Resolução da Ficha - Convexidade

1. Mostre que os seguintes conjuntos são convexos.

   (a) Subespaços de $\mathbb{R}^d$:

   Resolução: Seja D um subespaço de $\mathbb{R}^d$. Sejam quaisquer $u, v \in D$ e $t \in [0,1]$.
   Queremos mostrar que $tu + (1-t)v \in D$.
   Como D é um subespaço de $\mathbb{R}^d$, D é fechado para adição e multiplicação por escalares, temos $tu \in D$ e $(1-t)v \in D$. Portanto
   $$tu + (1-t)v \in D,$$
   donde se conclui que D é convexo.

   (b) Conjuntos unitários $S = \{w_0 \in \mathbb{R}^d\}$.

   Resolução: Sejam quaisquer $u, v \in S$ e $t \in [0,1]$.
   Queremos mostrar que $tu + (1-t)v \in S$.
   Como S contém apenas o ponto w_0, temos $u = v = w_0$. A combinação convexa destes pontos é
   $$tu + (1-t)v = tw_0 + (1-t)w_0 = tw_0 + 1w_0 - tw_0 = w_0 \in S$$
   donde se conclui que S é convexo.

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

   Resolução: Sejam quaisquer $u, v \in H$ e $t \in [0,1]$.
   Queremos mostrar que $tu + (1-t)v \in S$.
   Por definição de H, temos
   $$a^T u = b \quad \text{e} \quad a^T v = b.$$
   Logo,
   $$a^T\big(tu + (1-t)v\big) = t\,a^T u + (1-t)\,a^T v = tb + (1-t)b = tb + b - tb = b.$$
   Portanto $tu + (1-t)v \in H$, donde se conclui que H é convexo.

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

   Resolução: Sejam quaisquer $u, z \in S$ e $t \in [0,1]$.
   Queremos mostrar que $v := tu + (1-t)z \in S$, ou seja, $a^T v \leq b$.
   Por definição de S, temos
   $$a^T u \leq b, \qquad a^T z \leq b.$$
   Assim,
   $$a^T v = a^T\big(tu + (1-t)z\big) = t\,a^T u + (1-t)\,a^T z \leq tb + (1-t)b = b.$$
   Portanto $tu + (1-t)z \in S$, donde se conclui que S é convexo.

   (e) Bolas euclideanas: $B(w_c, r) = \{w \in \mathbb{R}^d : \|w - w_c\|_2 \leq r\}$ (com $w_c \in \mathbb{R}^d, r \in \mathbb{R}^+$)

   Resolução: Sejam quaisquer $u, z \in B(w_c, r)$ e $t \in [0,1]$.
   Queremos mostrar que $v = tu + (1-t)z \in S$, ou seja, $\|v - w_c\|_2 \leq r$.
   Por definição de $B(w_c, r)$, temos
   $$\|u - w_c\|_2 \leq r, \qquad \|z - w_c\|_2 \leq r.$$
   Assim, pela desigualdade triangular e pela homogeneidade da norma, temos:
   $$\begin{aligned}\big\|\,tu + (1-t)z - w_c\big\|_2 &= \big\|tu + (1-t)z - w_c - tw_c + tw_c\big\|_2 \\ &= \big\|tu - tw_c + (1-t)z - (1-t)w_c\big\|_2 \\ &= \big\|t(u - w_c) + (1-t)(z - w_c)\big\|_2 \\ &\leq \big\|t(u - w_c)\big\|_2 + \big\|(1-t)(z - w_c)\big\|_2 \\ &= t\|u - w_c\|_2 + (1-t)\|z - w_c\|_2 \\ &\leq tr + (1-t)r = r.\end{aligned}$$
   Portanto $v = tu + (1-t)z \in B(w_c, r)$, donde se conclui que $B(w_c, r)$ é convexo.

(f) Conjunto das matrizes semi-defindas positivas: $S = \{P \in \mathcal{M}_{\{n\times n\}} : s^T P s \geq 0, \forall s \in \mathbb{R}^n\}$.
Resolução: Sejam quaisquer $P, Q \in S$ e $t \in [0,1]$. Queremos mostrar que

$$R := tP + (1-t)Q \in S,$$

ou seja, que satisfaz a condição $s^T R s \geq 0$ para todo $s \in \mathbb{R}^n$.
Assim, para qualquer $s \in \mathbb{R}^n$ temos,

$$s^T R s = s^T \big(tP + (1-t)Q\big) s = t\, s^T P s + (1-t)\, s^T Q s.$$

Como $P, Q \in S$ são semidefinidas positivas, $s^T P s \geq 0$ e $s^T Q s \geq 0$. Como $t \geq 0$ e $1-t \geq 0$ temos

$$s^T R s \geq 0.$$

Portanto $R \in S$, donde se conclui que S é convexo.

