Capítulo 5
Funções Convexas

Conteúdo

FUNÇÕES convexas ou côncavas desempenham um papel especial em áreas como a Termodinâmica, a Mecânica Estatística, a Teoria das Probabilidades, na Teoria das Equações Diferenciais, no Cálculo Variacional e em diversas outras. Pretendemos neste breve capítulo apresentar suas definições e suas propriedades básicas para futuro uso e referência. Obtemos algumas desigualdades úteis envolvendo funções convexas e côncavas, a mais relevante sendo, talvez, a desigualdade de Jensen, apresentada na Proposição 5.10, página 260. Nestas Notas faremos uso de propriedades de funções convexas ou côncavas em diversos momentos, por exemplo, no tratamento da função Gama de Euler no Capítulo 7, página 282.

5.1 Funções Convexas. Definições e Propriedades Básicas

• Conjuntos convexos em espaços vetoriais reais

Seja $\mathcal{V}$ um espaço vetorial real. Um conjunto não-vazio $C \subset \mathcal{V}$ é dito ser um *conjunto convexo* se para todos $x, y \in C$ e todo $\lambda \in [0, 1]$ valer $\lambda x + (1-\lambda)y \in C$.

Se $C \subset \mathcal{V}$ é convexo e $z \in C$, dizemos que uma expressão do tipo $z = \lambda x + (1-\lambda)y$ com $x, y \in C$ e $\lambda \in [0, 1]$ é uma *decomposição convexa* de z. Há três situações nas quais uma tal decomposição convexa é trivial: quando $x = y = z$ e $\lambda \in [0, 1]$ é arbitrário, quando $x = z$, $\lambda = 1$ e $y \in C$ é arbitrário ou quando $y = z$, $\lambda = 0$ e $x \in C$ é arbitrário. Nesses casos a decomposição convexa é apenas $z = z$.

Seja $C \subset \mathcal{V}$ convexo. Dizemos que $z \in C$ é um *ponto interior* de C se existirem $x, y \in C$ com $x \neq y$ e $\lambda \in (0, 1)$ tais que $z = \lambda x + (1-\lambda)y$. Em outras palavras, $z \in C$ é um ponto interior de C se admitir ao menos uma decomposição convexa não-trivial.

Dizemos que $z \in C$ é um *ponto extremo* (ou *ponto extremal*) de C se não existirem $x, y \in C$ distintos e $\lambda \in (0, 1)$ tais que $z = \lambda x + (1-\lambda)y$. Em outras palavras, $z \in C$ é um ponto extremo de C se admitir somente decomposições convexas triviais, ou seja, se não for um ponto interior de C.

Exemplo 5.1 No caso em que $\mathcal{V}$ é o espaço $\mathbb{R}^n$, um conjunto $C \subset \mathbb{R}^n$ é convexo se e somente se o segmento de reta conectando dois pontos quaisquer de C estiver inteiramente contido em C. Um triângulo aberto ou fechado em $\mathbb{R}^2$ é um conjunto convexo. Se o triângulo for aberto, todos os seus pontos são interiores e não há pontos extremos. Se o triângulo for fechado todos os seus pontos são interiores, exceto seus três vértices, que são seus únicos pontos extremos. ⌑

• Funções convexas e côncavas em espaços vetoriais reais

Seja $\mathcal{V}$ um espaço vetorial real e seja $C \subset \mathcal{V}$ um conjunto convexo. Uma função $f : C \to \mathbb{R}$ é dita ser uma *função convexa* se para todos $x, y \in C$ e todo $\lambda \in [0, 1]$ valer a desigualdade

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y) . \tag{5.1}$$

Uma função $f : C \to \mathbb{R}$ é dita ser uma *função côncava* se para todos $x,\ y \in C$ e todo $\lambda \in [0,\ 1]$ valer a desigualdade

$$f\big(\lambda x + (1-\lambda)y\big) \ \geq \ \lambda f(x) + (1-\lambda) f(y) \ . \tag{5.2}$$

Uma função $f : C \to \mathbb{R}$ é dita ser uma *função estritamente convexa* se para todos $x,\ y \in C$ com $x \neq y$ e todo $\lambda \in (0,\ 1)$ valer $f\big(\lambda x + (1-\lambda)y\big) < \lambda f(x) + (1-\lambda) f(y)$. Uma função $f : C \to \mathbb{R}$ é dita ser uma *função estritamente côncava* se para todos $x,\ y \in C$ com $x \neq y$ e todo $\lambda \in (0,\ 1)$ valer $f\big(\lambda x + (1-\lambda)y\big) > \lambda f(x) + (1-\lambda) f(y)$.

É elementar constatar que uma função f é côncava se e somente se $-f$ for convexa. Com isso, propriedades de funções côncavas podem ser facilmente derivadas de propriedades correspondentes de funções convexas e, por isso, discutiremos majoritariamente as últimas. O mesmo vale para funções estritamente côncavas e estritamente convexas.

No que segue, estudaremos funções convexas definidas em conjuntos convexos de $\mathbb{R}$, de $\mathbb{R}^n$ e também de espaços vetoriais reais normados (não necessariamente de dimensão finita).

5.1.1 Funções Convexas de uma Variável

No que segue, consideraremos funções definidas em um conjunto convexo $I \subset \mathbb{R}$ de interior I^0 não-vazio. Podemos ter $I = \mathbb{R}$, ou um intervalo aberto, semiaberto ou fechado, como $[A,\ B]$, $(A,\ B)$, $[A,\ B)$, $(A,\ B]$, $[A,\ \infty)$, $(A,\ \infty)$, $(-\infty,\ A]$ ou $(-\infty,\ A)$, com $-\infty < A < B < \infty$. I^2 designa o produto Cartesiano $I \times I$ e I_d designa seu conjunto diagonal: $I_d := \{(x,\ x),\ x \in I\} \subset I^2$.

Uma função $f : I \to \mathbb{R}$ é dita ser uma *função convexa* se para todos $x,\ y \in I$ e todo $\lambda \in [0,\ 1]$ valer

$$f\big(\lambda x + (1-\lambda)y\big) \ \leq \ \lambda f(x) + (1-\lambda) f(y) \ . \tag{5.3}$$

Uma função $f : I \to \mathbb{R}$ é dita ser uma *função côncava* se para todos $x,\ y \in I$ e todo $\lambda \in [0,\ 1]$ valer

$$f\big(\lambda x + (1-\lambda)y\big) \ \geq \ \lambda f(x) + (1-\lambda) f(y) \ . \tag{5.4}$$

E. 5.1 *Exercício*. A noção de convexidade (de concavidade) possui uma interpretação geométrica muito simples para funções de uma variável real. Uma função $f : I \to \mathbb{R}$ é convexa (respectivamente, côncava) se e somente se dados dois pontos quaisquer $x < y$ de seu domínio o gráfico de f no intervalo $(x,\ y)$ ficar abaixo (respectivamente, acima) da linha reta que conecta o par $(x,\ f(x))$ ao par $(y,\ f(y))$, tal como expresso nos gráficos da Figura 5.1, página 247. Justifique essa afirmação com base nas definições. ✠

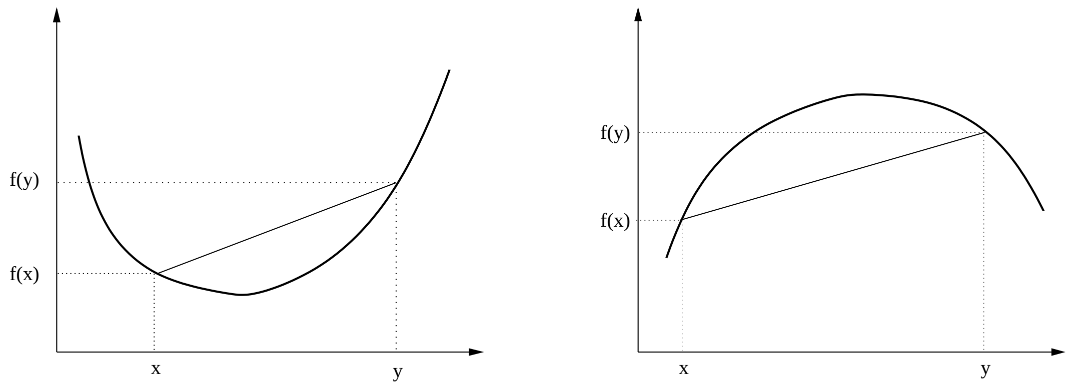

Figura 5.1: O gráfico de uma função convexa (à esquerda) e de uma função côncava (à direita). Em ambos é indicado o segmento de reta conectando par $(x,\ f(x))$ ao par $(y,\ f(y))$.

E. 5.2 *Exercício*. Seguindo a definição, mostre que $f(x) = |x|$ e $f(x) = x^2$ são funções convexas em $\mathbb{R}$. ✠

E. 5.3 *Exercício (fácil)*. Sejam $f : \mathbb{R} \to \mathbb{R}$ e $g : \mathbb{R} \to \mathbb{R}$ duas funções convexas. Mostre que se f é não-decrescente, então $f \circ g$ é convexa. ✠

Se $f : I \to \mathbb{R}$ é convexa, é muito fácil demonstrar, usando indução finita, que para todos $n \in \mathbb{N}$, todos $x_1, \ldots, x_n \in I$ e todos $\lambda_1, \ldots, \lambda_n \in [0, 1]$ tais que $\lambda_1 + \cdots + \lambda_n = 1$ vale

$$f\big(\lambda_1 x_1 + \cdots + \lambda_n x_n\big) \leq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n) . \tag{5.5}$$

Se f é côncava, temos

$$f\big(\lambda_1 x_1 + \cdots + \lambda_n x_n\big) \geq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n) . \tag{5.6}$$

Provaremos apenas o caso convexo, pois o outro é análogo. Suponhamos a afirmação válida para $n - 1$, com $n \geq 3$. Podemos supor que haja ao menos dois λ's não-nulos com $\lambda_1, \ldots, \lambda_n \in [0, 1]$ tais que $\lambda_1 + \cdots + \lambda_n = 1$, pois se houver apenas um, esse deve valer 1 e os demais 0 e não haveria o que se demonstrar. Sem perda de generalidade, suponhamos assim que $\lambda_{n-1} + \lambda_n > 0$. Então, como podemos escrever

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = \lambda_1 x_1 + \cdots + \lambda_{n-2} x_{n-2} + (\lambda_{n-1} + \lambda_n) \left[\frac{\lambda_{n-1} x_{n-1} + \lambda_n x_n}{\lambda_{n-1} + \lambda_n}\right]$$

temos, pela hipótese de indução,

$$\begin{aligned} f\big(\lambda_1 x_1 + \cdots + \lambda_n x_n\big) &= f\left(\lambda_1 x_1 + \cdots + \lambda_{n-2} x_{n-2} + (\lambda_{n-1} + \lambda_n) \left[\frac{\lambda_{n-1} x_{n-1} + \lambda_n x_n}{\lambda_{n-1} + \lambda_n}\right]\right) \\ &\leq \lambda_1 f(x_1) + \cdots + \lambda_{n-2} f(x_{n-2}) + (\lambda_{n-1} + \lambda_n) f\left(\frac{\lambda_{n-1} x_{n-1} + \lambda_n x_n}{\lambda_{n-1} + \lambda_n}\right) \\ &\leq \lambda_1 f(x_1) + \cdots + \lambda_{n-2} f(x_{n-2}) + \lambda_{n-1} f(x_{n-1}) + \lambda_n f(x_n) , \end{aligned}$$

sendo que, na última desigualdade, usamos a convexidade de f para obter

$$f\left(\frac{\lambda_{n-1} x_{n-1} + \lambda_n x_n}{\lambda_{n-1} + \lambda_n}\right) \leq \frac{\lambda_{n-1}}{\lambda_{n-1} + \lambda_n} f(x_{n-1}) + \frac{\lambda_n}{\lambda_{n-1} + \lambda_n} f(x_n) .$$

Isso provou (5.5) para todo $n \in \mathbb{N}$.

A desigualdade (5.5) (ou sua forma côncava (5.6)) é por vezes denominada *desigualdade de Jensen*[1]. É importante mencionar que a desigualdade de Jensen pode ser ainda generalizada e (5.5) é apenas sua versão mais simples (discreta). Para uma forma mais geral, vide Proposição 5.10, página 260, em especial, vide (5.29).

• Propriedades do conjunto de funções convexas em I

Se $f : I \to \mathbb{R}$ e $g : I \to \mathbb{R}$ são duas funções convexas, então para todos $\alpha, \beta \in [0, \infty)$ a função $\alpha f + \beta g$ é também convexa em I. A prova disso é elementar. Essa propriedade afirma que o conjunto das funções convexas em I é um cone convexo[2]. Essas afirmações valem também para funções côncavas.

Seja $f_n : I \to \mathbb{R}$, $n \in \mathbb{N}$ uma sequência de funções convexas que converge pontualmente a uma função $f : I \to \mathbb{R}$, ou seja, tal que para cada $x \in I$ valha $f(x) = \lim_{n\to\infty} f_n(x)$. Então, f é igualmente convexa. A prova dessa afirmação é elementar e deixada ao estudante. Essa afirmação vale também para funções côncavas.

