UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ - UTFPR
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL - PROFMAT

CARLOS KRASSOWSKI FILHO

UMA INTRODUÇÃO ÀS FUNÇÕES CONVEXAS

CURITIBA

2018

CARLOS KRASSOWSKI FILHO

UMA INTRODUÇÃO ÀS FUNÇÕES CONVEXAS

Dissertação apresentada ao Mestrado Profissional em Matemática em Rede Nacional da Universidade Tecnológica Federal do Paraná em Curitiba - PROFMAT-UTCT como requisito parcial para obtenção do grau de Mestre.
Orientador: Dr. Andrés David Báez Sánchez

CURITIBA

2018

Dados Internacionais de Catalogação na Publicação

Krassowski Filho, Carlos
Uma introdução às funções convexas / Carlos Krassowski Filho.-- 2018.
1 arquivo texto (63 f.): PDF; 1,32 MB

Modo de acesso: World Wide Web
Texto em português com resumo em inglês
Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná. Programa de Mestrado Profissional em Matemática em Rede Nacional, Curitiba, 2018
Bibliografia: f. 63

1. Matemática - Dissertações. 2. Funções (Matemática). 3. Método de decomposição. 4. Funções convexas. 5. Funções de variáveis reais. 6. Funções de várias variáveis reais. I. Báez-Sánchez, Andrés David. II. Universidade Tecnológica Federal do Paraná. Programa de Mestrado Profissional em Matemática em Rede Nacional. III. Título.

CDD: Ed. 23 -- 510

Biblioteca Central da UTFPR, Câmpus Curitiba
Bibliotecário: Adriano Lopes CRB-9/1429

TERMO DE APROVAÇÃO DE DISSERTAÇÃO Nº 56

A Dissertação de Mestrado intitulada "Uma Introdução às Funções Convexas", defendida em sessão pública pelo candidato Carlos Krassowski Filho, no dia 07 de dezembro de 2018, foi julgada para a obtenção do título de Mestre, área de concentração Matemática, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Matemática em Rede Nacional.

BANCA EXAMINADORA:

Prof. Dr. Andres David Baez Sanchez - Presidente – UTFPR
Prof. Dr. Marcio Rostirolla Adames – UTFPR
Prof. Dr. Cristiano Torezzan - UNICAMP

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 07 de dezembro de 2018.

Carimbo e Assinatura do(a) Coordenador(a) do Programa

AGRADECIMENTOS

Agradeço aos meus familiares que me apoiaram nessa caminhada. Em especial ao meu pai, Carlos, pelo integral apoio. À minha mãe, Sueli e meu irmão, Ricardo pelas confortáveis hospedagens na capital nesse período.

Agradeço, também, ao meu orientador, professor David, pois cada orientação foi uma aula muito rica com discussões relevantes repletas de entusiasmo.

Aos grandes amigos que fiz ao ingressar na turma 2016. Ali encontrei muito mais que colegas, pessoas boas que vestiram a camisa e se uniram, de maneira sem igual, para enfrentar os desafios.

Aos professores que tive em toda a caminhada acadêmica, em particular os que lecionaram as disciplinas do PROFMAT, pelas instruções e inspirações: Paula, Mari, Ronie, Denise, André, Roy, Márcio e Patrícia.

E todas as outras pessoas que, de maneira indireta, contribuiram para a realização deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“A persistência é o caminho do êxito”.

(Charles Chaplin)

RESUMO

KRASSOWSKI, Carlos Filho. UMA INTRODUÇÃO ÀS FUNÇÕES CONVEXAS. 63 f. Dissertação - Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, Universidade Tecnológica Federal do Paraná. Curitiba, 2018.

Este trabalho consiste em uma introdução às funções convexas de uma variável real. Após um estudo do conceito de conjunto convexo é considerada a noção de função convexa. São apresentados diversos resultados teóricos sobre a caraterização e propriedades das funções convexas. Como aplicação destes resultados são exploradas algumas generalizações e relações entre médias de números positivos e desenvolvidas demonstrações alternativas para diversas desigualdades entre médias consideradas no livro de Matemática Discreta da coleção do PROFMAT. Além disso, é feito uma exploração sobre o problema de decomposição de uma função como diferença de funções convexas, baseado no conceito de variação total de uma função.

Palavras-chave: Convexidade; Funções Convexas; Desigualdades entre Médias.

ABSTRACT

KRASSOWSKI, Carlos Filho. AN INTRODUCTION TO CONVEX FUNCTIONS. 63 pg. Dissertation - Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, Universidade Tecnológica Federal do Paraná. Curitiba, 2018

This dissertation presents an introduction to convex functions of one real variable. After studying the concept of convex set, the notion of convex function is considered. Several theoretical results related to characterizations and properties of convex functions are presented. As an application of these results, some generalizations and relationships between means of positive numbers are explored and alternative proofs of means inequalities presented in the book of Discrete Mathematics of PROFMAT collection, are developed. In addition, is presented an exploration on the problem of decomposition of a function as the difference of convex functions, based on the concept of total variation of a function.

Keywords: Convexity; Convex Functions; Mean Inequalities.

LISTA DE ILUSTRAÇÕES

SUMÁRIO

INTRODUÇÃO

As funções convexas, tema dessa dissertação, são relevantes em diversas áreas, e suas propriedades permitem demonstrações de resultados importantes, como algumas desigualdades que se fazem presentes neste trabalho.

O ramo da Matemática que estuda os conjuntos convexos e as funções convexas é denominado de Análise Convexa. Os conjuntos convexos representam, geometricamente, regiões de forma que um segmento definido por quaisquer dois pontos da região, está inteiramente contido nessa mesma região.

O objetivo deste trabalho é discutir algumas definições e resultados referentes às funções convexas. Para ilustrar alguns exemplos e resultados foram produzidas algumas ilustrações feitas com auxílio do GeoGebra.

Dentre os trabalhos do PROFMAT relacionados ao tema, destaca-se a dissertação intitulada *Uma Introdução à Análise Convexa, Conjuntos e Funções Convexas* de Amorin (2013). Nela são apresentadas as principais ideias concernentes aos conjuntos e funções convexas, realizando uma revisão bibliográfica acerca do tema. O autor destina um capítulo para fazer uma revisão básica de conceitos referentes à Álgebra Linear e à Análise Real. Na presente dissertação, estes assustos serão assumidos já conhecidos.

Este trabalho estrutura-se em quatro capítulos. No primeiro capítulo, intitulado **conjuntos convexos**, são apresentadas definições e propriedades básicas relacionadas ao conceito de conjunto convexo, envoltório convexo e combinações convexas.

O segundo capítulo desenvolve um estudo sobre as **funções convexas de uma variável real**. São discutidas algumas caracterizações e propriedades de funções convexas incluindo uma caracterização por meio de retas secantes e outra via derivadas.

No capítulo três são discutidas as desigualdades entre médias clássicas e médias mais gerais (p-Médias). Algumas desigualdades entre médias clássicas estão incluidas em livros da coleção do PROFMAT, Morgado e Carvalho (2015), e são demonstradas utilizando diversas estratégias não tão naturais. Neste trabalho estas demonstrações e suas generalizações são devenvolvidas utilizando os resultados de funções convexas.

No quarto capítulo é estudado o problema da decomposição de funções reais como diferença entre convexas. Embora a convexidade não seja característica comum a todas as funções reais, é possível decompor determinadas funções, sob algumas condições, em termos de funções convexas. Neste trabalho será explorado um procedimento para obter esta decomposição.