2. Seja $D = \{(w_1, w_2) \in \mathbb{R}^2 : w_2 \geq 0\}$.

Resolução: Sejam quaisquer $u, v \in D$ e $t \in [0,1]$.
Queremos mostrar que $z := tu + (1-t)v \in D$, ou seja, $z_2 \geq 0$.
Como $u = (u_1, u_2) \in D$, $v = (v_1, v_2) \in D$ então

$$u_2 \geq 0, \qquad v_2 \geq 0.$$

Assim, temos

$$\begin{aligned} z = tu + (1-t)v &= t(u_1,u_2) + (1-t)(v_1,v_2) \\ &= (\underbrace{tu_1 + (1-t)v_1}_{z_1}, \underbrace{tu_2 + (1-t)v_2}_{z_2}) \end{aligned}$$

Pela hipótese sobre u_2, v_2,

$$z_2 = tu_2 + (1-t)v_2 \geq t \times 0 + (1-t) \times 0, \quad \text{ou seja,} \quad z_2 \geq 0,$$

o que implica $z \in D$. Portanto o conjunto D é convexo.

Geometricamente: D é um semi-subespaço e consiste em todos os pontos (w_1, w_2) situados *acima* (ou sobre) da reta $w_2 \geq 0$.

3. Considere os conjuntos $D_1 = \{(w_1, w_2) \in \mathbb{R}^2 : w_2 \geq 0\}$ e $D_2 = \{(w_1, w_2) \in \mathbb{R}^2 : w_1 \geq w_2\}$. Mostre que $D = D_1 \cap D_2$ é um conjunto convexo e represente graficamente o conjunto D.

Resolução: Já mostramos no exercício anterior que D_1 é convexo. Como a interseção de conjuntos convexos é um conjunto convexo, basta então mostrar que D_2 é convexo.

Sejam quaisquer $u, v \in D_2$ e $t \in [0,1]$.
Queremos mostrar que $z := tu + (1-t)v \in D_2$, ou seja, $z_1 \geq z_2$.
Como $u = (u_1, u_2) \in D_2$, $v = (v_1, v_2) \in D_2$ então

$$u_1 \geq u_2, \qquad v_1 \geq v_2.$$

Assim, temos

$$\begin{aligned} z = tu + (1-t)v &= t(u_1,u_2) + (1-t)(v_1,v_2) \\ &= (\underbrace{tu_1 + (1-t)v_1}_{z_1}, \underbrace{tu_2 + (1-t)v_2}_{z_2}) \end{aligned}$$

Pela hipótese sobre u_1, v_1,

$$z_1 = tu_1 + (1-t)v_1 \geq tu_2 + (1-t)v_2, \quad \text{ou seja,} \quad z_1 \geq z_2,$$

o que implica $z \in D_2$. Portanto o conjunto D_2 é convexo.

Geometricamente: $D = D_1 \cap D_2 = \{(w_1, w_2) \in \mathbb{R}^2 : w_2 \geq 0, w_1 \geq w_2\}$ consiste em todos os pontos (w_1, w_2) situados *acima* da reta $w_2 \geq 0$ e abaixo da reta $w_2 \leq w_1$.

4. Mostre, usando a definição que as que seguintes funções $f : \mathbb{R} \to \mathbb{R}$ são convexas.

   (a) $f(w) = w^2$
   Resolução: Sejam quaisquer $u, z \in \mathbb{R}$ e $t \in [0, 1]$. Queremos mostrar que

   $$f(tu + (1-t)z) \leq tf(u) + (1-t)f(z).$$

   Temos:
   $$f(tu + (1-t)z) = (tu + (1-t)z)^2 = t^2u^2 + (1-t)^2z^2 + 2t(1-t)uz.$$