Seja $\{f_\omega : I \to \mathbb{R},\ \omega \in \Omega\}$ uma família de funções convexas definidas em I tal que para cada $x \in I$ exista $f(x) := \sup\{f_\omega(x),\ \omega \in \Omega\}$. Então, $f : I \to \mathbb{R}$ é também convexa. A prova dessa afirmação é muito simples e deixada ao estudante. Para funções côncavas valem as mesmas afirmações, com o supremo substituído pelo ínfimo.

• Uma condição equivalente à de convexidade

Para uma função $f : I \to \mathbb{R}$, considere-se a função simétrica de duas variáveis $R_f \equiv R : I^2 \setminus I_d$ dada por

$$R_f(x, y) \equiv R(x, y) := \frac{f(x) - f(y)}{x - y} , \qquad x \neq y .$$

[1]Johan Ludwig William Valdemar Jensen (1859–1925). A desigualdade de Jensen, assim como outros trabalhos do mesmo sobre funções convexas, data de 1906. Para a referência original, vide nota de rodapé 4 à página 255.

[2]Para entender essa nomenclatura o estudante deve recordar que se $\mathcal{C}$ é um cone convexo em $\mathbb{R}^3$, então se $\vec{v}$ e $\vec{u}$ são vetores de $\mathcal{C}$, segue que $\alpha\vec{v} + \beta\vec{u}$ é também um vetor de $\mathcal{C}$ para todos $\alpha, \beta \in [0, \infty)$.

A proposição que segue mostra que com a função R podemos apresentar uma definição alternativa de convexidade (uma outra caracterização distinta da noção de função convexa será encontrada na Proposição 5.4, página 254).

Proposição 5.1 *Uma função $f : I \to \mathbb{R}$ é convexa se e somente se para todos $x,\ y,\ z \in I$, distintos, com $y < z$ valer*

$$R(x,\ y)\ \leq\ R(x,\ z)\ , \tag{5.7}$$

ou seja, se e somente se, fixado um dos argumentos, R for monotonamente não-decrescente no outro argumento. □

No Exercício E. 5.4, página 250, apresenta-se uma interpretação geométrica da Proposição 5.1.

Prova da Proposição 5.1. **Parte I:** *supondo f convexa provamos (5.7).* Para provarmos (5.7) há três casos a se considerar: $x < y < z$, $y < x < z$ e $y < z < x$.

Caso 1: $x < y < z$. Como y fica entre x e z, podemos escrever $y = \lambda x + (1 - \lambda)z$ adotando para tal $\lambda = \frac{z-y}{z-x}$ (notar que $\lambda \in (0,\ 1)$). Da convexidade de f segue, então, que

$$f(y)\ \leq\ \lambda f(x) + (1 - \lambda)f(z)\ =\ \frac{z - y}{x - z}f(x) + \frac{y - x}{z - x}f(z)\ .$$

Subtraindo $f(x)$ de ambos os lados obtemos, após alguns cômputos elementares,

$$\frac{f(y) - f(x)}{y - x}\ \leq\ \frac{f(z) - f(x)}{z - x}\ ,$$

ou seja, $R(x,\ y) \leq R(x,\ z)$.

Caso 2: $y < x < z$. Como x fica entre y e z, podemos escrever $x = \lambda y + (1 - \lambda)z$ adotando para tal $\lambda = \frac{z-x}{z-y}$ (notar que $\lambda \in (0,\ 1)$). Da convexidade de f segue, então, que

$$f(x)\ \leq\ \lambda f(y) + (1 - \lambda)f(z)\ =\ \frac{z - x}{z - y}f(y) + \frac{x - y}{z - y}f(z)\ .$$

Subtraindo $f(x)$ de ambos os lados, podemos escrever

$$0\ \leq\ \frac{z - x}{z - y}\big(f(y) - f(x)\big) + \frac{x - y}{z - y}\big(f(z) - f(x)\big)$$

do que segue imediatamente que

$$\frac{f(y) - f(x)}{y - x}\ \leq\ \frac{f(z) - f(x)}{z - x}\ ,$$

ou seja, $R(x,\ y) \leq R(x,\ z)$.

Caso 3: $y < z < x$. Como z fica entre y e x, podemos escrever $z = \lambda y + (1 - \lambda)x$ adotando para tal $\lambda = \frac{x-z}{x-y}$ (notar que $\lambda \in (0,\ 1)$). Da convexidade de f segue, então, que

$$f(z)\ \leq\ \lambda f(y) + (1 - \lambda)f(x)\ =\ \frac{x - z}{x - y}f(y) + \frac{z - y}{x - y}f(x)\ .$$

Subtraindo-se $f(x)$ de ambos os lados, obtém-se após cômputos elementares

$$\frac{f(y) - f(x)}{y - x}\ \leq\ \frac{f(z) - f(x)}{z - x}\ ,$$

ou seja, novamente $R(x,\ y) \leq R(x,\ z)$. Com isso, (5.7) está estabelecida em todos os casos possíveis em que $y < z$.

Parte II: *supondo (5.7) provamos que f é convexa.* Por (5.7) sabemos que se $y < z$, então

$$\frac{f(x) - f(y)}{x - y}\ \leq\ \frac{f(x) - f(z)}{x - z}$$

para todo x, com $x \neq y$ e $x \neq z$. Tomemos, em particular, $x = \lambda y + (1 - \lambda)z$, com $\lambda \in (0,\ 1)$. A última expressão fica

$$\frac{f\big(\lambda y + (1 - \lambda)z\big) - f(y)}{(1 - \lambda)(z - y)}\ \leq\ -\frac{f\big(\lambda y + (1 - \lambda)z\big) - f(z)}{\lambda(z - y)}\ .$$

Cancelando-se o fator $z - y > 0$ de ambos os lados, obtemos $f\big(\lambda y + (1-\lambda)z\big) \leq \lambda f(y) + (1-\lambda)f(z)$, provando a convexidade de f. ■

E. 5.4 *Exercício*. A Proposição 5.1 possui a seguinte interpretação geométrica. Sejam $x,\ y,\ z \in I$ três pontos do domínio de definição de uma função f tais que $y < z$. Considere-se o segmento de linha reta L_1 conectando o ponto $\big(x,\ f(x)\big)$ ao ponto $\big(y,\ f(y)\big)$ (ambos no gráfico gráfico de f) e considere-se o segmento de linha reta L_2 conectando o ponto $\big(x,\ f(x)\big)$ ao ponto $\big(z,\ f(z)\big)$ (ambos também no gráfico de f). Então, o que a Proposição 5.1 afirma é que f é convexa se e somente se a inclinação de L_1 for menor ou igual à inclinação de L_2. Vide o gráfico no lado esquerdo da Figura 5.2, página 250. Justifique essa afirmativa com base na Proposição 5.1. ✠

• Algumas desigualdades de interesse

Antes de prosseguirmos apresentemos um resultado que será futuramente evocado nestas Notas.

Lema 5.1 *Seja $f : I \to \mathbb{R}$ convexa. Então, para quaisquer $w,\ x,\ y,\ z \in I$ com $w < x < y < z$ valem as desigualdades*

$$\frac{f(w)-f(x)}{w-x} \leq \frac{f(x)-f(y)}{x-y} \leq \frac{f(y)-f(z)}{y-z} \tag{5.8}$$

e

$$\frac{f(w)-f(x)}{w-x} \leq \frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x} . \tag{5.9}$$

□

Prova. Pelas hipóteses e pela Proposição 5.1, página 249, tem-se $R(w,\ x) \leq R(x,\ y) \leq R(y,\ z)$, assim como $R(x,\ w) \leq R(x,\ y) \leq R(x,\ z)$. Escrevendo-se explicitamente o que é a função R, obtemos disso (5.8) e (5.9), respectivamente. ■

O estudante deve atentar para as semelhanças e diferenças entre (5.8) e (5.9). A primeira desigualdade em ambas é a mesma. A diferença está na segunda desigualdade. A desigualdade (5.8) é graficamente representada no gráfico à direita da Figura 5.2, página 250. A desigualdade (5.9) é graficamente representada no gráfico à esquerda da Figura 5.2.

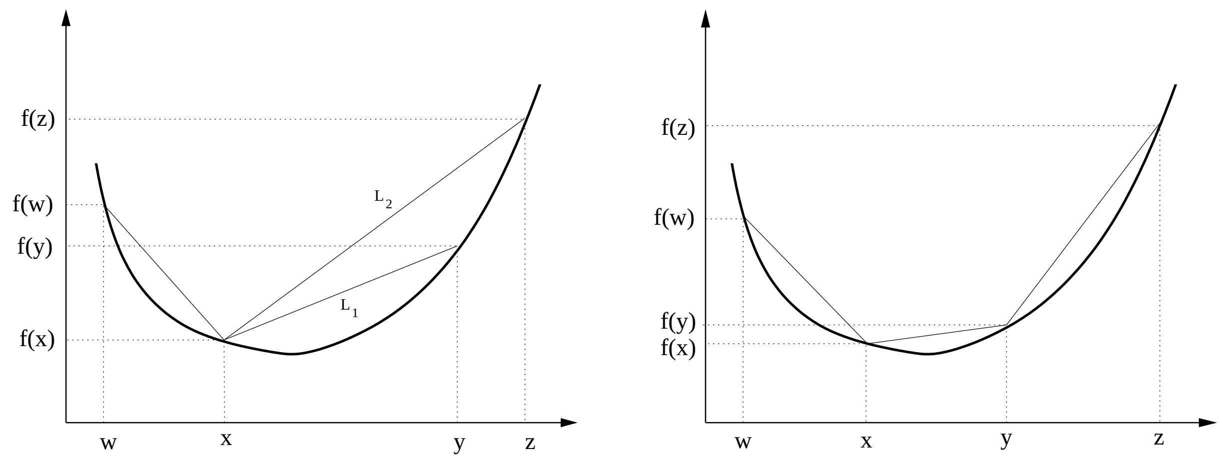

Figura 5.2: À esquerda: gráfico de uma função convexa. As inclinações dos segmentos de reta lá indicados são $R(w,\ x)$, $R(x,\ y)$ e $R(x,\ z)$. A figura deixa claro que $R(w,\ x) \leq R(x,\ y) \leq R(x,\ z)$. À direita: gráfico de uma função convexa. As inclinações dos segmentos de reta lá indicados são $R(w,\ x)$, $R(x,\ y)$ e $R(y,\ z)$. A figura deixa claro que $R(w,\ x) \leq R(x,\ y) \leq R(y,\ z)$.

• Convexidade e derivadas laterais

Seja uma função $g : I \to \mathbb{R}$. Para $x \in I^0$ definimos as derivadas laterais que denotamos por $g'_+(x)$ e $g'_-(x)$ por

$$g'_+(x) := \lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} \frac{g(x+\epsilon)-g(x)}{\epsilon} \qquad \text{e} \qquad g'_-(x) := \lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} \frac{g(x)-g(x-\epsilon)}{\epsilon} ,$$

caso esses limites existam. É relevante notar que caso ambos os limites existam, então g é contínua em x, pois teremos $\lim_{\substack{\epsilon\to 0\\ \epsilon>0}}\big(g(x+\epsilon)-g(x)\big)=0$ e $\lim_{\substack{\epsilon\to 0\\ \epsilon>0}}\big(g(x)-g(x-\epsilon)\big)=0$. É claro que g é diferenciável em x se e somente se $g'_+(x)$ e $g'_+(x)$ existirem em forem iguais. O proposição que segue revela mais fatos básicos importantes sobre funções convexas.

Proposição 5.2 *Se $f : I \to \mathbb{R}$ é convexa, então f é contínua em I^0 e possui em cada ponto de I^0 derivadas laterais à direita e à esquerda, as quais satisfazem a seguinte desigualdade:*

$$f'_-(x) \;\leq\; f'_+(x) \;\leq\; \frac{f(x')-f(x)}{x'-x} \;\leq\; f'_-(x') \;\leq\; f'_+(x') \tag{5.10}$$

para todos $x,\ x' \in I^0$ com $x < x'$. Isso diz que tanto f'_- quanto f'_+ são funções monotonamente não-decrescentes em I^0, sendo que $f'_- \leq f'_+$ em todo I^0. Outra afirmação que disso pode ser extraída é que f é não-diferenciável em uma coleção no máximo enumerável de pontos. □

Uma demonstração mais geral da continuidade de funções convexas no interior do seu domínio de definição será apresentada na Proposição 5.9, página 258.

Prova da Propossição 5.2. Sejam $w,\ x,\ y \in I^0$ com $w < x < y$. Pela Proposição 5.1, página 249, temos $R(x,\ w) \leq R(x,\ y)$.