1 CONJUNTOS CONVEXOS

A proposta deste capítulo é fazer um estudo inicial sobre conjuntos convexos, trazendo algumas definições e resultados com ilustrações. Os conceitos trabalhados nesse capítulo foram inspirados na obra de Krantz (2015).

Informalmente um conjunto C é convexo se, e somente se, para quaisquer pontos a e b em C, o segmento $\overline{ab}$ está inteiramente em contido C.

1.1 DEFINIÇÕES E PROPRIEDADES BÁSICAS

Definição 1. *Um conjunto C é **convexo** se dados pontos x e y em C e um escalar $\lambda \in [0, 1]$, o ponto dado por $\lambda x + (1 - \lambda)y \in C$.*

A figura que segue ilustra uma representação de um conjunto convexo.

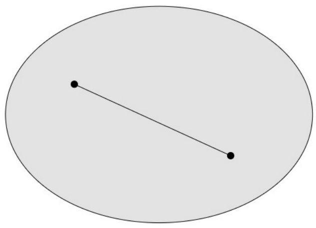

Figura 1 – Conjunto convexo

O conjunto ilustrado na figura que segue não é convexo pois existe parte do segmento fora do conjunto.

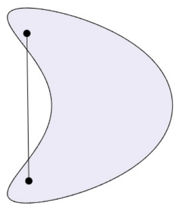

Figura 2 – Conjunto não convexo

Em seguida são apresentados exemplos de conjuntos com sua convexidade estudadas.

Exemplo 1.1.1. Toda reta é um conjunto convexo.

Exemplo 1.1.2. O conjunto dado por

$$C = \{x \in \mathbb{R}^n : Ax = y, A \in \mathbb{R}^{m \times n}, y \in \mathbb{R}^m\}$$

é convexo. De fato, considerando $x_1, x_2 \in C$, então para qualquer $\alpha \in [0,1]$,

$$A(\alpha x_1 + (1-\alpha)x_2) = \alpha A x_1 + (1-\alpha)Ax_2 = \alpha y + (1-\alpha)y = y$$

mostrando que $\alpha x_1 + (1-\alpha)x_2 \in C$.

Exemplo 1.1.3. Qualquer espaço vetorial é um conjunto convexo.

Se V é um espaço vetorial, a soma de dois vetores de V e a multiplicação de um vetor de V por escalar estão em V, logo, tomando quaisquer λ em $[0,1]$ e dois vetores x e y quaisquer em V, $\lambda x + (1-\lambda)y \in V$.

Exemplo 1.1.4. O conjunto $S = \{x \in \mathbb{R}^n : Ax \leq b\}$ com A uma matriz n por n e $b \in \mathbb{R}^n$ é convexo. Nesse contexto, a expressão $Ax \leq b$ diz respeito a um sistema de inequações lineares.

Sejam $x, y \in S$ e $\lambda \in [0,1]$, então

$$\begin{aligned}
A(\lambda x + (1-\lambda)y) &= A\lambda x + A(1-\lambda)y && \text{(distributividade)}\\
&= \lambda Ax + (1-\lambda)Ay\\
&\leq \lambda b + (1-\lambda)b && \text{(hipótese)}\\
&\leq b
\end{aligned}$$

Exemplo 1.1.5. O conjunto $B = \{x \in \mathbb{R}^n : \ ||x|| \leq r\}$, com $x \in \mathbb{R}^n$, $r \in \mathbb{R}_+$ é convexo.

$$\begin{aligned}
\| \lambda x + (1-\lambda)y \| &\leq \| \lambda x \| + \| (1-\lambda)y \| && \text{(desigualdade triangular)}\\
&= |\lambda| \ \| x \| + |(1-\lambda)| \ \| y \| && \text{(propriedades de norma)}\\
&= \lambda \| x \| + (1-\lambda) \| y \| && (\lambda, (1-\lambda) \geq 0)\\
&\leq \lambda r + (1-\lambda)r && \text{(hipótese)}\\
&= r.
\end{aligned}$$

Exemplo 1.1.6. O conjunto vazio é convexo. Supondo que o vazio não seja convexo, então haverão dois pontos tais que o segmento definido por eles pode estar parcialmente fora, no entanto não existem esses pontos pois o conjunto é vazio, configurando um absurdo.

Teorema 1. *Se C e D são conjuntos convexos de $\mathbb{R}^n$, então:*

1. $C + D = \{c + d : c \in C, d \in D\}$ *é convexo.*
2. $\alpha C := \{\alpha c : c \in C\}$ *com α fixado em $\mathbb{R}$ é convexo.*
3. *A intersecção de uma coleção de conjuntos convexos é convexa.*

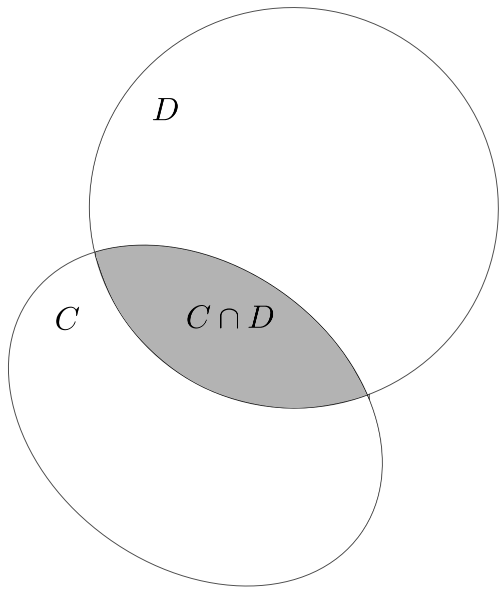

Figura 3 – A intersecção dos conjuntos C e D é convexa.

Demonstração. 1) Sejam $c_1, c_2 \in C, d_1, d_2 \in D$, além disso, considerando $x = c_1 + d_1$ e $y = c_2 + d_2$, com $x, y \in (C + D)$. Sabe-se da convexidade de C e de D que para algum $\lambda \in [0, 1]$, tem-se que $\lambda c_1 + (1 - \lambda)c_2 \in C$ e $\lambda d_1 + (1 - \lambda)d_2 \in D$. Nessas condições,

$$\lambda(c_1+d_1)+(1-\lambda)(c_2+d_2) = (\lambda c_1+(1-\lambda)c_2)+(\lambda d_1+(1-\lambda)d_2) \in (C+D).$$

2) Sejam $x, y \in \alpha C$, isto é, $x = \alpha c_1$ e $y = \alpha c_2$ com $c_1, c_2 \in C$. Da convexidade de C, tem-se que $c = \lambda c_1 + (1 - \lambda)c_2 \in C$.

Sabemos que

$$\begin{aligned}\lambda x + (1-\lambda)y &= \lambda\alpha c_1 + (1-\lambda)\alpha c_2\\ &= \alpha(\lambda c_1 + (1-\lambda)c_2)\\ &= \alpha c \in \alpha C.\end{aligned}$$

3) Se $x, y \in \cap_{i\in I} C_i$ então $x, y \in C_i$, para todo i. Como C_i é convexo, então para qualquer $\alpha \in [0,1], \alpha y + (1-\alpha)y \in C_i$, para todo i.

