R}^d$ convexo e seja $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ uma função diferenciável em Dom F. A função F é convexa em Dom F se e só se $\forall w, \bar{w} \in$ Dom F

$$F(w) \geq F(\bar{w}) + \nabla F(\bar{w})^T (w - \bar{w}). \tag{2}$$



Nota: Uma consequência desta proposição é a seguinte: Se F é uma função diferenciável, convexa e $\nabla F(\bar{w}) = 0$, então $\bar{w}$ minimiza F em $\mathbb{R}^d$.



Nota:

- $F(w) \geq F(\bar{w}) + \nabla F(\bar{w})^T (w - \bar{w})$ se e só se

$$(\nabla F(\bar{w}), -1)^T \left[(w, F(w)) - (\bar{w}, F(\bar{w}))\right] \leq 0, \forall w \in Dom\ F.$$

Logo, $H(w) = F(\bar{w}) + \nabla F(\bar{w})^T (w - \bar{w})$ é um hiperplano de suporte de epi F em $\bar{w}$.

Teorema 9

Seja Dom F um conjunto não vazio e convexo em $\mathbb{R}^d$ e seja $F : \mathbb{R}^d \longrightarrow \mathbb{R}$ uma função duas vezes continuamente diferenciável sobre um conjunto aberto que contém Dom F.

a) *Se $\nabla^2 F(w)$ é semi-definida positiva em todos os pontos do Dom F, então F é convexa em Dom F.*

b) *Se $\nabla^2 F(w)$ é definida positiva em todos os pontos do Dom F, então F é estritamente convexa em Dom F.*

c) *Se Dom F é aberto e F é convexa em Dom F, então $\nabla^2 F(w)$ é semi-definida positiva em todos os pontos do Dom F.*

Recordar: Seja $A \in \mathbb{R}^{d \times d}$ uma matriz simétrica.

- A é definida positiva sse $s^T A s > 0$ para todo $s \in \mathbb{R}^d, \; s \neq 0$
- A é semi-definida positiva sse $s^T A s \geq 0$ para todo $s \in \mathbb{R}^d, \; s \neq 0$.

Nota:

- Similarmente F é concâva se e só se $\nabla^2 F$ é semi-definida negativa todos os pontos do $Dom\, F$, em que $Dom\, F$ um conjunto não vazio, aberto e convexo.

- Se e $\nabla^2 F$ é definida positiva para todo $w \in Dom\, F$, então F é estritamente convexa. Contudo, o contrário não é verdade. Por exemplo a função $F : \mathbb{R} \to \mathbb{R}$ dada por $F(w) = w^4$ é estritamente convexa, mas tem segunda derivada igual a zero em $w = 0$

Demonstração.

a) Suponhamos que $w^T \nabla^2 F(\bar{w}) w \geq 0,\ \forall w \in \mathbb{R}^d$. Considere-se $w, \bar{w} \in Dom\ F$. Pelo Teorema do Valor Médio, temos

$$F(w) = F(\bar{w}) + \nabla F(\bar{w})^T (w - \bar{w}) + \frac{1}{2}(w - \bar{w})^T \nabla^2 F(\hat{w})(w - \bar{w}),$$

onde $\hat{w} = \lambda \bar{w} + (1 - \lambda) w \in Dom\ F$.
Portanto

$$F(w) \geq F(\bar{w}) + \nabla F(\bar{w})^T (w - \bar{w}),$$

que implica a convexidade.

b) É semelhante à prova do item [a)], obtemos $F(w) > F(\bar{w}) + \nabla F(\bar{w})^T (w - \bar{w})$, para todo $w, \bar{w} \in Dom\ F$ com $w \neq \bar{w}$.

$\square$

c) Assumindo que F é convexa e seja $\bar{w} \in Dom\ F$, queremos provar que $w^T \nabla^2 f(\bar{w}) w \geq 0, \forall w \in \mathbb{R}^d$. Uma vez que $Dom\ F$ é um conjunto aberto, para qualquer $w \in \mathbb{R}^d$, $\bar{w} + \lambda w \in Dom\ F$ for $|\lambda| \neq 0$ e suficientemente pequeno

$$F(\bar{w} + \lambda w) \geq F(\bar{w}) + \lambda \nabla F(\bar{w})^T w.$$

Pela expansão de Taylor, temos

$$\begin{aligned} F(\bar{w} + \lambda w) = \quad & F(\bar{w}) + \lambda \nabla F(\bar{w})^T w + \tfrac{1}{2}\lambda^2 w^T \nabla^2 F(\bar{w}) w + \\ & + \lambda^2 \|w\|^2 \theta(\bar{w}, \lambda w). \end{aligned}$$

Logo,

$$\frac{1}{2}\lambda^2 w^T \nabla^2 F(\bar{w}) w + \lambda^2 \|w\|^2 \theta(\bar{w}, \lambda w) \geq 0.$$

Dividindo por $\lambda^2 > 0$ e assumindo que $\lambda \to 0$, temos que

$$w^T \nabla^2 F(\bar{w}) w \geq 0$$

Exercício: Verifique que a função $F(w) = w \ln(w)$ é convexa em $\mathbb{R}^+$, utilizando o Teorema 9.

Resolução: Atendendo que estamos perante uma função $F : \mathbb{R}^+ \to \mathbb{R}$, basta verificar que a 2ª derivada é não negativa.

$$F'(w) = \ln(w) + 1, F''(w) = 1/w.$$

Como $F''(w) > 0$ para $w > 0$, a função é (estritamente) convexa.