Fixemos w e x. Sabemos, também pela Proposição 5.1, que a função $y \mapsto R(x,\ y)$ definida para $y > x$ é decrescente quando y diminui para x. Assim, o limite $\lim_{\substack{y\to x\\ y>x}} R(x,\ y)$ existe, por ser decrescente e limitado inferiormente por $R(x,\ w)$. Sucede que, pela definição de R, o limite $\lim_{\substack{y\to x\\ y>x}} R(x,\ y)$ é precisamente a derivada lateral à direita $f'_+(x)$.

Fixemos x e y. Sabemos, também pela Proposição 5.1, que a função $w \mapsto R(x,\ w)$ definida para $w > x$ é crescente quando w cresce para x. Assim, o limite $\lim_{\substack{w\to x\\ w<x}} R(x,\ w)$ existe, por ser crescente e limitado superiormente por $R(x,\ y)$. Sucede que, pela definição de R, o limite $\lim_{\substack{w\to x\\ w<x}} R(x,\ w)$ é precisamente a derivada lateral à esquerda $f'_-(x)$.

Isso estabeleceu a existência dos limites laterais para todo ponto de I^0 e estabeleceu que f é contínua em todo ponto de I^0.

Sejam agora $w,\ x,\ y,\ w',\ x',\ y'$ seis pontos de I^0 tais que $w < x < y < w' < x' < y'$. Fazendo uso da Proposição 5.1, temos

$$R(w,\ x) \;\leq\; R(x,\ y) \;\leq\; R(y,\ w') \;\leq\; R(w',\ x') \;\leq\; R(x',\ y')\ ,$$

ou seja,

$$\frac{f(x)-f(w)}{x-w} \;\leq\; \frac{f(y)-f(x)}{y-x} \;\leq\; \frac{f(w')-f(y)}{w'-y} \;\leq\; \frac{f(x')-f(w')}{x'-w'} \;\leq\; \frac{f(y')-f(x')}{y'-x'}\ . \tag{5.11}$$

Tomando-se em (5.11) os limites $\lim_{\substack{w\to x\\ w<x}}$, $\lim_{\substack{y\to x\\ y>x}}$, $\lim_{\substack{w'\to x'\\ w'<x'}}$, $\lim_{\substack{y'\to x'\\ y'>x'}}$ e usando-se a continuidade de f, obtemos

$$f'_-(x) \;\leq\; f'_+(x) \;\leq\; \frac{f(x')-f(x)}{x'-x} \;\leq\; f'_-(x') \;\leq\; f'_+(x')\ ,$$

que é (5.10).

Seja $N \subset I^0$ a coleção de todos os pontos de I^0 nos quais f não seja diferenciável, ou seja, nos quais $f'_-(x) \neq f'_+(x)$. (5.10) informa-nos que se $x,\ x' \in N$ com $x < x'$, então os intervalos $\big(f'_-(x),\ f'_+(x)\big)$ e $\big(f'_-(x'),\ f'_+(x')\big)$ são intervalos disjuntos de $\mathbb{R}$. Agora, $\mathbb{R}$ pode no máximo admitir uma família enumerável de intervalos disjuntos. Logo, N é no máximo enumerável. ■

É instrutivo chamar a atenção do leitor para o fato de a continuidade a que se refere a Proposição 5.2 ser garantida apenas no interior I^0 do domínio de definição I. As funções $f : [0,\ 1] \to \mathbb{R}$ definidas por

$$f(x) \;:=\; \begin{cases} 1\ , & x = 0\ , \\ x^2\ , & x \in (0,\ 1]\ , \end{cases} \qquad f(x) \;:=\; \begin{cases} 1\ , & x = 0\ , \\ 0\ , & x \in (0,\ 1]\ , \end{cases}$$

são convexas em todo o intervalo $[0,\ 1]$, mas não são contínuas em $x = 0$.

A Proposição 5.2, página 251, tem a seguinte consequência, que usaremos quando apresentarmos a demonstração de uma forma geral da desigualdade de Jensen na Proposição 5.10, página 260:

Corolário 5.1 *Se $f : I \to \mathbb{R}$ é convexa, então*

$$f(y) \;\geq\; (y-x)f'_{\pm}(x) + f(x) \tag{5.12}$$

para todos $x,\ y \in I$. □

Esse corolário tem a seguinte interpretação geométrica: o gráfico de uma função convexa está sempre acima das retas tangentes ao mesmo (e isso é verdade mesmo em pontos em que a derivada é descontínua, em cujo caso temos duas retas tangentes com inclinações f'_- e f'_+).

Prova do Corolário 5.1. Para $x' > x$ temos de (5.10) as desigualdades $f'_-(x) \leq f'_+(x) \leq \frac{f(x')-f(x)}{x'-x}$, que implicam que

$$f(x') \;\geq\; (x'-x)f'_{\pm}(x) + f(x)\ , \qquad x' \geq x\ . \tag{5.13}$$

Incluímos acima o caso $x' = x$ devido à continuidade de f (também demonstrada na Proposição 5.2).

Também para $x' > x$ temos de (5.10) as desigualdades $\frac{f(x')-f(x)}{x'-x} \leq f'_-(x') \leq f'_+(x')$, que implicam que

$$f(x) \;\geq\; (x-x')f'_{\pm}(x') + f(x')\ , \qquad x' > x\ .$$

Trocando as letras $x \leftrightarrow x'$, isso fica

$$f(x') \;\geq\; (x'-x)f'_{\pm}(x) + f(x)\ , \qquad x > x'\ . \tag{5.14}$$

Contemplando (5.13) e (5.14), vemos que estabelecemos que $f(x') \geq (x'-x)f'_{\pm}(x) + f(x)$ para todos $x,\ x' \in I$. ■

A Proposição 5.2, página 251, afirma que se uma função é convexa, então suas derivadas laterais são crescentes. Sob hipóteses adequadas é possível garantir a recíproca dessa afirmação. A proposição que segue mostra a forma mais simples dessa recíproca.

Proposição 5.3 *Seja $f : I \to \mathbb{R}$ contínua em I e diferenciável em I^0. Então, uma condição necessária e suficiente para que f seja convexa é que f' seja monotonamente não-decrescente, ou seja, que $f'(x) \leq f'(y)$ para todos $x,\ y \in I^0$ com $x \leq y$.* □

Prova. Da Proposição 5.2, página 251, é evidente que convexidade e diferenciabilidade implicam que f' é monotonamente não-decrescente, de modo que resta apenas provar a recíproca.

Sejam $x_0,\ x_1 \in I$ com $x_0 < x_1$ e seja $\lambda \in (0,\ 1)$. Definamos $x_\lambda := \lambda x_0 + (1-\lambda)x_1$. É claro que $x_0 < x_\lambda < x_1$. Com as hipóteses, podemos evocar o Teorema do Valor Médio e afirmar que existem $\xi_0 \in \big(x_0,\ x_\lambda\big)$ e $\xi_1 \in \big(x_\lambda,\ x_1\big)$ tais que valem

$$\begin{aligned} f(x_\lambda) - f(x_0) &= f'(\xi_0)\big(x_\lambda - x_0\big) \;=\; (1-\lambda)(x_1-x_0)f'(\xi_0)\ , \\ f(x_1) - f(x_\lambda) &= f'(\xi_1)\big(x_1 - x_\lambda\big) \;=\; \lambda(x_1-x_0)f'(\xi_1)\ . \end{aligned}$$

Note-se que $\xi_0 < \xi_1$ e, portanto, $f'(\xi_0) \leq f'(\xi_1)$. Temos, assim,

$$f(x_\lambda) - \lambda f(x_0) - (1-\lambda)f(x_1) \;=\; \lambda\big(f(x_\lambda) - f(x_0)\big) + (1-\lambda)\big(f(x_\lambda) - f(x_1)\big) \;=\; \lambda(1-\lambda)(x_1-x_0)\big(f'(\xi_0) - f'(\xi_1)\big) \;\leq\; 0\ ,$$

pois $f'(\xi_0) \leq f'(\xi_1)$. Isso estabeleceu que $f\big(\lambda x_0 + (1-\lambda)x_1\big) \leq \lambda f(x_0) + (1-\lambda)f(x_1)$ com $\lambda \in (0,\ 1)$. Para $\lambda \in \{0,\ 1\}$ essa relação é trivial e isso demonstra a convexidade de f. ■

O seguinte corolário é agora evidente e dispensa demonstração.

Corolário 5.2 *Seja $f : I \to \mathbb{R}$ contínua em I e duas vezes diferenciável em I^0. Então, uma condição necessária e suficiente para que f seja convexa é que $f''(x) \geq 0$ para todo $x \in I^0$.* □

Esse corolário permite confortavelmente determinar se uma função f contínua e duas vezes diferenciável em seu domínio de definição I é convexa (o que se dá caso $f''(x) \geq 0$ para todo $x \in I^0$) ou côncava (o que se dá caso $f''(x) \leq 0$ para todo $x \in I^0$).

Exemplos 5.2 De posse do critério estabelecido no Corolário 5.2, é fácil demonstrar os seguintes fatos. A função $\ln x$ é côncava em $(0, \infty)$. As funções $e^{\pm x}$ são convexas em $\mathbb{R}$. As funções x^n com $n \in \mathbb{N}$ par são convexas em $\mathbb{R}$. As funções x^n com $n \in \mathbb{N}$ ímpar são convexas em $[0, \infty)$ e côncavas em $(-\infty, 0]$. A função $1/(1-x^2)$ é convexa no intervalo $(-1, 1)$ e diverge para $x \to \pm 1$. A função $x + 1/x$ é convexa em $(0, \infty)$ e côncava em $(-\infty, 0)$. As funções $\cos x$ e $\operatorname{sen} x$ são côncavas nos intervalos em que são positivas e convexas nos intervalos em que são negativas. ⌑

E. 5.5 *Exercício (fácil).* Considere-se a função $s : [0, 1] \to \mathbb{R}$ dada por

$$s(x) := \begin{cases} 0, & \text{se } x = 0, \\ -x\ln(x), & \text{se } x \in (0, 1]. \end{cases}$$

Constate que $s(x) \geq 0$ para todo $x \in [0, 1]$, que s é uma função contínua, duas vezes diferenciável em $(0, 1]$ e côncava. ✠

A Proposição 5.3 tem ainda um outro corolário digno de nota:

Corolário 5.3 *Seja $I \subset \mathbb{R}$ um intervalo convexo do tipo que aqui consideramos e seja $g : I \to \mathbb{R}$ uma função contínua, e não-decrescente. Então, a função $h : I \to \mathbb{R}$ definida por*

$$h(x) := \int_a^x g(u)\, du$$

com $a \in I$, fixo, é uma função convexa da variável x. □

Prova. É claro que h é contínua e que $h'(x) = g(x)$ para todo $x \in I$. Logo, h' é estritamente crescente e, pela Proposição 5.3, página 252, h é convexa. ■

• A condição do ponto médio

Vamos agora apresentar mais uma caracterização de funções convexas e contínuas, a qual possui diversas aplicações. O que mostraremos é que uma função contínua $f : I \to \mathbb{R}$ é convexa se e somente se satisfizer a desigualdade $f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$ para todos $x, y \in I$. Como comentaremos, foi essa caracterização que deu origem histórica à teoria das funções convexas.

Começamos com o seguinte resultado, cuja demonstração é assaz interessante:

Lema 5.2 *Seja $f : I \to \mathbb{R}$ tal que*

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} \tag{5.15}$$

para todos $x, y \in I$. Então, vale

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n} \tag{5.16}$$

para todo $n \in \mathbb{N}$ e para todos $x_1, \ldots, x_n \in I$. □

Prova. A prova é feita seguindo uma curiosa estratégia de indução[3]: primeiramente mostramos que se a proposição vale para n, então vale para todo número da forma $2^k n$ com $k \in \mathbb{N}$ (indução para a frente!). Em seguida, provamos que se a proposição vale para m ela vale para $m - 1$ (indução para trás!). Com isso, todos os naturais são varridos pelo procedimento indutivo.

[3]Essa estratégia foi inventada por Cauchy para demonstrar (5.33) e foi empregada no presente contexto por Jensen (para a referência original, vide nota de rodapé 4 à página 255). O texto original de Cauchy é *Cours d'analyse de l'Ecole Royale Polytechnique, premier partie, Analyse algebrique*, Paris 1821.

Vamos assumir (5.16) válida para algum $n \in \mathbb{N}$ (ela vale para $n = 2$, por hipótese). Sejam $x_1, \ldots, x_{2n} \in I$. Então,

$$f\left(\frac{x_1 + \cdots + x_{2n}}{2n}\right) = f\left(\frac{\frac{x_1+x_2}{2} + \cdots + \frac{x_{2n-1}+x_{2n}}{2}}{n}\right) \overset{\text{hipótese}}{\leq} \frac{1}{n}\left(f\left(\frac{x_1 + x_2}{2}\right) + \cdots + f\left(\frac{x_{2n-1} + x_{2n}}{2}\right)\right)$$

$$\overset{(5.15)}{\leq} \frac{f(x_1) + \cdots + f(x_{2n})}{2n} .$$

Assim, se (5.16) vale para n, vale também para $2n$ e, consequentemente, para todo número da forma $2^k n$ com $k \in \mathbb{N}$. Isso prova os passos indutivos para a frente.