Portanto, para todo $\alpha \in [0,1], \alpha p_1 + (1-\alpha)p_2 \in \cap_{i\in I} C_i$. $\square$

É possível mostrar que o lugar geométrico da soma de dois círculos C e D é outro círculo cujo centro e raio são dados pela soma dos centros e raios de C e D, respectivamente. Na Figura 4 o círculo C tem raio igual a 1 e centro $(2,1)$ e o círculo D tem o mesmo raio e centro $(4,0)$. A soma $C + D$ é representada por um círculo de centro $(6,1)$ e raio igual a 2. A Figura 4, ainda ilustra que a união de dois conjuntos convexos nem sempre é convexa, como ocorre com os círculos C e D.

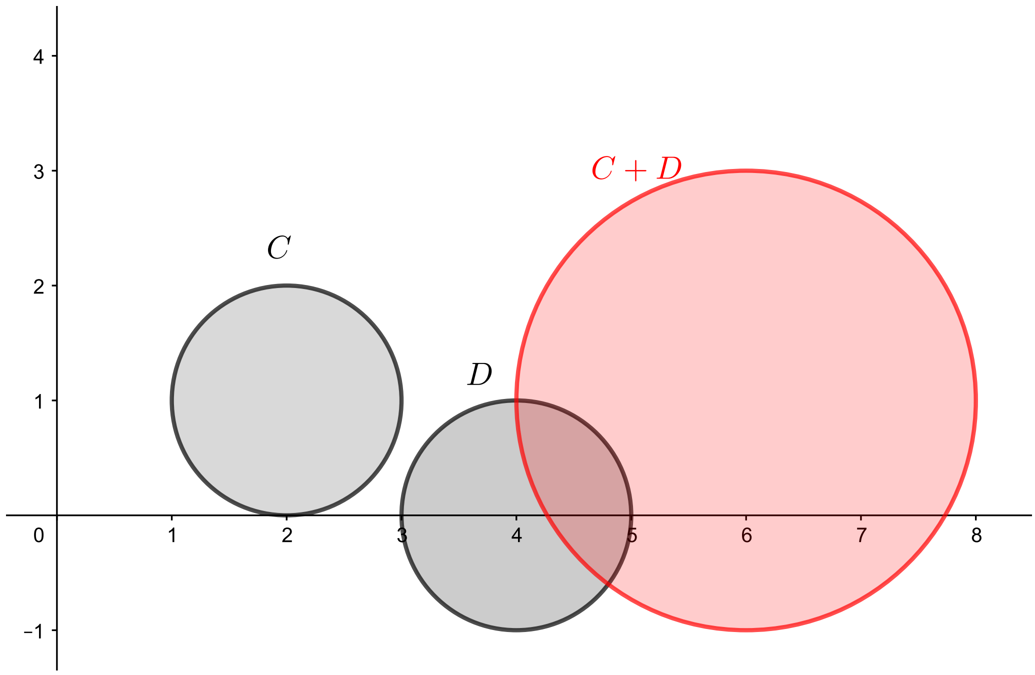

Figura 4 – $C + D$ é convexo. $C \cup D$ não é convexo.

1.2 ENVOLTÓRIO E COMBINAÇÕES CONVEXAS

Informalmente o envoltório convexo de um conjunto C pode ser entendido como o menor conjunto convexo que contém C.

Definição 2. *O **envoltório convexo** de C, denotado co(C), é definido como a intersecção de todos os conjuntos convexos que contém C, ou seja,*

$$co(C) = \cap_{B \in I} B,$$

com $I = \{B \text{ é convexo}, C \subseteq B\}$.

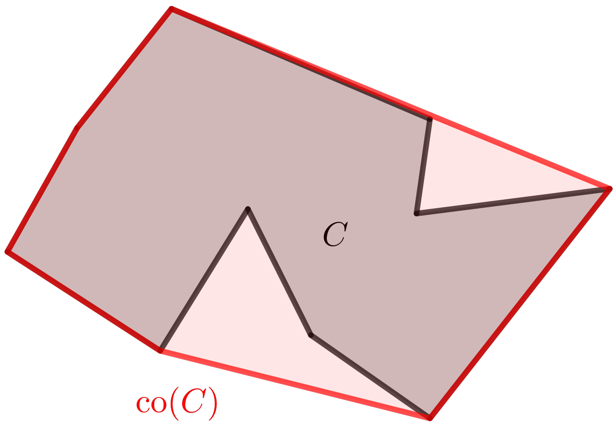

Figura 5 – Envoltório convexo de um conjunto C

A definição que segue permite generalizar o conceito de convexidade levando em conta n pontos em $\mathbb{R}^n$:

Definição 3. *Sejam $p_1, p_2, \ldots, p_n \in \mathbb{R}^n$. A expressão*

$$\alpha_1 p_1 + \alpha_2 p_2 + ... + \alpha_n p_n$$

é dita combinação convexa de $p_1, p_2, \ldots, p_n$, quando $\alpha_1, \alpha_2, \ldots, \alpha_n \in [0, 1]$ e $\sum_{i=1}^{n} \alpha_i = 1$.

Proposição 1.2.1. *B é convexo se e somente se para todo $p_1, \ldots, p_n \in B$ e todo $\lambda_i \in [0, 1]$ com $\sum \lambda_i = 1$, tem-se:*

$$\lambda_1 p_1 + \cdots + \lambda_n p_n \in B \tag{1.1}$$

Demonstração. Suponha B convexo. O resultado expressado por (1.1) será mostrado via indução matemática. O caso base é trivial pois a combinação convexa com um único ponto de B está em B. É necessário que $\lambda_1 = 1$ e um ponto $p_1 \in B$ qualquer, que $\lambda_1 p_1 = p_1 \in B$.

Agora, considere os pontos $p_1, \ldots, p_n, p_{n+1} \in B$ e que (1.1) seja válida para algum n.

Do fato da convexidade de B e de que os pontos p_{n+1} e $\lambda_1 p_1 + \cdots + \lambda_n p_n$ pertencem a B, tem-se que para qualquer $\alpha \in [0, 1]$ ocorre que:

$$(1 - \alpha)(\lambda_1 p_1 + \cdots + \lambda_n p_n) + \alpha p_{n+1} \in B. \tag{1.2}$$

Seja $\alpha_1, \ldots, \alpha_{n+1} \in [0, 1]$ com $\sum \alpha_i = 1$, então:

$$\alpha_1 + \cdots + \alpha_n = 1 - \alpha_{n+1}$$

Supondo $1 - \alpha_{n+1} \neq 0$, tomando $\lambda_i = \dfrac{\alpha_i}{1 - \alpha_{n+1}}$ e $\alpha = \alpha_{n+1}$, então de (1.2),

$$(1 - \alpha_{n+1})\left(\frac{\alpha_1}{1 - \alpha_{n+1}} p_1 + \cdots + \frac{\alpha_n}{1 - \alpha_{n+1}} p_n\right) + \alpha_{n+1} p_{n+1} \in B.$$

Disto, tem-se que

$$\alpha_1 p_1 + \cdots + \alpha_n p_n + \alpha_{n+1} p_{n+1} \in B, \tag{1.3}$$

Caso $1 - \alpha_{n+1} = 0$, então $\alpha_{n+1} = 1$, logo $\alpha_1 = \cdots = \alpha_n = 0$, ainda assim, a expressão (1.3) será válida pois se reduzirá a $p_{n+1} \in B$, concluindo a indução.

A recíproca é verdadeira. Basta tomar $n = 2$, pois recai diretamente na definição de conjunto convexo. □

A proposição que segue afirma que o envoltório convexo é o mesmo que todas as combinações convexas de um conjunto.