   Por outro lado, temos:
   $$tf(u) + (1-t)f(z) = tu^2 + (1-t)z^2.$$

   Subtraindo os dois lados, obtemos:

   $$\begin{aligned}
   &tf(u) + (1-t)f(z) - f(tu + (1-t)z) \\
   &= tu^2 + (1-t)z^2 - (t^2u^2 + (1-t)^2z^2 + 2t(1-t)uz) \\
   &= t(1-t)u^2 + (1-t)(1-(1-t)z^2 - 2t(1-t)uz) \\
   &= t(1-t)u^2 + (1-t)(1-1+t)z^2 - 2t(1-t)uz) \\
   &= t(1-t)(u-z)^2 \geq 0
   \end{aligned}$$

   Portanto, temos que
   $$tf(u) + (1-t)f(z) - f(tu + (1-t)z) \geq 0$$

   ou seja,
   $$f(tu + (1-t)z) \leq tf(u) + (1-t)f(z),$$

   o que prova que $f(w) = w^2$ é convexa em $\mathbb{R}$.

   (b) $f(w) = |w|$
   Resolução: Sejam quaisquer $u, z \in \mathbb{R}$ e $t \in [0, 1]$. Queremos mostrar que

   $$f(tu + (1-t)z) \leq tf(u) + (1-t)f(z).$$

   Usando a desigualdade triangular para o valor absoluto e a homogeneidade do módulo para escalares não-negativos, obtemos

   $$\big|tu + (1-t)z\big| \leq \big|tu\big| + \big|(1-t)z\big| = t|u| + (1-t)|z|$$

   ou seja,
   $$f(tu + (1-t)z) \leq tf(u) + (1-t)f(z).$$

   Portanto f é convexa em $\mathbb{R}$.

   (c) $f(w) = \begin{cases} 1, & w = 0, \\ w^2, & w > 0. \end{cases}$

   Resolução: Há três casos:
   (i) Sejam quaisquer $u > 0$ e $z > 0$. Neste caso $f(u) = u^2$, $f(z) = z^2$ e $f\big(tu + (1-t)z\big) = (tu + (1-t)z)^2$. Já vimos no exercício (a) que $f(w) = w^2$ é convexa,

   $$(tu + (1-t)z)^2 \leq tu^2 + (1-t)z^2,$$

logo a desigualdade de convexidade é satisfeita.
(ii) $u = 0$ **e** $z = 0$**.** Então $f(u) = f(z) = 1$ e para todo $t \in [0, 1]$

$$f\big(t \cdot 0 + (1-t) \cdot 0\big) = f(0) = 1 = tf(0) + (1-t)f(0) = t \cdot 1 + (1-t) \cdot 1,$$

logo a desigualdade é verdadeira (igualdade).
(iii) Um dos pontos é 0 e o outro > 0. Sem perda de generalidade suponhamos $u = 0$ e $z > 0$.

Temos:

$$f\big(tu + (1-t)z\big) = f\big((1-t)z\big) = ((1-t)z)^2.$$

Por outro lado, temos

$$tf(u) + (1-t)f(z) = t \cdot 1 + (1-t)z^2.$$

Calculamos a diferença

$$\begin{aligned} & tf(u) + (1-t)f(z) - f\big(tu + (1-t)z\big) \\ = \ & t + (1-t)z^2 - (1-t)^2 z^2 \\ = \ & t + t(1-t)z^2 \\ = \ & t\big(1 + (1-t)z^2\big) \geq 0, \end{aligned}$$

pois $t \geq 0$ e $1 + (1-t)z^2 > 0$. Portanto, segue que

$$tf(u) + (1-t)f(z) - f(tu + (1-t)z) \geq 0.$$

ou seja,

$$f(tu + (1-t)z) \leq tf(u) + (1-t)f(z).$$

Portanto f é convexa.
Como em todos os casos a desigualdade de convexidade se verifica, concluímos que f é convexa em $[0, \infty)$.

5. Mostre, usando o Teorema da Função convexa que as que seguintes funções $f : \mathbb{R} \to \mathbb{R}$ são convexas.

   (a) $f(w) = aw + b; \quad a, b \in \mathbb{R};\ a \neq 0$
   Resolução: Queremos mostrar:

   $$f''(w) \geq 0 \quad \forall w \in \mathbb{R} \Rightarrow f(w) \text{ é convexa}$$

   Vejamos:
   dom $f = \mathbb{R}$ é um conjunto convexo, aberto e não vazio.
   Temos:

   $$f'(w) = a \quad \text{e} \quad f''(w) = 0$$

   Como $f'(w)$ e $f''(w)$ são contínuas em $\mathbb{R}$ e $f''(w) = 0\ \forall w \in \mathbb{R}$, então $f(w)$ é convexa em $\mathbb{R}$.