Vamos agora assumir que (5.16) valha para algum $n \geq 4$. Seja $x_1, \ldots, x_{n-1} \in I$ e defina-se $x_n := \frac{1}{n-1}\big(x_1+\cdots+x_{n-1}\big)$. É evidente que $x_n \in I$ e teremos,

$$\frac{x_1 + \cdots + x_n}{n} = \frac{x_1 + \cdots + x_{n-1}}{n-1} .$$

Verifique! Logo,

$$\begin{aligned} f\left(\frac{x_1 + \cdots + x_{n-1}}{n-1}\right) &= f\left(\frac{x_1 + \cdots + x_n}{n}\right) \\ &\overset{\text{hipótese}}{\leq} \frac{f(x_1) + \cdots + f(x_n)}{n} \\ &= \frac{f(x_1) + \cdots + f(x_{n-1})}{n} + \frac{1}{n} f\left(\frac{x_1 + \cdots + x_{n-1}}{n-1}\right) . \end{aligned}$$

Passando o termo $\frac{1}{n} f\left(\frac{x_1+\cdots+x_{n-1}}{n-1}\right)$ para o lado esquerdo da desigualdade, obtemos imediatamente que

$$f\left(\frac{x_1 + \cdots + x_{n-1}}{n-1}\right) = \frac{f(x_1) + \cdots + f(x_{n-1})}{n-1} ,$$

provando que (5.16) vale para $n - 1$. Isso prova os passos de indução retrógrada e completa a demonstração. ■

O seguinte resultado apresenta uma caracterização muito útil de funções convexas:

Proposição 5.4 *Seja $f : I \to \mathbb{R}$, contínua. Então, f é convexa em I se e somente se satisfizer*

$$f\left(\frac{x + y}{2}\right) \leq \frac{f(x) + f(y)}{2} \tag{5.17}$$

para todos $x, y \in I$. □

Prova. Se f for convexa em I, (5.17) é um caso particular da definição de convexidade (tome-se $\lambda = 1/2$). Vamos provar que se f é contínua em I e lá satisfaz (5.17), então f é convexa em I.

Sejam $a_1, a_2 \in \mathbb{N}$ dois números naturais. Sejam também $x_1, x_2 \in I$. É claro que $\big(a_1x_1 + a_2x_2\big)/\big(a_1 + a_2\big)$ é um elemento de I. Como

$$a_1x_1 + a_2x_2 = \underbrace{x_1 + \cdots + x_1}_{a_1 \text{vezes}} + \underbrace{x_2 + \cdots + x_2}_{a_2 \text{vezes}}$$

é uma soma de $a_1 + a_2$ elementos de I, podemos evocar o Lema 5.2, página 253, em particular, a relação (5.16) com $n = a_1 + a_2$, e escrever

$$f\left(\frac{a_1x_1 + a_2x_2}{a_1 + a_2}\right) \leq \frac{\overbrace{f(x_1) + \cdots + f(x_1)}^{a_1 \text{vezes}} + \overbrace{f(x_2) + \cdots + f(x_2)}^{a_2 \text{vezes}}}{a_1 + a_2} = \frac{a_1 f(x_2) + a_2 f(x_2)}{a_1 + a_2} . \tag{5.18}$$

Afirmamos que isso implica que se $r \in \mathbb{Q} \cap [0,\ 1]$, então

$$f\big(rx_1 + (1-r)x_2\big) \ \leq \ rf(x_1) + (1-r)f(x_2) \ . \tag{5.19}$$

É suficiente tomarmos $r \in \mathbb{Q} \cap (0,\ 1)$. Como tal, r é da forma $r = p/q$ com $0 < p < q$, sendo $p,\ q \in \mathbb{N}$. Se definirmos $a_1 := p$ e $a_2 := q - p$, teremos $a_1,\ a_2 \in \mathbb{N}$ e poderemos escrever $r = \frac{a_1}{a_1+a_2}$ e $1 - r = \frac{a_2}{a_1+a_2}$. Com essas observações torna-se evidente que a validade de (5.19) segue de (5.18).

Em (5.19), façamos agora o racional r convergir a $\lambda \in [0,\ 1]$, arbitrário. A continuidade da f implica que teremos $f\big(\lambda x_1 + (1-\lambda)x_2\big) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$. Como isso é válido para qualquer $\lambda \in [0,\ 1]$ e quaisquer $x_1,\ x_2 \in I$, a convexidade de f fica estabelecida. ■

Nota histórica. A noção de função convexa foi introduzida por Jensen[4] em 1906. A definição original de Jensen para função convexa era a relação (5.17). Assumindo continuidade para f, Jensen então seguiu os passos que apresentamos na demonstração da Proposição 5.4 e demonstrou (5.3) assim como (5.5) (que ficou conhecida como *desigualdade de Jensen*). ♣

Comentamos, por fim, que a condição de continuidade não pode ser dispensada da Proposição 5.4. Com o uso de bases de Hamel (vide discussão à página 138 e seguintes e, em particular, vide a discussão que sucede a Proposição 2.6, página 140) é possível construir funções não-contínuas em $\mathbb{R}$ satisfazendo (5.17) para todos $x,\ y \in \mathbb{R}$ mas que não são convexas (pois seu gráfico é denso em todo $\mathbb{R}^2$ (!)).

• Mais duas caracterizações de funções contínuas convexas

O lema a seguir contém uma informação útil que será evocada adiante.

Lema 5.3 *Seja $f : I \to \mathbb{R}$ uma função convexa. Então, dado qualquer intervalo compacto $[a,\ b] \subset I$ (com $-\infty < a < b < \infty$), o máximo de f nesse intervalo é alcançado em um dos seus pontos extremos, a ou b, ou seja,* $\max\{f(x),\ x \in [a,\ b]\} = \max\{f(a),\ f(b)\}$. □

Prova. Para todo $\lambda \in [0,\ 1]$ tem-se, evidentemente, $\lambda f(a) + (1-\lambda)f(b) \leq \max\{f(a),\ f(b)\}$. Se $x \in [a,\ b]$, então $x = \lambda_0 a + (1-\lambda_0)b$ para algum $\lambda_0 \in [0,\ 1]$. Logo, $f(x_0) = f\big(\lambda_0 a + (1-\lambda_0)b\big) \leq \lambda_0 f(a) + (1-\lambda_0)f(b) \leq \max\{f(a),\ f(b)\}$. Claro é que essa desigualdade $f(x_0) \leq \max\{f(a),\ f(b)\}$ é uma igualdade se x_0 for igual a a ou b. ■

O estudante deve notar que a recíproca do Lema 5.3 não é verdadeira: a função cosseno restrita ao intervalo $I = [0,\ 2\pi]$ é contínua e satisfaz $\max\{\cos(x),\ x \in [a,\ b]\} = \max\{\cos(a),\ \cos(b)\}$ para todo compacto $[a,\ b] \subset [0,\ 2\pi]$. Mas a função cosseno não é convexa em $[0,\ 2\pi]$. Em verdade, toda função f definida em um intervalo fechado $[A,\ B]$, contínua, sem nenhum máximo local exceto, eventualmente, os pontos A e B satisfaz $\max\{f(x),\ x \in [a,\ b]\} = \max\{f(a),\ f(b)\}$ para todo $[a,\ b] \subset [A,\ B]$.

A proposição que segue, mencionada em [137], apresenta uma caracterização de convexidade para funções contínuas. Ela mostra o que é necessário supor adicionalmente para que se obtenha uma recíproca à afirmação do Lema 5.3.

Proposição 5.5 *Seja $f : I \to \mathbb{R}$ uma função contínua. Então, f é convexa se e somente se para todo $\alpha \in \mathbb{R}$ o máximo da função $\phi_\alpha(x) := f(x) + \alpha x$ em um intervalo compacto arbitrário $[a,\ b] \subset I$ for sempre assumido em um dos extremos do mesmo. Em outras palavras, f é convexa se e somente se para todo $\alpha \in \mathbb{R}$ valer* $\max\{\phi_\alpha(x),\ x \in [a,\ b]\} = \max\{\phi_\alpha(a),\ \phi_\alpha(b)\}$ *em todo compacto $[a,\ b] \in I$.* □

Prova. É trivial constatar que $\phi_\alpha(x) := f(x) + \alpha x$ é convexa se e somente se f o for. Logo, pelo Lema 5.3, página 255, os máximos da função ϕ_α em um intervalo compacto arbitrário $[a,\ b] \subset I$ são assumidos em um dos extremos do mesmo.

Vamos agora supor a recíproca, ou seja, que os máximos da função ϕ_α em um intervalo compacto arbitrário $[a,\ b] \subset I$ são assumidos em um dos extremos do mesmo. Assim, temos $f(x) + \alpha x \leq M$ para todo $x \in [a,\ b]$, onde $M := \max\{f(a) + \alpha a,\ f(b) + \alpha b\}$. Escrevamos $x \in [a,\ b]$ na forma $x = \lambda a + (1-\lambda)b$ com $\lambda \in [a,\ b]$. Temos, portanto, $f\big(\lambda a + (1-\lambda)b\big) + \alpha\big(\lambda a + (1-\lambda)b\big) \leq M$, ou seja,

$$f\big(\lambda a + (1-\lambda)b\big) \ \leq \ M - \alpha\big(\lambda a + (1-\lambda)b\big) \ .$$

[4] Johan Ludwig William Valdemar Jensen (1859–1925). O trabalho original de Jensen é: J. L. W. V. Jensen, *"Sur les fonctions convexes et les inégalités entre les valeurs moyennes"*, Acta Math. **30**, 175–193 (1906).

Subtraindo $\lambda f(a)+(1-\lambda)f(b)$ de ambos os lados dessa desigualdade, teremos

$$f\big(\lambda a+(1-\lambda)b\big)-\lambda f(a)-(1-\lambda)f(b) \ \leq \ M-\lambda\big(f(a)+\alpha a\big)-(1-\lambda)\big(f(b)+\alpha b\big) \ .$$

Observe-se agora que essa desigualdade deve, por hipótese, ser verdadeira para todo $\alpha \in \mathbb{R}$ e que seu lado esquerdo independe de α. Se escolhermos $\alpha = \big(f(a)-f(b)\big)/(b-a)$, teremos $f(a)+\alpha a = f(b)+\alpha b = M$. Nesse caso, o lado direito da desigualdade anula-se e concluímos que $f\big(\lambda a+(1-\lambda)b\big) \leq \lambda f(a)+(1-\lambda)f(b)$. Como a e b são arbitrários em I (com $a<b$) e $\lambda \in [0,\ 1]$ é igualmente arbitrário, isso estabeleceu que f é convexa em todo I. ■

A proposição que segue, também mencionada em [137], apresenta mais uma condição equivalente à convexidade para funções contínuas.

Proposição 5.6 *Seja $f: I \to \mathbb{R}$ uma função contínua. Então, f é convexa se e somente se satisfizer*

$$f(x) \ \leq \ \frac{1}{2h}\int_{x-h}^{x+h} f(t)\,dt \ . \tag{5.20}$$

para todo $x \in I^0$ e todo $h>0$ tal que $x \pm h \in I$. □

Comentário. A Proposição 5.6 está na raíz da definição das chamadas *funções sub-harmônicas*, tema do qual não trataremos aqui. ♣

Prova da Proposição 5.6. Se f é contínua e convexa então, por (5.15), vale $f(x) \leq \frac{1}{2}\big(f(x+t)+f(x-t)\big)$ para todos $x \in I^0$ e $x \pm t \in I$. Integrando-se ambos os lados t na variável t com t no intervalo $[0,\ h]$, teremos $hf(x) \leq \frac{1}{2}\int_{-h}^{h} f(x+t)\,dt$, como facilmente se vê, e isso equivale a (5.20).

Vamos agora supor que f seja contínua e satisfaça (5.20). Seja $\alpha \in \mathbb{R}$. É elementar constatar que para todo $h>0$ vale $\alpha x = \frac{1}{2h}\int_{x-h}^{x+h}\alpha t\,dt$. Logo,

$$\phi_\alpha(x) \ \leq \ \frac{1}{2h}\int_{x-h}^{x+h}\phi_\alpha(t)\,dt \ , \tag{5.21}$$

com $\phi_\alpha(x) := f(x)+\alpha x$. Seja $[a,\ b] \subset I$ um intervalo compacto e seja um intervalo $[x_0-h,\ x_0+h] \subset [a,\ b]$ com as seguintes propriedades: 1º $\phi_\alpha(y) \leq \phi_\alpha(x_0)$ para todo $y \in [x_0-h,\ x_0+h]$ e 2º existe $y_0 \in [x_0-h,\ x_0+h]$ tal que $\phi_\alpha(y_0) < \phi_\alpha(x_0)$. Como ϕ_α é contínua, teremos $\frac{1}{2h}\int_{x_0-h}^{x_0+h}\phi_\alpha(t)\,dt < \phi_\alpha(x_0)$. Isso contraria (5.21) e, portanto, devemos ter $\phi_\alpha(y) \geq \phi_\alpha(x_0)$ para todo $y \in [x_0-h,\ x_0+h]$. Como isso vale para todo intervalo $[x_0-h,\ x_0+h] \subset [a,\ b]$, segue que $\max\{\phi_\alpha(x),\ x \in [a,\ b]\} = \max\{\phi_\alpha(a),\ \phi_\alpha(b)\}$. Pela Proposição 5.5, página 255, isso implica que f é convexa. ■

• A desigualdade de Hermite-Hadamard

Proposição 5.7 *Seja $f: I \to \mathbb{R}$ uma função convexa. Então, f satisfaz*

$$f\left(\frac{a+b}{2}\right) \ \leq \ \frac{1}{b-a}\int_a^b f(t)\,dt \ \leq \ \frac{f(a)+f(b)}{2} \tag{5.22}$$

para todos $a,\ b \in I^0$ com $a<b$. □

A desigualdade (5.22) é conhecida como *desigualdade de Hermite*[5]*-Hadamard*[6], ou simplesmente como *desigualdade de Hadamard*.