Proposição 1.2.2. *Seja C um conjunto convexo. O envoltório convexo de C é o conjunto de todas as combinações convexas de C.*

Demonstração. Seja U o conjunto das combinações convexas de C, ou seja,

$$U = \{\lambda_1 p_1 + \cdots + \lambda_n p_n : p_i \in C, \lambda_i \in [0,1], \sum \lambda_i = 1, n \in \mathbb{Z}^+\}.$$

Será provado que co(C)= U.

Primeiramente, será feita a prova de que co(C)$\subseteq U$. Seja $p \in C$, é possível observar que todo $p \in U$, considerando $n = 1$, $p_1 = p$, $\lambda_1 = 1$. Logo $C \subset U$.

Será provado que U é convexo. Tomados quaisquer duas combinações $x, y \in U$,

$$x = \alpha_1 p_1 + \cdots + \alpha_n p_n,$$

$$y = \beta_{n+1} p_{n+1} + \cdots + \beta_m p_m$$

com $\alpha_i, \beta_i \in [0,1]$, $\sum \alpha_i = \sum \beta_i = 1$, considere $m > n$ e $\alpha_i = 0$, para todo i e $n+1, \ldots, m$.

Os pontos do segmento xy podem ser escritos como

$$\begin{aligned} \lambda x + (1-\lambda) y &= \lambda \sum_{i=1}^{m} \alpha_i p_i + (1-\lambda) \sum_{i=1}^{m} \beta_i p_i \\ &= \sum_{i=1}^{m} (\lambda \alpha_i + (1-\lambda)\beta_i) p_i, \end{aligned}$$

para $\lambda \in [0,1]$. No entanto,

$$\begin{aligned} \sum_{i=1}^{m} (\lambda \alpha_i + (1-\lambda)\beta_i) &= \lambda \sum_{i=1}^{m} \alpha_i + (1-\lambda) \sum_{i=1}^{m} \beta_i \\ &= \lambda \cdot (1) + (1-\lambda) \cdot (1) \\ &= 1, \end{aligned}$$

e $\lambda \alpha_i + (1-\lambda)\beta_i) \in [0,1]$, portanto, os pontos segmento xy são combinações convexas de elementos de C, ou seja, o segmento xy está contido em U, provando que U é convexo.

Foi mostrado que o conjunto U contém C e é convexo, dessa maneira ele é um dos conjuntos convexos cuja intersecção produzem co(C), portanto, co(C)$\subseteq U$.

Para monstrar que $U \subseteq$ co(C) é necessário que qualquer ponto de U esteja, também, em co(C).

Considere uma combinação convexa arbitrária de elementos de C, isto é, $\alpha_1 p_1 + \cdots + \alpha_n p_n \in U$. Seja B um conjunto convexo qualquer que contém C.

Como $p_i \in C$, segue que $p_i \in B$. Deste modo $\alpha_1 p_1 + \cdots + \alpha_n p_n$ é uma combinação convexa de elementos de B. Como B é convexo, pela Proposição 1.2.1 segue que $\alpha_1 p_1 + \cdots + \alpha_n p_n \in B$.

Então $U \subset B$ para qualquer B convexo que contenha C, portanto U está na intersecção que define co(C) isto é, $U \subseteq$ co(C). $\square$

Tendo definido e estudado algumas propriedades de conjuntos convexos, no capítulo seguinte será feito um estudo sobre funções convexas em uma variável real.

2 FUNÇÕES CONVEXAS DE UMA VARIÁVEL REAL

O objetivo deste capítulo é estudar as funções convexas de uma variável. Alguns resultados fundamentais serão apresentados, demonstrados e ilustrados com exemplos.

Serão consideradas algumas caracterizações de funções convexas: uma por meio de retas secantes e outra por meio de tangentes, usando derivadas. Alguns dos resultados deste capítulo foram baseados na teoria presente no terceiro capítulo da obra de Boyd e Vandenberghe (2004).

2.1 FUNÇÕES CONVEXAS. DEFINIÇÕES E ALGUMAS PROPRIEDADES

A definição que segue é a mais importante deste trabalho, sua compreensão é fundamental para dar prosseguimento à teoria.

Definição 4. *Seja I um intervalo. Uma função $f : I \subset \mathbb{R} \to \mathbb{R}$ é dita convexa em I se, para todo $x, y \in I$ e todo $\lambda \in [0, 1]$, satisfaz a desigualdade:*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Daqui em diante, sempre que não houver ambiguidade, será omitida a referência ao intervalo I em que a função seja convexa.

Uma função $f : I \subset \mathbb{R} \to \mathbb{R}$ é convexa se e somente se para quaisquer par de pontos x, y no domínio, com $x > y$, o segmento de reta com extremos $(x, f(x))$ e $(y, f(y))$ fica acima do gráfico de f no intervalo (x, y). Na Figura 6 é ilustrado o conceito de função convexa de uma variável.

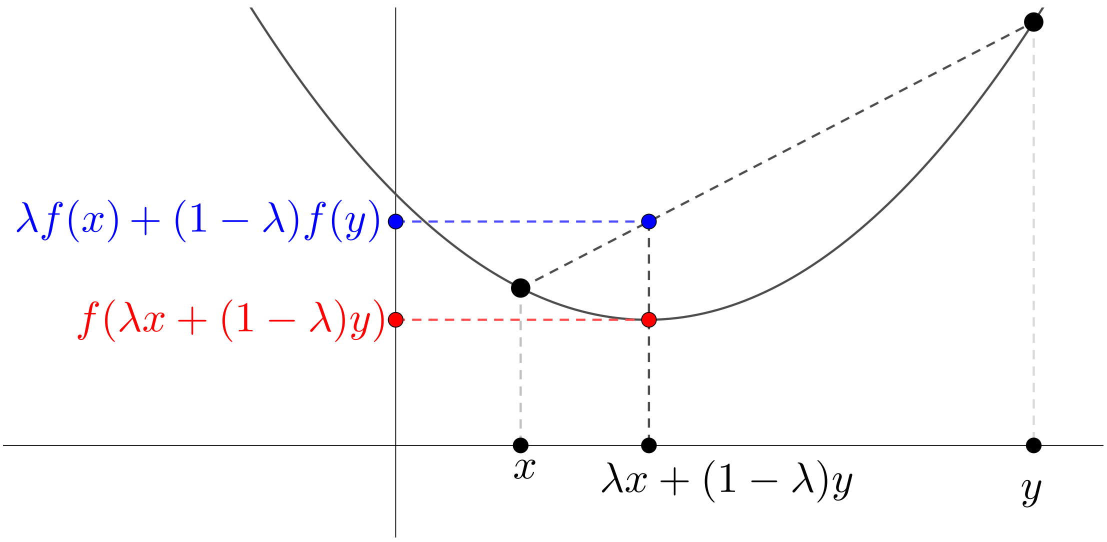

Figura 6 – Função convexa

Definição 5. *Uma função $f : I \subset \mathbb{R} \to \mathbb{R}$ é dita concava quando, $-f$ é convexa.*

Exemplo 2.1.1. A partir da análise da Figura 8, poderia se concluir que a função $f(x) = x^2$ é, de fato, convexa, mas para assegurar tal afirmação, será feita uma verificação das condições de convexidade.