   Nota: Se f é uma função linear, então $-f$ é também linear. Este exercício mostra que as funções lineares são simultaneamente convexas e côncavas.

   (b) $f(w) = a^2 w + bw + c; \quad a \in \mathbb{R}^+;\ b, c \in \mathbb{R};\ a \neq 0$
   Resolução: Queremos mostrar:

   $$f''(w) \geq 0 \quad \forall w \in \mathbb{R} \Rightarrow f(w) \text{ é convexa}$$

   Vejamos:
   dom $f = \mathbb{R}$ é um conjunto convexo, aberto e não vazio.
   Temos:

   $$f'(w) = 2aw + b \quad \text{e} \quad f''(w) = 2a$$

   Como $f'(w)$ e $f''(w)$ são contínuas em $\mathbb{R}$ e $f''(w) = 2a > 0\ \forall w \in \mathbb{R}$ (a é positivo), então $f(w)$ é convexa em $\mathbb{R}$.

(c) $f(w) = e^{aw}$, $a \in \mathbb{R}$.
Resolução: Queremos mostrar:

$$f''(w) \geq 0 \quad \forall w \in \mathbb{R} \Rightarrow f(w) \text{ é convexa}$$

Vejamos:
dom $f = \mathbb{R}$ é um conjunto convexo, aberto e não vazio.
Temos:

$$f'(w) = ae^{aw} \quad \text{e} \quad f''(w) = a^2e^{aw}$$

Notar que $a^2 \geq 0$ e $e^{aw} > 0$ $\forall w \in \mathbb{R}$.

Como $f'(w)$ e $f''(w)$ são contínuas em $\mathbb{R}$ e $f''(w) = a^2e^{aw} \geq 0 \forall w \in \mathbb{R}$, então $f(w)$ é convexa em $\mathbb{R}$.

(d) $f(w) = w^a$ (definida em $\mathbb{R}^+$), $a \geq 1$ ou $a \leq 0$.
Resolução: Queremos mostrar:

$$f''(w) \geq 0 \quad \forall w \in \mathbb{R}^+ \Rightarrow f(w) \text{ é convexa}$$

Vejamos:
dom $f = \mathbb{R}^+$ é um conjunto convexo, aberto e não vazio.
Temos:

$$f'(w) = aw^{a-1} \quad \text{e} \quad f''(w) = a(a-1)w^{a-2}$$

Notar que $w^{a-2} > 0$ para $\forall w \in \mathbb{R}^+$. Assim, $f''(w) = a(a-1)w^{a-2} \geq 0$ se $a(a-1) \geq 0$, ou seja, se $(a \geq 0 \wedge a-1 \geq 0$ ou $a \leq 0 \wedge a-1 \leq 0)$. Desta condição resulta: $a \geq 1$ ou $a \leq 0$, como é indicada no enunciado. Assim, $f''(w) \geq 0$ para $a \geq 1$ ou $a \leq 0$, $\forall w \in \mathbb{R}^+$.
Portanto $f(w)$ é convexa em $\mathbb{R}^+$.

(e) $f(w) = -w^a$ (definida em $\mathbb{R}^+$), $0 \leq a \leq 1$.
Resolução: Queremos mostrar:

$$f''(w) \geq 0 \quad \forall w \in \mathbb{R} \Rightarrow f(w) \text{ é convexa}$$

Vejamos:
dom $f = \mathbb{R}^+$ é um conjunto convexo, aberto e não vazio.
Temos:

$$f'(w) = -aw^{a-1} \quad \text{e} \quad f''(w) = -a(a-1)w^{a-2}$$

Notar que $w^{a-2} > 0$ para $\forall w \in \mathbb{R}^+$. Assim, $f''(w) = -a(a-1)w^{a-2} \geq 0$ se $-a(a-1) \geq 0$, isto é, se $a(a-1) \leq 0$. Esta condição é verificada para $a \geq 0 \wedge a-1 \leq 0$, ou seja, para $0 \leq a \leq 1$ 0, como é indicada no enunciado. Assim, $f''(w) \geq 0$ para $0 \leq a \leq 1$ $\forall w \in \mathbb{R}^+$.
Portanto $f(w)$ é convexa em $\mathbb{R}^+$.