Prova da Proposição 5.7. Tomando $h=(b-a)/2$ e $x=(a+b)/2$ em (5.20), obtemos a primeira desigualdade $f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(t)\,dt$. Com a mudança de variáveis $t=\lambda(b-a)+a=\lambda b+(1-\lambda)a$ para $\lambda \in [0,\ 1]$, temos, pela convexidade de f,

$$\frac{1}{b-a}\int_a^b f(t)\,dt \ = \ \int_0^1 f\big(\lambda b+(1-\lambda)a\big)\,d\lambda \ \stackrel{\text{convex.}}{\leq} \ \int_0^1 \Big[f(b)\lambda+f(a)(1-\lambda)\Big]\,d\lambda \ = \ \frac{f(a)+f(b)}{2} \ ,$$

completando a prova. ■

[5] Charles Hermite (1822–1901).
[6] Jacques Salomon Hadamard (1865–1963).

5.1.2 Funções Convexas de Várias Variáveis

• **Funções convexas em $\mathbb{R}^n$**

No que segue, consideraremos conjuntos convexos $C \subset \mathbb{R}^n$ que sejam também de dimensão n, isto é, que sejam tais que o menor subespaço que contém C seja $\mathbb{R}^n$.

Seja $C \subset \mathbb{R}^n$ um conjunto convexo. Uma função $f : C \to \mathbb{R}$ é dita ser uma *função convexa* se para todos $x, y \in C$ e todo $\lambda \in [0, 1]$ valer a desigualdade

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y) . \tag{5.23}$$

Uma função $f : C \to \mathbb{R}$ é dita ser uma *função côncava* se para todos $x, y \in C$ e todo $\lambda \in [0, 1]$ valer a desigualdade

$$f\big(\lambda x + (1-\lambda)y\big) \geq \lambda f(x) + (1-\lambda)f(y) . \tag{5.24}$$

É elementar constatar que f é côncava se e somente se $-f$ for convexa, sendo, portanto, suficiente estudar propriedades gerais de funções convexas.

Mais adiante, no Corolário 5.4, página 259, provaremos que se $f : C \to \mathbb{R}$ é convexa ou côncava, então f é contínua em C^0, o interior de C.

Se $f : C \to \mathbb{R}$ é uma função convexa, então ela é uma função convexa de cada uma de suas variáveis individualmente: tomemos, por exemplo, $x, y \in C$ na forma $x \equiv (x_1, x_2, \ldots, x_n)$ e $y \equiv (y_1, x_2, \ldots, x_n)$. Então, a convexidade de f implica que $f\big(\lambda x_1 + (1-\lambda)y_1, x_2, \ldots, x_n\big) \leq \lambda f\big(x_1, x_2, \ldots, x_n\big) + (1-\lambda)f\big(y_1, x_2, \ldots, x_n\big)$, que corresponde à afirmação que f é convexa enquanto função de sua primeira variável. Para as demais variáveis individualmente tem-se o mesmo.

A recíproca, porém, não é necessariamente verdadeira: se f é convexa em cada uma de suas variáveis independentemente, ela não é necessariamente convexa enquanto função de suas n variáveis. Como contraexemplo, tome-se a função $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x_1, x_2) := x_1 x_2$. É fácil ver que f é convexa como função de x_1 e de x_2 separadamente, mas ela não pode satizfazer (5.23) se, por exemplo, tomarmos $x_1 > 0$, $x_2 = 0$, $y_1 = 0$, $y_2 > 0$ e $\lambda \in (0, 1)$. Verifique!

O exercício a seguir mostra uma maneira de se obter uma função convexa em $n+1$ variáveis a partir de uma função convexa de n variáveis.

E. 5.6 *Exercício*. Seja C um conjunto convexo em $\mathbb{R}^n$ e seja $f : C \to \mathbb{R}$ uma função convexa de n variáveis em C. Seja $g(x, t) := tf(t^{-1}x)$ uma função de $n+1$ variáves definida em $D := \big\{(x, t) \in \mathbb{R}^{n+1} \big| \, t > 0 \text{ e } t^{-1}x \in C\big\} \subset \mathbb{R}^{n+1}$. Mostre que D é um conjunto convexo em $\mathbb{R}^{n+1}$ (em verdade, um cone convexo em $\mathbb{R}^{n+1}$) e mostre que g é uma função convexa de $n+1$ variáveis em D, ou seja, mostre que

$$g\Big(\lambda(x, t) + (1-\lambda)(x', t')\Big) \equiv g\Big(\lambda x + (1-\lambda)x', \lambda t + (1-\lambda)t'\Big) \leq \lambda g(x, t) + (1-\lambda)g(x', t')$$

para todos (x, t), $(x', t') \in D$ e $\lambda \in [0, 1]$. *Sugestão:* Use o fato que

$$\frac{1}{\lambda t + (1-\lambda)t'}\Big(\lambda x + (1-\lambda)x'\Big) = \left(\frac{\lambda t}{\lambda t + (1-\lambda)t'}\right)\frac{1}{t}\,x + \left(\frac{(1-\lambda)t'}{\lambda t + (1-\lambda)t'}\right)\frac{1}{t'}\,x' ,$$

note que a soma dos termos entre parênteses do lado direito vale 1 e use a convexidade de C e de f. ✠

E. 5.7 *Exercício*. Mostre que toda semi-norma em um espaço vetorial real é uma função convexa sobre o mesmo. ✠

No caso em que $f : C \to \mathbb{R}$ é duas vezes diferenciável, a proposição que segue fornece condições necessárias e suficientes à convexidade.

Proposição 5.8 *Seja $C \subset \mathbb{R}^n$ um conjunto convexo e seja $f : C \to \mathbb{R}$ uma função contínua e duas vezes diferenciável no interior C^0 de C. Então, f é convexa se e somente se*

$$\sum_{i, j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\, a_i a_j \geq 0 \tag{5.25}$$

para todo $x \in C^0$ e para todo $(a_1, \ldots, a_j) \in \mathbb{R}^n$. □

Nota. A condição (5.25) é satisfeita se e somente se todos os auto-valores da matriz Hessiana[7] $H_{ij}(x) := \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$, $i, j \in \{1, \ldots, n\}$, forem não-negativos. ♣

Prova da Proposição 5.8. Para $x, y \in C^0$ defina-se a função de uma variável

$$g_{x,\,y}(t) \;:=\; f\big(x + t(y-x)\big), \qquad t \in [0,\,1]\;.$$

Para $t_1, t_2 \in [0, 1]$ temos

$$\begin{aligned} g_{x,\,y}\big(\lambda t_1 + (1-\lambda)t_2\big) &= f\Big(\lambda\big(x + t_1(y-x)\big) + (1-\lambda)\big(x + t_2(y-x)\big)\Big)\;, \\ g_{x,\,y}(t_1) &= f\big(x + t_1(y-x)\big)\;, \\ g_{x,\,y}(t_2) &= f\big(x + t_2(y-x)\big)\;, \end{aligned}$$

para todo $\lambda \in [0, 1]$. Dessas três relações prova-se facilmente que f é convexa em C se e somente todas as funções $g_{x,\,y}$, $x, y \in C^0$, forem funções convexas de uma variável.

A função g é duas vezes diferenciável, pois f o é. Assim, pelo Corolário 5.2, página 252, f é convexa se e somente se para todos $x, y \in C^0$ valer $g''_{x,\,y} \geq 0$. Agora, pela regra da cadeia,

$$g''_{x,\,y}(t) \;=\; \sum_{i,\,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}\big(x + t(y-x)\big)\;(y_i - x_i)(y_j - x_j)\;.$$

Logo, teremos $g''_{x,\,y} \geq 0$ para todos $x, y \in C^0$ se e somente se

$$\sum_{i,\,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\; a_i a_j \;\geq\; 0$$

para todo $x \in C^0$ e para todo $(a_1, \ldots, a_j) \in \mathbb{R}^n$. ■

• Funções convexas em espaços vetoriais normados e sua continuidade

Vamos agora considerar uma importante generalização da nossa discussão para espaços vetoriais reais normados. A leitura do que segue requer noções de topologia de espaços métricos, como espaços vetoriais normados, temas tratados em capítulos posteriores. O principal resultado é:

Proposição 5.9 *Seja $\mathcal{V}$ um espaço vetorial real dotado de uma norma $\|\cdot\|$, seja $C \subset \mathcal{V}$ convexo, com interior C^0 não-vazio, e seja $f : C \to \mathbb{R}$ uma função convexa em C. Se f for limitada superiormente em C, ou seja, se possuir um majorante superior finito $\mathsf{S} := \sup\{f(x),\ x \in C\} < \infty$, então f é contínua em C^0, o interior de C.*

Se $f : C \to \mathbb{R}$ côncava em C e limitada inferiormente em C, ou seja, tal que $\mathsf{I} := \inf\{f(x),\ x \in C\} > -\infty$, então, f é contínua em C^0, o interior de C. □

Prova. Provaremos apenas a afirmação para funções convexas, pois a afirmação para funções côncavas decorre imediatamente da mesma (trocando-se f por $-f$).

Vamos supor que existe $z \in C^0$ tal que f não seja contínua em z. Então, existe $\epsilon > 0$ tal que para cada $\delta > 0$ é possível encontrar um $x \in \mathcal{V}$ tal que $\|z - x\| < \delta$ mas com $|f(z) - f(x)| > \epsilon$ (contrariamente, f seria contínua em z).

Como $z \in C^0$ e C^0 é um conjunto aberto, existe $r > 0$ tal que a bola aberta de raio r centrada em z, $B(r, z) := \big\{y \in \mathcal{V} \big|\ \|y - z\| < r\big\}$, está inteiramente contida em C^0: $B(r, z) \subset C^0$.

Vamos escolher $\lambda \in (0, 1]$ tal que

$$\frac{1-\lambda}{\lambda} \;>\; \frac{\mathsf{S} - f(z)}{\epsilon}\;. \tag{5.26}$$

[7] Ludwig Otto Hesse (1811–1874).

Que uma tal escolha sempre é possível segue do fato que a imagem da função $(0,\ 1] \ni \lambda \mapsto \lambda^{-1} - 1$ é $[0,\ \infty)$, como se vê facilmente. Segue de (5.26) que

$$\frac{\epsilon}{\lambda} \ > \ \mathsf{S} - f(z) + \epsilon \ , \tag{5.27}$$

relação que usaremos adiante.

Tomemos $\delta = r\lambda$. Então, pelo que afirmamos acima, existe $x \in \mathcal{V}$ tal que $\|x - z\| < \delta = r\lambda$ mas $|f(z) - f(x)| > \epsilon$. Note-se que, como $\lambda \in (0,\ 1]$, tem-se que $\delta \leq r$ e, portanto, $x \in B(r,\ z) \subset C^0$. Como $|f(z) - f(x)| > \epsilon$, há duas situações possíveis: *situação a:* $f(x) > f(z) + \epsilon$ e *situação b:* $f(x) < f(z) - \epsilon$.

Situação a: $f(x) > f(z) + \epsilon$. Seja $y \in \mathcal{V}$ dado por

$$y \ := \ \frac{1}{\lambda}x - \frac{1-\lambda}{\lambda}z \ .$$

Teremos $\|y - z\| = \frac{1}{\lambda}\|x - z\| < \frac{\delta}{\lambda} = r$, de modo que $y \in B(r,\ z) \subset C^0$. Porém, temos $x = \lambda y + (1 - \lambda)z$ e como f é convexa, temos $f(x) \leq \lambda f(y) + (1 - \lambda)f(z)$. Logo,

$$f(y) \ \geq \ \frac{1}{\lambda}f(x) - \frac{1-\lambda}{\lambda}f(z) \ > \ \frac{1}{\lambda}(f(z)+\epsilon) - \frac{1-\lambda}{\lambda}f(z) \ = \ f(z) + \frac{\epsilon}{\lambda} \stackrel{(5.27)}{>} \mathsf{S} + \epsilon \ .$$

Assim, obtivemos, $f(y) > \mathsf{S} + \epsilon > \mathsf{S}$, o que contraria a definição de S, mostrando que a situação a é impossível.