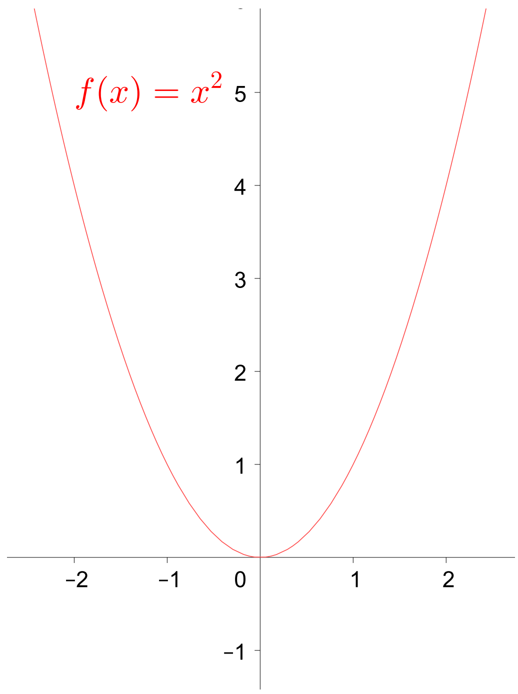

Figura 7 – $f(x) = x^2$

Primeiramente é preciso considerar que como $\lambda \in [0,1]$, então $\lambda - 1$ é negativo, além disso, a expressão $(x-y)^2 \geq 0$ é sempre verdadeira, então:

$$\lambda(\lambda - 1)(x - y)^2 \leq 0,$$

disto, seguem as seguintes desigualdades equivalentes:

$$\begin{aligned}
\lambda(\lambda - 1)(x^2 - 2xy + y^2) &\leq 0 \\
\lambda(\lambda - 1)x^2 + \lambda(\lambda - 1)y^2 - 2\lambda(\lambda - 1)xy &\leq 0 \\
\lambda x^2(\lambda - 1) + (1 - \lambda)y^2(1 + \lambda - 1) + 2\lambda(1 - \lambda)xy &\leq 0 \\
\lambda^2 x^2 - \lambda x^2 + (1 - \lambda)^2 y^2 - (1 - \lambda)y^2 + 2\lambda(1 - \lambda)xy &\leq 0 \\
\lambda^2 x^2 + 2\lambda x(1 - \lambda)y + (1 - \lambda)^2 y^2 &\leq \lambda x^2 + (1 - \lambda)y^2 \\
(\lambda x + (1 - \lambda)y)^2 &\leq \lambda x^2 + (1 - \lambda)y^2 \\
f(\lambda x + (1 - \lambda)y) &\leq \lambda f(x) + (1 - \lambda)f(y)
\end{aligned}$$

Portanto, com base na última desigualdade, $f(x)$ é convexa.

Exemplo 2.1.2. Como $f(x) = x^2$ é convexa, então $g(x) = -x^2$ é côncava para todo $x \in \mathbb{R}$.

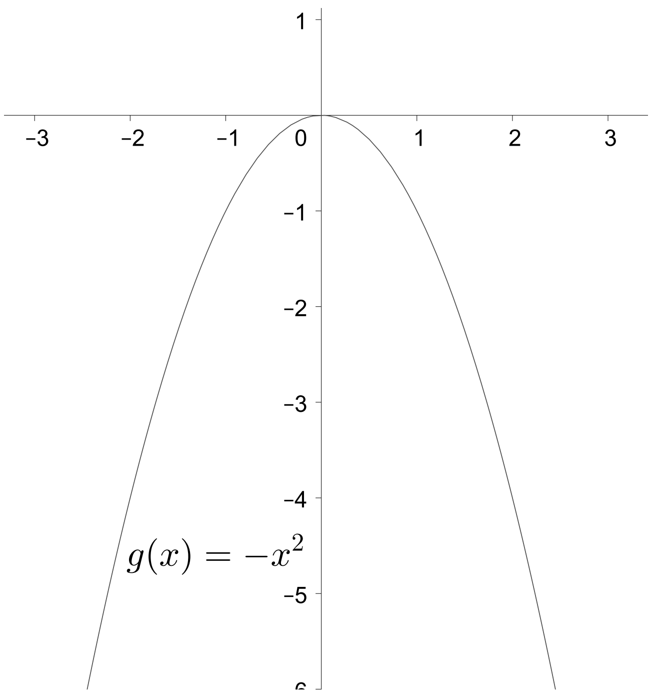

Figura 8 – Representação gráfica da função côncava $g(x) = -x^2$.

Exemplo 2.1.3. A função $f(x) = x^3$ não é convexa e nem côncava em $\mathbb{R}$. Como contra exemplo para a convexidade, basta considerar $x = -1, y = 1, \lambda = \frac{3}{4}$.

Supondo que $f(x) = x^3$ é convexa, então,

$$\begin{aligned} f(\lambda(-1) + (1-\lambda)\cdot 1) &\leq \lambda f(-1) + (1-\lambda) f(1) \\ f\left(\frac{3}{4}(-1) + \left(1 - \frac{3}{4}\right)\cdot 1\right) &\leq \frac{3}{4} f(-1) + \left(1 - \frac{3}{4}\right) f(1) \\ f\left(-\frac{1}{2}\right) &\leq \frac{3}{4} f(-1) + \frac{1}{4} f(1) \\ \left(-\frac{1}{2}\right)^3 &\leq \frac{3}{4}(-1)^3 + \frac{1}{4}\cdot 1^3 \\ -\frac{1}{8} &\leq -\frac{1}{2}, \end{aligned}$$

porém, a última desigualdade constitui um absurdo, assim $f(x) = x^3$ não é convexa. A função também não é côncava, pois basta considerar o caso em que $\lambda = \frac{1}{4}$, para os mesmos x e y. Estes fatos são ilustrados na Figura 9.

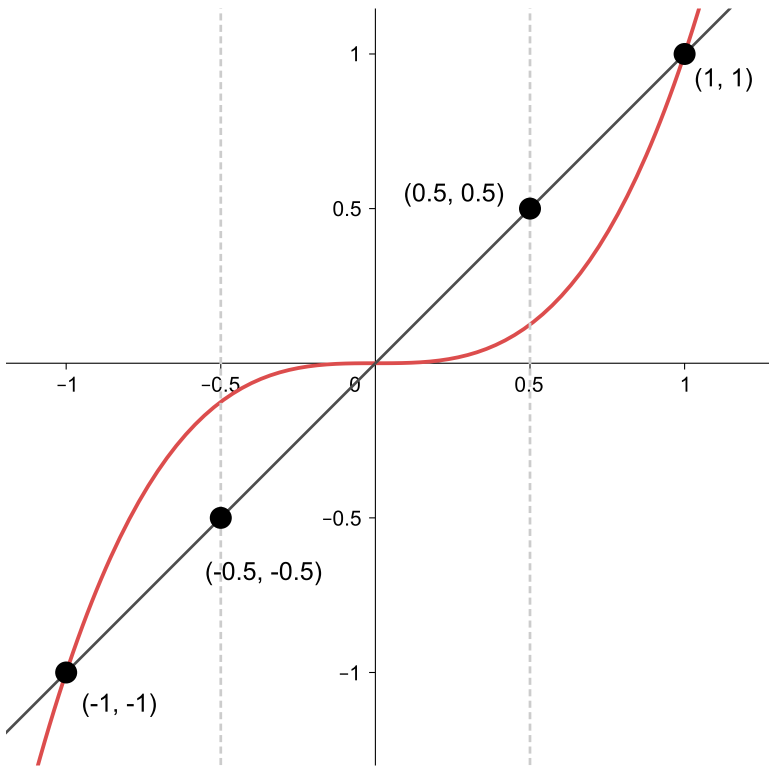

Figura 9 – $f(x) = x^3$, nem convexa, nem côncava.