6. Mostre, usando o Teorema 3 que $F(w_1, w_2) = \ln(e_1^w + e_2^w)$, $F : \mathbb{R}^2 \to \mathbb{R}$, é uma função convexa.
Resolução: Queremos mostrar:

$$\nabla^2 F(w_1, w_2) \text{ é semi-definida positiva } \forall w \in \mathbb{R} \Rightarrow F \quad \text{é convexa.}$$

Vejamos:
$\mathbb{R}^2$ é um conjunto convexo, aberto e não vazio.

Derivadas parciais de primeira ordem:

$$\begin{aligned}
\frac{\partial F}{\partial w_1}(w_1, w_2) &= \frac{e^{w_1}}{e^{w_1} + e^{w_2}}, \\
\frac{\partial F}{\partial w_2}(w_1, w_2) &= \frac{e^{w_2}}{e^{w_1} + e^{w_2}}.
\end{aligned}$$

Derivadas parciais de segunda ordem:

$$\begin{aligned}
\frac{\partial^2 F}{\partial w_1^2}(w_1,w_2) &= \frac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2},\\
\frac{\partial^2 F}{\partial w_1 \partial w_2}(w_1,w_2) &= -\frac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2},\\
\frac{\partial^2 F}{\partial w_2^2}(w_1,w_2) &= \frac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2},
\end{aligned}$$

e, portanto, a matriz Hessiana é: $\nabla^2 F(w_1,w_2) = \begin{bmatrix} \dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} & -\dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} \\ -\dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} & \dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} \end{bmatrix}$.

Vamos agora calcular o determinante de $\nabla^2 F(w_1,w_2)$ usando o Critério de Sylvester:

- $\left|\dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2}\right| = \frac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} > 0.$
- $\begin{vmatrix} \dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} & -\dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} \\ -\dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} & \dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} \end{vmatrix} = \left(\frac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} \cdot \frac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2}\right) - \left(-\frac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} \cdot -\frac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2}\right) = 0$

Como $\left|\dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2}\right| > 0$ e $\begin{vmatrix} \dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} & -\dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} \\ -\dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} & \dfrac{e^{w_1}e^{w_2}}{(e^{w_1}+e^{w_2})^2} \end{vmatrix} = 0$, a matriz Hessiana é semidefinida positiva. Portanto F é convexa em $\mathbb{R}^2$.

7. Mostre, usando o Teorema 3 que a função quadrática $G(w) = \frac{1}{2}w^T Q w + c^T w + b$; ($Q$ é uma matriz simétrica $d \times d$, $c \in \mathbb{R}^d$, $b \in \mathbb{R}$), $G : \mathbb{R}^d \to \mathbb{R}$, é:

   (a) Convexa se e só se a matriz Q é semi-definida positiva.

   (b) Estritamente convexa se e só se Q é definida positiva.

   (c) Côncava se e só se a matriz Q é semi-definida negativa.

   (d) Estritamente côncava se e só se Q é definida negativa.

   Resolução:
   $\mathbb{R}^d$ é um conjunto convexo, aberto e não vazio.

   Gradiente de G (vetor das derivadas parciais de primeira ordem):
   $\nabla G(w) = (\frac{1}{2}Qw + \frac{1}{2}Q^T w) + c = Qw + c$,
   pois Q é simétrica.

   Hessiana de G (matriz das derivadas parciais de segunda ordem):
   $\nabla^2 G(w) = Q$, para todo $w \in \mathbb{R}^d$.

   Pelo Teorema 3, uma função $F : \mathbb{R}^d \to \mathbb{R}$ é convexa se e só se a sua Hessiana é semi-definida positiva em todo o domínio, e é estritamente convexa se e só se a Hessiana é definida positiva. Analogamente, F é côncava (resp. estritamente côncava) se e só se a Hessiana é semi-definida negativa (resp. definida negativa).

   Como $\nabla^2 G(w) = Q$, temos que:

   (a) G é convexa se e só se Q é semi-definida positiva.

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

(c) G é côncava se e só se Q é semi-definida negativa.

(d) G é estritamente côncava se e só se Q é definida negativa.

Isto conclui a prova.