Situação b: $f(x) < f(z) - \epsilon$. Seja $y \in \mathcal{V}$ dado por

$$y \ := \ \frac{1}{\lambda}z - \frac{1-\lambda}{\lambda}x \ .$$

Teremos $\|y - z\| = \frac{1-\lambda}{\lambda}\|x - z\| < (1-\lambda)\frac{\delta}{\lambda} = (1-\lambda)r < r$, de modo que $y \in B(r,\ z) \subset C^0$. Porém, temos $z = \lambda y + (1-\lambda)x$ e como f é convexa, temos $f(z) \leq \lambda f(y) + (1 - \lambda)f(x)$. Logo,

$$f(y) \ \geq \ \frac{1}{\lambda}f(z) - \frac{1-\lambda}{\lambda}f(x) \ > \ \frac{1}{\lambda}f(z) - \frac{1-\lambda}{\lambda}(f(z)-\epsilon) \ = \ f(z) + \epsilon\frac{1-\lambda}{\lambda} \stackrel{(5.26)}{>} \mathsf{S} \ .$$

Assim, obtivemos, $f(y) > \mathsf{S}$, o que novamente contraria a definição de S, mostrando que a situação b também é impossível.

A resolução dessas contradições é que um tal ponto $z \in C^0$ onde f é descontínua não pode existir. ■

• Comentários à Proposição 5.9

Se $\mathcal{V}$ for um espaço vetorial real normado e $f : \mathcal{V} \to \mathbb{R}$ for um funcional linear em $\mathcal{V}$, então f é uma função convexa (e côncava) no conjunto convexo $\mathcal{V}$. Se $\mathcal{V}$ não for um espaço de dimensão finita, linearidade não necessariamente faz de f uma função contínua. Há exemplos bem conhecidos de funcionais lineares descontínuos de um espaço normado em $\mathbb{R}$ ou $\mathbb{C}$, como o funcional delta de Dirac discutido à página 2056. Assim, podem existir funções convexas ou côncavas não-continuas em espaços normados de dimensão infinita[8] e aí reside a relevância de resultados como os da Proposição 5.9.

É de se notar também que a condição de limitação superior (inferior) listada na Proposição 5.9 é suficiente, mas não é necessária para que uma função convexa (côncava) seja contínua no interior do seu domínio convexo de definição. Segundo a Proposição 41.1, página 2055, se $\mathcal{V}$ é um espaço vetorial normado, um funcional linear $f : \mathcal{V} \to \mathbb{R}$ é contínuo se e somente se $\sup\{|f(x)|,\ x \in \mathcal{V},\ \|x\| = 1\} < \infty$. Funcionais lineares $f : \mathcal{V} \to \mathbb{R}$ que satisfaçam essa condição de continuidade não são limitados nem superior nem inferiormente em $\mathcal{V}$ (segundo a definição que usamos acima).

• Continuidade no caso de dimensão finita

No caso em que $\mathcal{V}$ é um espaço vetorial real normado de dimensão finita (como $\mathbb{R}^n$), a condição de uma função $f : \mathcal{V} \to \mathbb{R}$ convexa (côncava) ser limitada superiormente (inferiormente) é dispensável para a continuidade:

Corolário 5.4 *Seja $C \subset \mathbb{R}^n$, convexo, e seja $f : C \to \mathbb{R}$ uma função convexa ou côncava em C. Então, f é contínua em C^0, o interior de C.* □

[8] Todo o Capítulo 42, página 2239, é dedicado a operadores lineares não-contínuos.

Prova. Consideramos apenas o caso em que f é convexa, pois o outro é análogo. Seja $a \in C^0$. Então, é possível encontrar $r > 0$ tal que $B(r, \ a)$, a bola aberta de raio r centrada em a, está inteiramente contida em C^0. Dentro dessa bola é possível encontrar um conjunto finito $v_1, \ \ldots, \ v_{n+1}$ de pontos tais que o conjunto $\mathcal{C} := \{\lambda_1 v_1 + \cdots + \lambda_{n+1} v_{n+1}, \ \lambda_1, \ \ldots \ \lambda_{n+1} \in [0, \ 1]$ com $\lambda_1 + \cdots + \lambda_{n+1} = 1\}$ é uma vizinhança convexa de a. Pela convexidade de f, temos para todo ponto $x = \lambda_1 v_1 + \cdots + \lambda_{n+1} v_{n+1}$ de $\mathcal{C}$ que $f(x) \leq \lambda_1 f(v_1) + \cdots + \lambda_{n+1} f(v_{n+1}) \leq \max\{f(v_1), \ \ldots, \ f(v_{n+1})\}$. Logo, f é limitada superiormente no convexo $\mathcal{C}$ e, pela Proposição 5.9, página 258, f é contínua em $\mathcal{C}$ e, em particular em a. Como a é um ponto arbitrário de C^0 a demonstração está completa. ■

5.2 Algumas Consequências da Convexidade e da Convavidade

Da propriedade de convexidade ou concavidade de funções é possível obter desigualdades muito úteis das quais faremos uso em outros momentos nestas Notas. Nesta Seção apresentaremos e demonstramos algumas delas, como a importate desigualdade de Jensen, a (primeira) desigualdade de Young e algumas outras desigualdades decorrentes da concavidade da função logaritmo ou da convexidade da função exponencial. Todas essas desigualdades são relevantes e possuem aplicações diversas da Mecânica Estatística e Termodinâmica à Análise Funcional.

5.2.1 A Desigualdade de Jensen

Vamos agora apresentar uma importante generalização da desigualdade (5.5). Por simplicidade, consideraremos aqui funções convexas $f : \mathbb{R} \to \mathbb{R}$ definidas em todo $\mathbb{R}$, mas é fácil perceber que é possível também considerar funções convexas definidas em intervalos menores, desde que as devidas restrições sejam feitas às demais funções envolvidas.

Proposição 5.10 (Desigualdade de Jensen) *Seja $f : \mathbb{R} \to \mathbb{R}$ convexa. Seja $p : \mathbb{R} \to [0, \ \infty)$ uma função positiva e integrável, ou seja, tal que $P := \int_{-\infty}^{\infty} p(t)\, dt < \infty$. Seja $g : \mathbb{R} \to \mathbb{R}$ uma função integrável na medida $p(t)dt$, ou seja, tal que $\int_{-\infty}^{\infty} |g(t)|\, p(t)\, dt < \infty$. Então, vale*

$$P\, f\left(\int_{-\infty}^{\infty} g(t)\, p(t)\, dt\right) \ \leq \ \int_{-\infty}^{\infty} f\big(P\, g(t)\big)\, p(t)\ dt\ . \tag{5.28}$$

Essa desigualdade é denominada desigualdade de Jensen. *Um caso particular importante é aquele no qual a função p representa uma distribuição de probabilidades e temos $P = 1$. Nele, a desigualdade de Jensen assume a forma*

$$f\left(\int_{-\infty}^{\infty} g(t)\, p(t)\, dt\right) \ \leq \ \int_{-\infty}^{\infty} f\big(g(t)\big)\, p(t)\ dt\ . \tag{5.29}$$

□

O estudante deve perceber que (5.5) é um caso particular de (5.29). O estudante mais avançado deve também perceber ao longo da demonstração que (5.28) permacece válida se $p(t)dt$ for substituida por qualquer medida positiva e finita $d\mu$ em $\mathbb{R}$ (com $P := \int_{\mathbb{R}} d\mu$) e que (5.29) permanece válida se $p(t)dt$ for substituida por qualquer medida de probabilidade $d\mu$ em $\mathbb{R}$.

Prova da Proposição 5.10. A desigualdade (5.12), página 252, afirma que, para todo $y \in \mathbb{R}$ podemos escrever

$$f(y) \ \geq \ \alpha_x y + \beta_x\ ,$$

onde $\alpha_x := f'_{\pm}(x)$ e $\beta_x := f(x) - x f'_{\pm}(x)$, sendo $x \in \mathbb{R}$, arbitrário. Tomando-se $y \equiv Pg(t)$, temos

$$f\big(Pg(t)\big) \ \geq \ \alpha_x P g(t) + \beta_x\ .$$

Logo, como p é não-negativa, temos também

$$f\big(Pg(t)\big)p(t) \ \geq \ \alpha_x P g(t) p(t) + \beta_x p(t)\ .$$

Integrando-se, obtemos

$$\int_{-\infty}^{\infty} f\big(Pg(t)\big)\, p(t)\, dt \;\geq\; \alpha_x P \int_{-\infty}^{\infty} g(t)\, p(t)\, dt + \beta_x P \;=\; f'_{\pm}(x) P \left(\int_{-\infty}^{\infty} g(t)\, p(t)\, dt\right) + Pf(x) - Pxf'_{\pm}(x)\;,$$

com $x \in \mathbb{R}$, arbitrário. Tomando-se e $x \equiv \int_{-\infty}^{\infty} g(t)p(t)\, dt$, o primeiro e o terceiro termo do lado direito cancelam-se e obtemos, finalmente,

$$\int_{-\infty}^{\infty} f\big(Pg(t)\big)\, p(t)\, dt \;\geq\; Pf\left(\int_{-\infty}^{\infty} g(t)\, p(t)\, dt\right)\;,$$

que é a desigualdade de Jensen (5.28). ■

5.2.2 A Primeira Desigualdade de Young

Nosso resultado na corrente seção será uma desigualdade útil que denominamos *primeira desigualdade de Young*[9].

Proposição 5.11 (Primeira Desigualdade de Young) *Sejam dois intervalos $[0,\ \alpha)$ e $[0,\ \beta)$, onde $0 < \alpha \leq \infty$ e $0 < \beta \leq \infty$ e seja $F : [0,\ \alpha) \to [0,\ \beta)$ uma função contínua, não-negativa, crescente, bijetora e satisfazendo $F(0) = 0$. Então, para todos $x \in [0,\ \alpha)$ e $x \in [0,\ \beta)$ vale*

$$xy \;\leq\; \int_0^x F(s)\, ds + \int_0^y F^{-1}(t)\, dt\;. \tag{5.30}$$

Essa desigualdade é denominada aqui "primeira desigualdade de Young". □

A desigualdade (5.30) tem uma interpretação geométrica muito simples, que descrevemos na Figura 5.3, página 262. Para generalizações da desigualdade (5.30), vide e.g. [137].

Prova da Proposição 5.11. Como $F : [0,\ \alpha) \to [0,\ \beta)$ uma função contínua, não-negativa, crescente, bijetora e satisfaz $F^{-1}(0) = 0$, sua inversa $F^{-1} : [0,\ \beta) \to [0,\ \alpha)$ é igualmenete contínua, não-negativa, crescente, bijetora e satisfaz $F^{-1}(0) = 0$.

Defina-se, para $x \in [0,\ \alpha)$ e $y \in [0,\ \beta)$ as funções

$$A(x) \;:=\; \int_0^x F(s)\, ds \qquad \text{e} \qquad B(y) \;:=\; \int_0^y F^{-1}(t)\, dt\;.$$

A e B são contínuas e diferenciáveis (com $A'(x) = F(x)$ e $B'(y) = F^{-1}(y)$) e, pelo Corolário 5.3, página 253, ambas são convexas, já que F e F^{-1} são crescentes.

Afirmamos que vale a identidade $B(y) = C(y)$ para todo $y \in [0,\ \beta)$, onde

$$C(y) \;:=\; \int_0^{F^{-1}(y)} \Big(y - F(s)\Big)\, ds\;.$$

De fato, C é contínua e diferenciável (pois F^{-1} é contínua) e tem-se[10]

$$C'(y) \;=\; y - F\big(F^{-1}(y)\big) + \int_0^{F^{-1}(y)} \frac{\partial}{\partial y}\Big(y - F(s)\Big)\, ds \;=\; \int_0^{F^{-1}(y)} ds \;=\; F^{-1}(y) \;=\; B'(y)\;,$$

e como $B(0) = 0$ e $C(0) = 0$ (pois $F^{-1}(0) = 0$), segue que $B(y) = C(y)$ para todo $y \in [0,\ \beta)$.

[9] William Henry Young (1863–1942).

[10] Aqui, usa-se a identidade (prove-a!)

$$\frac{d}{dy}\int_0^{h(y)} J(y,\ s)\, ds \;=\; J\big(y,\ h(y)\big) + \int_0^{h(y)} \frac{\partial J}{\partial y}(y,\ s)\, ds\;,$$

válida para $h(y)$ contínua e $J(y,\ s)$ contínua e diferenciável, com $\frac{\partial}{\partial y}J(y,\ s)$ contínua.

Agora, a identidade $B(y) = C(y)$ permite escrever

$$B\big(F(x)\big) \;=\; C\big(F(x)\big) \;=\; \int_0^{F^{-1}(F(x))} \Big(F(x) - F(s)\Big)\, ds \;=\; \int_0^x \Big(F(x) - F(s)\Big)\, ds \;=\; xF(x) - A(x)\ .$$

Estabelecemos, portanto, que

$$xF(x) \;=\; A(x) + B\big(F(x)\big) \tag{5.31}$$

para todo $x \in [0,\ \alpha)$, relação que logo usaremos.