O exemplo anterior ilustra que o fato de que uma função não ser convexa não implica, necessariamente, que ela seja concava.

A Proposição 2.1.4 extende a ideia da definição de convexidade, considerando n pontos ao invés de apenas dois.

Proposição 2.1.4. *Uma função $f : I \subset \mathbb{R} \to \mathbb{R}$ é convexa, se e somente se, para todo $n \in \mathbb{N}$ com $n > 2$, se satisfaz que, se $x_1, \dots x_n \in I$ e $\lambda_1, \dots, \lambda_n \in [0, 1]$ tais que $\lambda_1 + \cdots + \lambda_n = 1$, então*

$$f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n). \tag{2.1}$$

Demonstração. Será utilizado o Princípio da Indução Matemática. A verificação para $n = 2$ recai na definição de convexidade.

Supondo que f seja convexa e que

$$f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n)$$

seja válida para algum $n \in \mathbb{N}, n > 2$, e para quaisquer $0 \leq \lambda_i \leq 1$ com $\sum_{i=1}^{n} \lambda_i = 1$ e $x_i \in I$ com $i = 1, 2, \dots, n$.

Será mostrado que a desiguldade

$$f(\lambda_1 x_1 + \cdots + \lambda_{n+1} x_{n+1}) \leq \lambda_1 f(x_1) + \cdots + \lambda_{n+1} f(x_{n+1})$$

também é válida, para $x_1, \dots, x_{n+1} \in I$ e $\lambda_1, \dots, \lambda_{n+1} \geq 0$ com $\lambda_1 + \cdots + \lambda_{n+1} = 1$.

Se $\lambda_{n+1} = 1$, então $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$ o que reduziria a análise ao caso $f(1 \cdot x_{n+1}) \leq 1 \cdot f(x_{n+1})$, a desigualdade se satisfaz trivialmente.

Supondo, então, que $\lambda_{n+1} \neq 1$, tomando

$$y = \frac{\lambda_1 x_1}{1 - \lambda_{n+1}} + \cdots + \frac{\lambda_n x_n}{1 - \lambda_{n+1}}, \tag{2.2}$$

e observando que $\lambda_1 + \cdots + \lambda_n = 1 - \lambda_{n+1}$ implica em

$$\frac{\lambda_1}{1 - \lambda_{n+1}} + \cdots + \frac{\lambda_n}{1 - \lambda_{n+1}} = 1,$$

da convexidade de f e da hipótese indutiva, segue que:

$$
\begin{aligned}
&f(\lambda_1 x_1 + \cdots + \lambda_n x_n + \lambda_{n+1} x_{n+1}) = \\
&\qquad = f((1-\lambda_{n+1})y + \lambda_{n+1}x_{n+1}) \\
&\qquad \leq (1-\lambda_{n+1})f(y) + \lambda_{n+1}f(x_{n+1}) \\
&\qquad = (1-\lambda_{n+1})f\left(\frac{\lambda_1}{1-\lambda_{n+1}}x_1 + \cdots + \frac{\lambda_n}{1-\lambda_{n+1}}x_n\right) + \lambda_{n+1}f(x_{n+1}) \\
&\qquad \leq (1-\lambda_{n+1})\left(\frac{\lambda_1}{1-\lambda_{n+1}}f(x_1) + \cdots + \frac{\lambda_n}{1-\lambda_{n+1}}f(x_n)\right) + \lambda_{n+1}f(x_{n+1}) \\
&\qquad = \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n) + \lambda_{n+1}f(x_{n+1})
\end{aligned}
$$

Reciprocamente, supondo que a desigualdade (2.1) é válida para qualquer $n \in \mathbb{N}$, em particular para $n = 2$ se tem que, por definição, f é convexa. □

Observação 2.1.5. Essa proposição pode ser interpretada através da média aritmética ponderada. Se uma uma função é convexa, então a função aplicada na média ponderada de uma coleção de pontos é menor do que ou igual a média ponderada da função aplicada em cada um dos pontos dessa coleção. Os pesos que ponderam essas médias são os λ_i.

A figura que segue representa essa desigualdade.

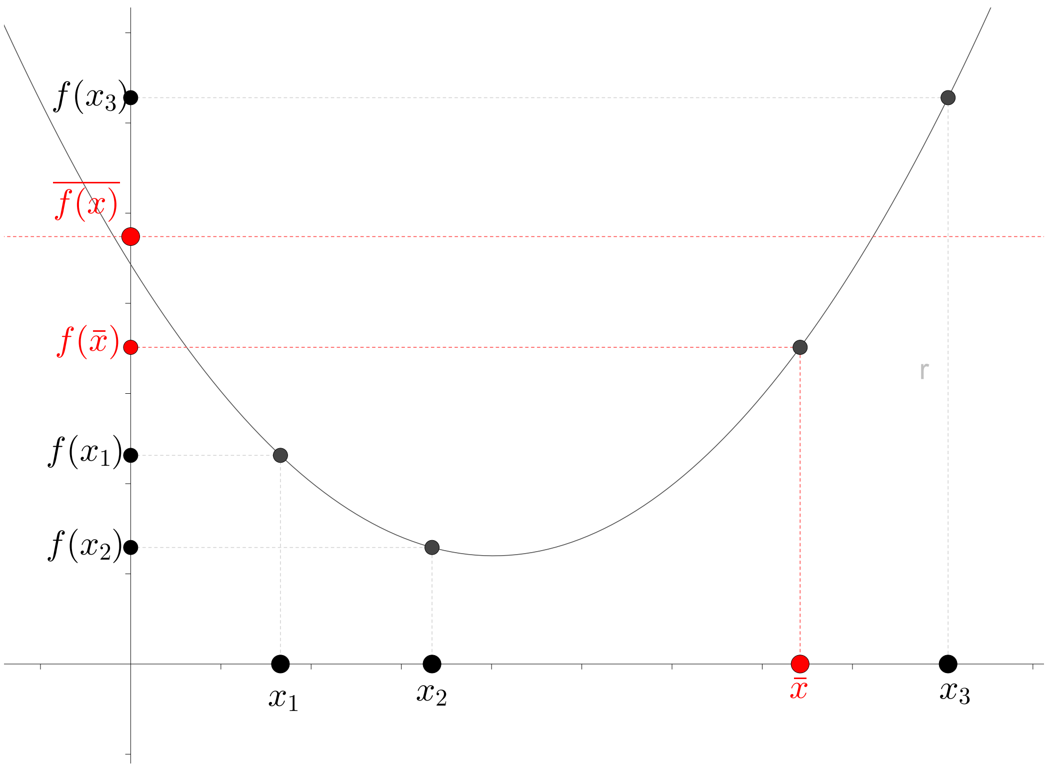

Figura 10 – Representação gráfica da Proposição 2.1.4.

Na Figura 10 são considerados $\bar{x} = \lambda_1 x_1 + \cdots + \lambda_n x_n$ e $\overline{f(x)} = \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n)$.