Exercício: Utilize o Teorema 9 para provar que as seguintes funções são convexas:

a) $f(w) = e^{aw}$ com $w \in \mathbb{R}$; $a \in \mathbb{R}$
b) $f(w) = w^a$ com $w \in \mathbb{R}^+$; $a \geq 1$ ou $a \leq 0$
c) $f(w) = -w^a$ com $w \in \mathbb{R}^+$; $0 \leq a \leq 1$
d) $f(w) = |w|^a$; com $w \in \mathbb{R}$; $a \geq 1$

Exercício: Verifique que a função $F : \mathbb{R}^2 \to \mathbb{R}$

a) $F(w,z) = \frac{w^2}{z}$ é convexa para $z > 0$.

b) $F(w,z) = \ln(e^w + e^z)$ é convexa.

Exercício: Verifique que a seguinte função quadrática $F : \mathbb{R}^d \to \mathbb{R}$, $F(w) = w^T Q w + a^T w + b$ com Q uma matriz simétrica $d \times d$ é:

a) convexa se e só se Q é semi-definida positiva;

b) estritamente convexa se e só se Q é definida positiva.

Operações de preservação de convexidade de funções

Proposição 7 (Soma ponderada não negativa)

Seja $F_1, \ldots, F_m : \mathbb{R}^d \longrightarrow \mathbb{R}$ *funções dadas, seja* $\lambda_1, \ldots, \lambda_m$ *escalares positivos, e considere a função* $G : \mathbb{R}^d \longrightarrow \mathbb{R}$ *dada por*

$$G(w) = \lambda_1 F_1(w) + \ldots + \lambda_m F_m(w).$$

Se $F_1, \ldots, F_m$ *são convexas, então* G *também é convexa,*

Demonstração.

Seja $F_1, \ldots, F_m$ funções convexas. Pela definição de convexidade, tem-se que para qualquer $w, z \in \mathbb{R}^d$ e $\alpha \in [0,1]$.

$$\begin{aligned}
G(tw + (1-t)z) &= \sum_{i=1}^{m} \lambda_i F_i(tw + (1-t)z) \\
&\leq \sum_{i=1}^{m} \lambda_i \left(tF_i(w) + (1-t)F_i(z) \right) \\
&= t \sum_{i=1}^{m} \lambda_i F_i(w) + (1-t) \sum_{i=1}^{m} \lambda_i F_i(z) \\
&= tG(w) + (1-t)G(z).
\end{aligned}$$

Logo, G é convexo. □

Proposição 8 (Composição com uma função afim)

Seja $F : \mathbb{R}^d \longrightarrow \mathbb{R}$ *uma função dada, seja* A *uma matriz* $n \times d$, $b \in \mathbb{R}^n$ *e considere a função* $G : \mathbb{R}^d \longrightarrow \mathbb{R}$ *dada por*

$$G(w) = F(Aw + b).$$

Se F *é convexo, então* G *também é convexo.*

Exercício: Prove a Proposição 8.

Proposição 9 (Supremo Pontual)

Seja $F_i : \mathbb{R}^d \longrightarrow \mathbb{R}$ *uma função dada para* $i \in I$, *onde* I *é um arbitrário conjunto de indíces e considere a função* $G : \mathbb{R}^d \longrightarrow \mathbb{R}$ *dada por*

$$G(w) = \sup_{i \in I} F_i(w).$$

Se $F_i, i \in I$, *é convexo.*

Demonstração.

Considere

$$\text{epi } G = \{(w, z) : G(w) \leq z\}.$$

Um par $(w, z) \in \text{epi } G$ se e só se $F_i(w) \leq z$ para todo $i \in I$, que é equivalente a dizer que $(w, z) \in \cap_{i \in I} \text{epi } F_i$. Portanto,

$$\text{epi } G = \cap_{i \in I} \text{epi } F_i.$$

Se F_i são funções todas convexas, os epigrafes epi F_i são convexos, portanto epi G é convexo G é convexo pela Proposição 5.

Resumo: diferentes modos para provar que uma função é convexa

- Usar a definição de função convexa.
- Mostrar que o epigrafo de função é convexo.
- 0ª ordem: verificar que $\forall w \in DomF,\ \forall s \in \mathbb{R}^d$, $g(\alpha) = F(w + \alpha s)$ é convexa (no seu domínio $\{\alpha : w + \alpha s \in Dom\ F\}$).
- 1º ordem: verificar que $\forall w, \bar{w} \in Dom\ F,\ F(w) \geq F(\bar{w}) + \nabla F(\bar{w})^T (w - \bar{w})$.
- 2º ordem: verificar que $\nabla^2 F(w)$ é semi-definida positiva em todos os pontos do $Dom\ F$.
- Construindo a partir de funções convexas usando operações de preservação de convexidade.

Os apontamentos foram baseados na seguinte bibliografia: [1], [2], [3] e [4]

A. S. Bazaraa, H. Sherali, and C. Shetty.
Nonlinear Programming.
John Wiley and Sons, Inc., 1993.

S. Boyd and L. Vandenberghe.
Convex Optimization.
Cambridge University Press, 2004.

R. T. Rockafellar.
Convex Analysis.
Princeton University Press, Princeton, New Jersey, 1970.

D. P. B. with Angelia Nedic and A. E. Ozdaglar.
Convex Analysis and Optimization.
Athena Scientific, 2003.