Da convexidade estrita de B e do Corolário 5.1, página 252, segue que

$$B(y) \;\geq\; (y - y')F^{-1}(y') + B(y')$$

para todos $y,\ y' \in [0,\ \beta)$. Tomando-se $y' = F(x)$, teremos

$$B(y) \;>\; \big(y - F(x)\big)x + B\big(F(x)\big) \;\stackrel{(5.31)}{=}\; xy - B\big(F(x)\big) - A(x) + B\big(F(x)\big)$$

e, portanto, provamos que

$$xy \;\leq\; A(x) + B(y)$$

para todos $x \in [0,\ \alpha)$ e $y \in [0,\ \beta)$. ■

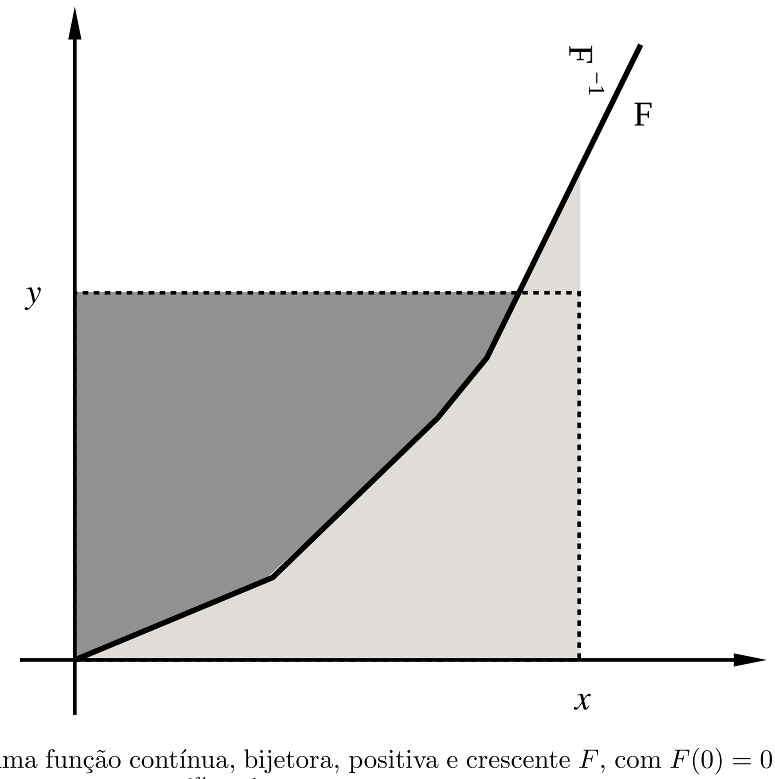

Figura 5.3: O gráfico de uma função contínua, bijetora, positiva e crescente F, com $F(0) = 0$. A área em cinza claro vale $\int_0^x F(s)\, ds$. A área em cinza escuro vale $\int_0^y F^{-1}(t)\, dt$. O retângulo de lados x e y é representado em linhas tracejadas e sua área xy é claramente menor ou igual à soma das duas áreas acinzentadas, que vale $\int_0^x F(s)\, ds + \int_0^y F^{-1}(t)\, dt$. Essa é a interpretação geométrica da Primeira Desigualdade de Young (5.30).

E. 5.8 *Exercício*. Tomando-se $F(x) = e^x - 1$ em (5.30), obtenha a desigualdade

$$(1 + x)(1 + y) \;\leq\; e^x + (1 + y)\ln(1 + y)\ , \qquad \forall x,\ y \geq 0\ .$$

Demonstre alternativamente sua validade estudando os mínimos da função $e^x + (1+y)\ln(1+y) - (1+x)(1+y)$ pelo método convencional. ✠

Na desigualdade (5.30), o lado esquerdo é uma função simétrica de x e y, mas o lado direito não é. Isso pode ser remediado simetrizando-se o lado direito no intervalo comum a F e F^{-1}, a saber, $[0, \gamma)$, onde $\gamma = \min\{\alpha, \beta\}$. Obtem-se, então, nesse intervalo

$$xy \leq \frac{1}{2}\left(\int_0^x F(s)\, ds + \int_0^y F(s)\, ds\right) + \frac{1}{2}\left(\int_0^x F^{-1}(t)\, dt + \int_0^y F^{-1}(t)\, dt\right) .$$

5.2.3 Médias Geométricas, Aritméticas e Desigualdades Correlatas

• **A desigualdade entre média geométrica e média aritmética**

A concavidade da função $\ln x$ tem algumas consequências relevantes. Para $n \in \mathbb{N}$, sejam $x_1, \ldots, x_n$ números positivos e sejam $\lambda_1, \ldots, \lambda_n \in [0, 1]$ tais que $\lambda_1 + \cdots + \lambda_n = 1$. Como $\ln x$ é côncava, (5.6) garante-nos que

$$\ln(\lambda_1 x_1 + \cdots + \lambda_n x_n) \geq \lambda_1 \ln(x_1) + \cdots + \lambda_n \ln(x_n) = \ln\left(x_1^{\lambda_1} \cdots x_n^{\lambda_n}\right) .$$

Tomando-se a exponencial dessa desigualdade, obtemos

$$x_1^{\lambda_1} \cdots x_n^{\lambda_n} \leq \lambda_1 x_1 + \cdots + \lambda_n x_n . \tag{5.32}$$

Note-se que se $\lambda_1, \ldots, \lambda_n \in (0, 1)$ essa desigualdade é válida mesmo que alguns x_k's sejam nulos.

A expressão $\lambda_1 x_1 + \cdots + \lambda_n x_n$ é denominada *média aritmética ponderada* do conjunto de números não-negativos $\{x_1, \ldots, x_n\}$ (ponderada pelos fatores $\lambda_1, \ldots, \lambda_n$) e a expressão $x_1^{\lambda_1} \cdots x_n^{\lambda_n}$ é denominada *média geométrica ponderada* do conjunto $\{x_1, \ldots, x_n\}$. A desigualdade (5.32) afirma, portanto, que a média geométrica ponderada de um conjunto finito de números não-negativos é sempre menor que a sua média aritmética ponderada. Em [137] o leitor poderá encontrar diversas outras demonstrações da desigualdade (5.32).

Sejam $x_1, \ldots, x_n$ números não-negativos. Definimos suas médias geométrica e aritmética (simples) por $\left(x_1 \cdots x_n\right)^{\frac{1}{n}}$ e $\frac{x_1+\cdots+x_n}{n}$, respectivamente. Elas correspondem ao caso em que $\lambda_1 = \cdots = \lambda_n = \frac{1}{n}$. Temos de (5.32), portanto,

$$\left(x_1 \cdots x_n\right)^{\frac{1}{n}} \leq \frac{x_1 + \cdots + x_n}{n} , \tag{5.33}$$

ou seja, a média geométrica de uma coleção finita de números não-negativos é sempre menor ou igual à média aritmética da mesma coleção. Esse resultado é originalmente atribuido a Cauchy[11] e pode ser provado de diversas formas. Cauchy obteve-o, não usando a concavidade do logaritmo, como fizemos, mas por indução (para uma tal prova, vide e.g. [137]).

• **A desigualdade de Young**

Na demonstração da chamada desigualdade de Hölder em espaços L^p (assunto discutido nas Seções 27.5.1 e 33.4.1, páginas 1339 e 1532, respectivamente) faz-se uso de uma desigualdade elementar conhecida como *desigualdade de Young*[12]. Como a desigualdade de Young tem interesse por si só e algumas outras aplicações, vamos apresentar sua demonstração.

Como já discutimos, a função logaritmo é côncava como função do intervalo $(0, \infty)$ sobre $\mathbb{R}$. Assim, para todos $a, b \in (0, \infty)$ e todo $\lambda \in [0, 1]$ tem-se $\ln\left(\lambda a + (1-\lambda)b\right) \geq \lambda \ln a + (1-\lambda)\ln b = \ln\left(a^\lambda b^{1-\lambda}\right)$. Tomando-se a exponencial de ambos os lados, obtemos

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b . \tag{5.34}$$

Note-se que se $\lambda \in (0, 1)$, então (5.34) é também válida caso $a = 0$ e/ou $b = 0$. A desigualdade (5.34) é por vezes denominada desigualdade de Young. Como se vê, trata-se meramente de um caso particular de (5.32) (para $n = 2$).

Por vezes a desigualdade de Young é apresentada de forma ligeiramente diferente. Sejam p e q ambos tais que $1 < p < \infty$ e $1 < q < \infty$, mas tais que $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Então, por (5.34), temos para $a, b \in [0, \infty)$.

$$a^{1/p}\, b^{1/q} \leq \frac{a}{p} + \frac{b}{q} , \tag{5.35}$$

[11] Augustin Louis Cauchy (1789–1857).
[12] William Henry Young (1863–1942).

sendo que a igualdade só é válida caso $a = b$ (isso segue do fato de o logaritmo ser estritamente côncavo. Vide, porém, nota logo abaixo).

A desigualdade dita de Young (5.35) é bastante elementar, mas ela foi originalmente provada a partir do resultado mais geral e não-trivial expresso na *primeira desigualdade de Young*, a desigualdade (5.30) da Proposição 5.11, página 261.

E. 5.9 *Exercício.* Tomando $F(s) = s^{p-1}$ com $1 < p < \infty$, para $s \geq 0$, obtenha da primeira desigualdade de Young (5.30) a desigualdade (5.35). ✠

Há ainda uma terceira desigualdade para produtos de convolução também denominada "desigualdade de Young" e que é vagamente relacionada às desigualdades (5.35) e (5.30). Vide, e.g., [216].

Nota. Por completeza, apresentemos uma segunda demonstração de (5.35) sem uso da concavidade. Notemos, em primeiro lugar, que se $a = 0$ ou $b = 0$ a (5.35) acima é trivialmente satisfeita, pois o lado esquerdo é sempre zero, enquanto que o lado direito é sempre maior ou igual a zero. Vamos, então, supor que a e b sejam ambos não-nulos. Tudo o que queremos é provar que $-a^{1/p}b^{1/q} + \frac{a}{p} + \frac{b}{q}$ é sempre maior ou igual a zero. Podemos escrever a última expressão como $b\left(-t^{\alpha} + \alpha t + \frac{1}{q}\right)$, onde $\alpha = 1/p$ e $t = a/b$. Como $1 < p < \infty$, temos que $0 < \alpha < 1$ enquanto que $t \geq 0$. Note-se que a função

$$f(x) \;=\; -x^{\alpha} + \alpha x + \frac{1}{q} \;,$$

é contínua para $x \in [0, \infty)$ e que, para $x > 0$, tem-se $f'(x) = \alpha\left(1 - x^{\alpha-1}\right)$ e $f''(x) = \alpha(1-\alpha)x^{\alpha-2} > 0$. Assim, $f(x)$ tem um único mínimo local em $x = 1$, onde $f(1) = 0$ (verifique). Fora isso, $f(0) = \frac{1}{q} > 0$ e $\lim_{x\to\infty} f(x) = +\infty$. Desses fatos concluímos facilmente que $f(x) \geq 0$ para todo $x \geq 0$, a igualdade só se dando caso $x = 1$. Isso fecha o que queríamos provar. ♣

E. 5.10 *Exercício.* Mostre que no caso $0 < p < 1$ a desigualdade (5.35) se reverte ($\leq$ deve ser substituído por $\geq$). Nesse caso $q < 0$. ✠

• Desigualdades envolvendo somas de potências

As desigualdades apresentadas na seguinte proposição são muito úteis, especialmente para o propósito de demonstrar que os conjuntos de sequências ℓ_p são espaços vetoriais (vide Seção 27.5.1, página 1339), o mesmo se dando com os conjuntos de funções $L_p(M, d\mu)$ dos quais trataremos no Capítulo 33, página 1495.

Proposição 5.12 *Sejam $a \geq 0$ e $b \geq 0$ dois números reais não-negativos.*

I. *Para todo p tal que $0 < p < 1$ tem-se*

$$2^{p-1}\big(a^p + b^p\big) \;\leq\; (a+b)^p \;\leq\; a^p + b^p \;. \tag{5.36}$$

II. *Para todo p tal que $p \geq 1$ tem-se*

$$a^p + b^p \;\leq\; (a+b)^p \;\leq\; 2^{p-1}\left(a^p + b^p\right) \;. \tag{5.37}$$

□

Prova. Apresentamos separadamente as demonstrações para os casos I e II.