Exemplo 2.1.6. Tomando $f(x) = x^2$, avaliada nos pontos $x_1 = 1, x_2 = 2, x_3 = 5$ e $\lambda_1 = \frac{2}{9}, \lambda_2 = \frac{3}{9}, \lambda_3 = \frac{4}{9}$, tem-se:

$$f\left(\frac{2 \cdot 1 + 3 \cdot 2 + 4 \cdot 5}{9}\right) \leq \frac{2f(1) + 3f(2) + 4f(5)}{9}$$
$$\frac{256}{81} \leq \frac{114}{9}$$

Seguem algumas proposições sobre somas e produtos escalares de funções convexas.

Proposição 2.1.7. *Se $f : I \to \mathbb{R}$ e $g : I \to \mathbb{R}$ são convexas, então $f + g$ também será convexa em I.*

Demonstração. De fato, observar que

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \text{ e}$$
$$g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda) g(y)$$

Somando ambas desigualdades se tem:

$$f(\lambda x + (1-\lambda)y) + g(\lambda x + (1-\lambda)y) \leq \lambda(f(x) + g(x)) + (1-\lambda)(f(y) + g(y))$$

□

Proposição 2.1.8. *Se $f : I \to \mathbb{R}$ é convexa, então αf, com $\alpha \in [0, \infty)$ é, também, convexa.*

Demonstração. Se $f : I \to \mathbb{R}$ é convexa, então

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

Multiplicando por $\alpha \in \mathbb{R}^+$ em ambos os lados, tem-se que

$$\alpha f(\lambda x + (1-\lambda)y) \leq \lambda \alpha f(x) + (1-\lambda) \alpha f(y)$$

. □

Corolário 2.1.9. *Se $f : I \to \mathbb{R}$ e $g : I \to \mathbb{R}$ são convexas, então $\forall \alpha, \beta \in \mathbb{R}^+, \alpha f + \beta g$ também será convexa em I.*

Demonstração. Se $f : I \to \mathbb{R}$ e $g : I \to \mathbb{R}$ são convexas, então pela Proposição 2.1.8, αf e βg com $\alpha, \beta \in \mathbb{R}^+$ serão convexas. Logo, pela Proposição 2.1.7, $\alpha f + \beta g$ será convexa, também. □

Da definição de funções côncavas (Definição 5), segue seguinte corolário:

Corolário 2.1.10. *Se $f : I \to \mathbb{R}$ é convexa, então αf, com $\alpha \in (-\infty, 0]$ é côncava.*

Observação 2.1.11. *O conjunto das funções convexas em um intervalo I* ***não forma um campo vetorial****, pois ao multiplicar uma função convexa por um escalar negativo a função resultante não será, em geral, convexa. No entanto, pelo corolário anterior, a união do conjunto das funções convexas com as funções côncavas, forma um espaço vetorial V_I, definido como:*

$$V_I := \{f + g : f \text{ é convexa em } I \text{ e } g \text{ é côncava em } I\},$$

ou de forma equivalente:

$$V_I = \{f - g : f \text{ e } g \text{ convexas em } I\}.$$

Algumas questões relevantes relacionadas a V_I são: Quais são as funções que podem ser escritas como diferença entre funções convexas? Existe algum processo para determinar essa decomposição? Essa decomposição é única? Essas questões, mais a frente, serão discutidas no último capítulo.

2.1.1 CONVEXIDADE E COMPOSIÇÃO DE FUNÇÕES

Nesta subseção são estudadas algumas propriedades de convexidade em funções invertíveis.

Teorema 2. *Seja $f : I \subset \mathbb{R} \to \mathbb{R}$ uma função convexa, e invertível, então:*

(i) Se f é crescente, sua inversa é côncava;

(ii) Se f é decrescente, sua inversa é convexa.

Demonstração. Supondo f crescente, convexa e invertível, com g sua inversa, do fato de que g também é crescente, segue que

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &\leq \lambda f(x) + (1-\lambda)f(y) \\ g(f(\lambda x + (1-\lambda)y)) &\leq g(\lambda f(x) + (1-\lambda)f(y)) \\ \lambda x + (1-\lambda)y &\leq g(\lambda f(x) + (1-\lambda)f(y)) \\ \lambda g(f(x)) + (1-\lambda)g(f(y)) &\leq g(\lambda f(x) + (1-\lambda)f(y)). \end{aligned}$$

Da última desigualdade se conclui que g é côncava.

Analogamente, supondo f decrescente, com argumentos similares, g será convexa.

□

Exemplo 2.1.12. A função $f(x) = x^2$ com $x > 0$ é convexa e sua inversa, $f^{-1} = \sqrt{x}$, côncava.

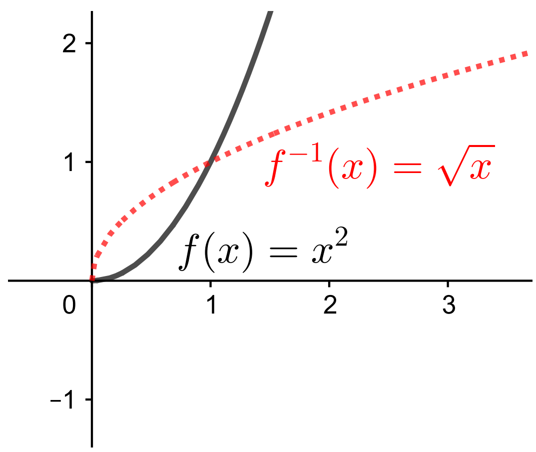

Figura 11 – x^2 convexa e $\sqrt{x}$ côncava.

Corolário 2.1.13. *Seja $f : I \subset \mathbb{R} \to \mathbb{R}$ uma função côncava, e invertível, então:*

(i) Se f é crescente, sua inversa é convexa;

(ii) Se f é decrescente, sua inversa é côncava.

Exemplo 2.1.14. A função $f(x) = \frac{1}{x}$ tem uma peculiaridade visto que $f^{-1}(x) = f(x)$, ou seja, ela é a própria inversa.

Para $x > 0$, $f(x)$ é decrescente e $f'(x)$ não-decrescente, logo, a inversa é convexa. Para $x < 0$, $f(x)$ é côncava.

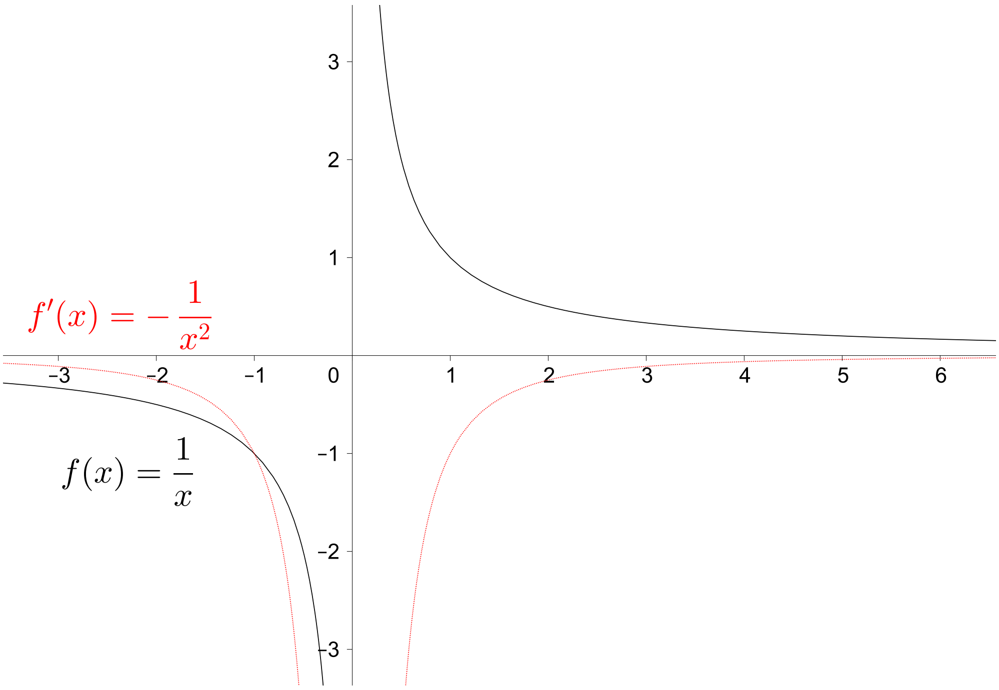

Figura 12 – $f(x) = \frac{1}{x}$ e sua derivada.