Caso I. Tomemos $0 < p < 1$ fixo. Vamos primeiramente provar a seguinte desigualdade: para quaisquer $a,\ b \geq 0$ vale

$$(a+b)^p \;\leq\; a^p + b^p \;. \tag{5.38}$$

Para $a = 0$ isso é óbvio. Seja, então, $a > 0$. Nesse caso, podemos fatorar a^p e a desigualdade acima ficaria,

$$\left(1 + \frac{b}{a}\right)^p \;\leq\; 1 + \left(\frac{b}{a}\right)^p \;.$$

Para provar isso, tudo o que desejamos é provar que $f(x) := (1+x)^p - 1 - x^p$ satisfaz $f(x) \leq 0$ para todo $x \geq 0$. De fato, tem-se,

$$f'(x) \;=\; -px^{p-1}\left[1 - \frac{1}{\left(1+\frac{1}{x}\right)^{1-p}}\right] \;. \tag{5.39}$$

Como $1 + \frac{1}{x} \geq 1$ e $1 - p > 0$, segue que $f'(x) \leq 0$ para todo $x \geq 0$. Com isso, provamos que f é não-crescente. Como $f(0) = 0$, segue que $f(x) \leq 0$ para todo $x \geq 0$. Isso provou (5.38).

Vamos agora provar que

$$2^{p-1}\big(a^p + b^p\big) \leq (a+b)^p . \tag{5.40}$$

Para $x \geq 0$ e $0 < p < 1$ a função $\varphi(x) = x^p$ é côncava. Portanto, para qualquer λ com $0 \leq \lambda \leq 1$, tem-se

$$\lambda\varphi(a) + (1-\lambda)\varphi(b) \leq \varphi\left(\lambda a + (1-\lambda)b\right) .$$

Para $\lambda = 1/2$, isso fica $\frac{a^p+b^p}{2} \leq \left(\frac{a+b}{2}\right)^p$, que é (5.40), e a prova de (5.36) está completa.

Caso II. Para o caso $p = 1$ a desigualdade (5.37) é evidente. Tomemos, então, $p > 1$ fixo. Vamos primeiramente provar a seguinte desigualdade: para quaisquer $a,\ b \geq 0$ vale

$$a^p + b^p \leq (a+b)^p . \tag{5.41}$$

Para $a = 0$ isso é óbvio. Seja, então, $a > 0$. Nesse caso, podemos fatorar a^p e a desigualdade acima ficaria,

$$\left(1 + \frac{b}{a}\right)^p \geq 1 + \left(\frac{b}{a}\right)^p .$$

Para provar isso, tudo o que desejamos é provar que $f(x) := (1+x)^p - 1 - x^p$ satisfaz $f(x) \geq 0$ para todo $x \geq 0$. Agora, por (5.39),

$$f'(x) = -px^{p-1}\left[1 - \left(1 + \frac{1}{x}\right)^{p-1}\right] .$$

Como $1 + \frac{1}{x} \geq 1$ e $p - 1 > 0$, segue que $f'(x) \geq 0$ para todo $x \geq 0$. Com isso provamos que f é crescente. Como $f(0) = 0$, segue que $f(x) \geq 0$ para todo $x \geq 0$, provando o que queríamos.

Vamos agora provar que

$$(a+b)^p \leq 2^{p-1}\left(a^p + b^p\right) . \tag{5.42}$$

Para $x \geq 0$ e $p > 1$ a função $\varphi(x) = x^p$ é convexa. Portanto, para qualquer λ com $0 \leq \lambda \leq 1$, tem-se

$$\varphi\left(\lambda a + (1-\lambda)b\right) \leq \lambda\varphi(a) + (1-\lambda)\varphi(b) .$$

Para $\lambda = 1/2$, isso fica $\left(\frac{a+b}{2}\right)^p \leq \frac{a^p+b^b}{2}$, que é (5.42), e a prova de (5.37) está completa. ■

Um corolário útil é:

Corolário 5.5 *Para $0 < q < 2$ e para todos $z,\ w \in \mathbb{C}$ vale*

$$|z+w|^q + |z-w|^q \leq 2\big(|z|^q + |w|^q\big) . \tag{5.43}$$

Para $q \geq 2$ e para todos $z,\ w \in \mathbb{C}$ vale

$$|z+w|^q + |z-w|^q \leq 2^{q-1}\big(|z|^q + |w|^q\big) . \tag{5.44}$$

□

O Corolário 5.5 pode ser usado para obter-se certas generalizações da identidade do paralelogramo (na forma de desigualdades) em espaços ℓ_p, com $p \geq 1$. Vide relações (27.49) e (27.50), página 1348.

Prova do Corolário 5.5. Para $z,\ w \in \mathbb{C}$ e com $0 < p < 1$, tomemos $a = |z+w|^2$ e $b = |z-w|^2$ na primeira desigualdade em (5.36). Obtemos,

$$\frac{|z+w|^{2p} + |z-w|^{2p}}{2^{1-p}} \leq \big(|z+w|^2 + |z-w|^2\big)^p .$$

Agora, $|z+w|^2 + |z-w|^2 = 2|z|^2 + 2|w|^2$. Assim,

$$\frac{|z+w|^{2p} + |z-w|^{2p}}{2^{1-p}} \leq 2^p\big(|z|^2 + |w|^2\big)^p .$$

Agora, tomemos $a = |z|^2$ e $b = |w|^2$ na segunda desigualdade em (5.36). Teremos $\big(|z|^2 + |w|^2\big)^p \leq |z|^{2p} + |w|^{2p}$. Assim, estabelecemos que

$$\frac{|z+w|^{2p} + |z-w|^{2p}}{2^{1-p}} \;\leq\; 2^p\big(|z|^{2p} + |w|^{2p}\big)$$

que é (5.43) com $q = 2p$.

Provemos agora (5.44). Para $z,\ w \in \mathbb{C}$ e com $p \geq 1$, tomemos $a = |z+w|^2$ e $b = |z-w|^2$ na primeira desigualdade em (5.37). Obtemos,

$$|z+w|^{2p} + |z-w|^{2p} \;\leq\; \big(|z+w|^2 + |z-w|^2\big)^p \;=\; 2^p\big(|z|^2 + |w|^2\big)^p \;.$$

Agora, adotemos $a = |z|^2$ e $b = |w|^2$ na segunda desigualdade em (5.37). Obtemos, $\big(|z|^2 + |w|^2\big)^p \leq 2^{p-1}\big(|z|^{2p} + |w|^{2p}\big)$. Assim, temos

$$|z+w|^{2p} + |z-w|^{2p} \;\leq\; 2^{2p-1}\big(|z|^{2p} + |w|^{2p}\big) \;.$$

Tomando-se $q = 2p$ essa é a desigualdade (5.44). ■

5.2.3.1 A Desigualdade de Minkowski

Uma importante desigualdade empregada do estudo de espaços métricos e na Análise Funcional é a chamada *desigualdade de Minkowski*[13]. Apresentaremos adiante demonstrações dessa degualdade am algumas instâncias, fazendo uso central da convexidade da função $f(x) = x^p$ na região $x > 0$, sendo $p \geq 1$. Outras demonstrações dessa desigualdade podem ser encontradas na Seção 27.5.1, página 1339, e na Seção 33.4.1, página 1532. A estratégia que seguimos na presente seção provém de [274].

Proposição 5.13 (Desigualdade de Minkowski para sequências finitas) *Seja $p \geq 1$. Para $n \in \mathbb{N}$, sejam, $a_k \in \mathbb{C}$ e $b_k \in \mathbb{C}$, $k = 1, \ldots, n$. Então, vale a desigualdade*

$$\left[\sum_{k=1}^{n} \big|a_k + b_k\big|^p\right]^{1/p} \;\leq\; \left[\sum_{k=1}^{n} \big|a_k\big|^p\right]^{1/p} + \left[\sum_{k=1}^{n} \big|b_k\big|^p\right]^{1/p} , \tag{5.45}$$

conhecida como desigualdade de Minkowski *para sequências finitas.* □

Prova. Se todos os a_k's forem nulos ou se todos os b_k's forem nulos, então (5.45) é evidente. Vamos, então, supor que

$$A \;:=\; \left[\sum_{k=1}^{n} \big|a_k\big|^p\right]^{1/p} \;\neq\; 0 \qquad \text{e} \qquad B \;:=\; \left[\sum_{k=1}^{n} \big|a_k\big|^p\right]^{1/p} \;\neq\; 0 \,.$$

É claro que $\big|a_k + b_k\big| \leq \big|a_k\big| + \big|b_k\big|$ e, portanto, $\big|a_k + b_k\big|^p \leq \Big(\big|a_k\big| + \big|b_k\big|\Big)^p$. Assim, podemos escrever que

$$\big|a_k + b_k\big|^p \;\leq\; \Big(\big|a_k\big| + \big|b_k\big|\Big)^p \;=\; (A+B)^p \left(\frac{A}{A+B}\frac{\big|a_k\big|}{A} + \frac{B}{A+B}\frac{\big|b_k\big|}{B}\right)^p \,. \tag{5.46}$$

Agora, como $\frac{A}{A+B} + \frac{B}{A+B} = 1$, com ambos os termos positivos, vemos que $\frac{A}{A+B}\frac{|a_k|}{A} + \frac{B}{A+B}\frac{|b_k|}{B}$ é uma combinação linear convexa dos números $\frac{|a_k|}{A}$ e $\frac{|b_k|}{B}$. Como a função $f(x) = x^p$ é convexa na região $x > 0$ quando $p \geq 1$, segue que

$$\left(\frac{A}{A+B}\frac{\big|a_k\big|}{A} + \frac{B}{A+B}\frac{\big|b_k\big|}{B}\right)^p \;\leq\; \frac{A}{A+B}\left(\frac{\big|a_k\big|}{A}\right)^p + \frac{B}{A+B}\left(\frac{\big|b_k\big|}{B}\right)^p \,.$$

Retornando com isso a (5.46), ficamos com

$$\big|a_k + b_k\big|^p \;\leq\; (A+B)^{p-1}\Big[A^{1-p}\big|a_k\big|^p + B^{1-p}\big|b_k\big|^p\Big] \,.$$

[13] Hermann Minkowski (1864–1909).

Somando-se em k, teremos

$$\sum_{k=1}^{n}|a_k+b_k|^p \;\leq\; (A+B)^{p-1}\left[A^{1-p}\underbrace{\sum_{k=1}^{n}|a_k|^p}_{=A^p}+B^{1-p}\underbrace{\sum_{k=1}^{n}|b_k|^p}_{=B^p}\right] \;=\; (A+B)^{p-1}\Big[A+B\Big] \;=\; (A+B)^p\,.$$

Logo, tomando-se a p-ésima raíz, obtemos $\left[\sum_{k=1}^{n}|a_k+b_k|^p\right]^{1/p}\leq A+B$, que é a desigualdade (5.45). ■

Um corolário imediato é

Corolário 5.6 (Desigualdade de Minkowski para sequências p-somáveis) *Seja $p\geq 1$ e sejam $\{a_k,\ k\in\mathbb{N}\}$ e $\{b_k,\ k\in\mathbb{N}\}$, duas sequências de números complexos tais que*

$$\sum_{k=1}^{\infty}|a_k|^p \;<\; \infty \qquad e \qquad \sum_{k=1}^{\infty}|b_k|^p \;<\; \infty\,. \tag{5.47}$$

Então, vale a desigualdade

$$\left[\sum_{k=1}^{\infty}|a_k+b_k|^p\right]^{1/p} \;\leq\; \left[\sum_{k=1}^{\infty}|a_k|^p\right]^{1/p}+\left[\sum_{k=1}^{\infty}|b_k|^p\right]^{1/p}\,, \tag{5.48}$$

conhecida como desigualdade de Minkowski *para sequências p-somáveis.* □

Prova. Use que (5.45) vale para todo $n\in\mathbb{N}$ e tome o limite $n\to\infty$ levando em conta (5.47). ■

Uma segunda demonstração da desigualdade (5.48) será apresentada na Seção 27.5.1, página 1339.

E. 5.11 *Exercício.* Sejam f e g duas funções contínuas definidas em um intervalo fechado $[a,\ b]\subset\mathbb{R}$ assumindo valores em $\mathbb{C}$. Mostre, imitando a demonstração da Proposição 5.13, que vale a desigualdade

$$\left[\int_a^b|f(x)+g(x)|^p\,dx\right]^{1/p} \;\leq\; \left[\int_a^b|f(x)|^p\,dx\right]^{1/p}+\left[\int_a^b|g(x)|^p\,dx\right]^{1/p}\,, \tag{5.49}$$

também conhecida como *desigualdade de Minkowski.* ✠

Os resultados acima podem ser muitíssimo generalizados (para funções p-integráveis em espaços mensuráveis). Vide Seção 33.4.1, página 1532.

• **O caso $0<p<1$**

E. 5.12 *Exercício.* Usando o fato que $f(x)=x^p$ é uma função côncava na região $x>0$ quando $0<p<1$, mostre que para tais valores de p vale

$$\left[\sum_{k=1}^{n}(a_k+b_k)^p\right]^{1/p} \;\geq\; \left[\sum_{k=1}^{n}(a_k)^p\right]^{1/p}+\left[\sum_{k=1}^{n}(b_k)^p\right]^{1/p}\,, \tag{5.50}$$

onde os a_k's e b_k's são números **reais não-negativos**. Note a reversão da desigualdade em (5.50) comparada a (5.45). ✠