A Proposição 2.1.15 a convexidade da composição de duas funções convexas.

Proposição 2.1.15. *Sejam f e g funções convexas. Se g é crescente, então $g \circ f$ é convexa;*

Demonstração. Seja f e g funções convexas, então:

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &\leq \lambda f(x) + (1-\lambda)f(y) && \text{(pois } f \text{ é convexa)} \\ g(f(\lambda x + (1-\lambda)y)) &\leq g(\lambda f(x) + (1-\lambda)f(y)) && \text{(supondo } g \text{ crescente)} \\ g(f(\lambda x + (1-\lambda)y)) &\leq \lambda g(f(x)) + (1-\lambda)g(f(y)) && \text{(pois } g \text{ convexa)} \end{aligned}$$

Portanto, $g \circ f$ é convexa.

$\square$

Exemplo 2.1.16. A função $h(x) = e^{x^2}$ é convexa, pois, tomando $f(x) = x^2$ e $g(x) = e^x$, $h(x) = g(f(x))$ são satisfeitas as condições da Proposição (2.1.15).

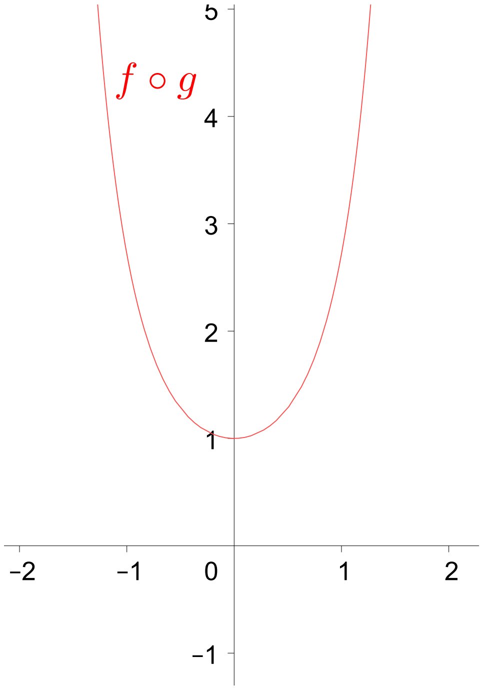

Figura 13 – $f(x) = x^2$ (convexa), $g(x) = e^x$ (convexa e crescente)

2.2 CARACTERIZAÇÃO DE FUNÇÕES CONVEXAS POR MEIO DA RETA SECANTE

O objetivo desta seção é exibir uma caracterização de funções convexas a partir do comportamente das inclinações das retas secantes.

A abordagem desta seção está baseada nas ideias discutidas em Barata (2018).

Para uma função $f : I \subset \mathbb{R} \to \mathbb{R}$, considere-se a função de duas variáveis R_f dada por

$$R_f(x, y) := \frac{f(x) - f(y)}{x - y} \text{ para } x \neq y.$$

A função R_f dá a inclinação da reta secante ao gráfico de f entre os pontos $(x, f(x))$ e $(y, f(y))$.

Note que $R_f(x, y) = R_f(y, x)$, portanto, R_f é uma função simétrica.

Para simplificar a notação será utilizado $R(x, y)$ no lugar de $R_f(x, y)$ quando não houver ambiguidade sobre a função.

A proposição que segue mostra como a função R se relaciona com a convexidade de f.

Proposição 2.2.1. *Uma função $f : I \subset \mathbb{R} \to \mathbb{R}$ é convexa se e somente se para todo $x, y, z \in I$, distintos, com $y < z$ se satisfaz*

$$R(x, y) \leq R(x, z)$$

Demonstração. Primeiramente, supondo que f é convexa, será provado que $R(x, y) \leq R(x, z)$. Para tal, há três casos que devem ser considerados: $x < y < z, y < x < z$ e $y < z < x$.

Caso 1: $x < y < z$. Do fato de que y está entre x e z, tem-se que $y = \lambda x + (1 - \lambda)z$. para algum $\lambda \in (0, 1)$. É possível determinar o valor de λ da seguinte forma:

$$\begin{aligned} y &= \lambda x + (1 - \lambda)z \\ \Rightarrow \ y - z &= \lambda(x - z) \\ \Rightarrow \ \lambda &= \frac{y - z}{x - z} \\ \Rightarrow \ \lambda &= \frac{z - y}{z - x} \end{aligned}$$

Além disso,

$$\begin{aligned} (1 - \lambda) &= 1 - \frac{z - y}{z - x} \\ &= \frac{z - x - (z - y)}{z - x} \\ &= \frac{y - x}{z - x}. \end{aligned}$$

Se f é convexa em I, então

$$f(y) \leq \lambda f(x) + (1 - \lambda)f(z).$$

Subtraindo $f(x)$ em ambos os lados, tem-se:

$$\begin{aligned} f(y) - f(x) &\leq \lambda f(x) + (1-\lambda)f(z) - f(x) \\ &\leq -(1-\lambda)f(x) + (1-\lambda)f(z) \\ &\leq (1-\lambda)(f(z) - f(x)) \\ &\leq \frac{y-x}{z-x}(f(z) - f(x)) \end{aligned}$$

como $y - x > 0$, se segue:

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

logo, $R(x, y) \leq R(x, z)$.

Caso 2: $y < x < z$. Do fato de que x está entre y e z, tem-se que $x = \lambda y + (1-\lambda)z$. com $\lambda \in (0, 1)$. Analogamente ao Caso 1, é possível mostrar que $\lambda = \frac{z-x}{z-y}$.

Se f é convexa, segue 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 $f(x)$ em ambos os lados, tem-se:

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

segue que

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

logo, $R(x, y) \leq R(x, z)$.

Caso 3: $y < z < x$. Do fato de que z está entre x e y, tem-se que $z = \lambda y + (1-\lambda)x$. com $\lambda \in (0, 1)$. Analogamente é possível mostrar que $\lambda = \frac{x-z}{x-y}$. Se f é convexa, segue 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 $f(x)$ em ambos os lados, tem-se:

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

logo, $R(x,y) \leq R(x,z)$.

Reciprocamente, supondo que para todo $x \in I$, $R(x,y) \leq R(x,z)$, quando $y < z$. Será provado que f é convexa.

Considerando $x = \lambda y + (1-\lambda)z$, para algum $\lambda \in (0,1)$ tem-se que

$$\begin{aligned} x - y &= \lambda y + (1-\lambda)z - y \\ &= -(1-\lambda)y + (1-\lambda)z \\ &= (1-\lambda)(z-y), \end{aligned}$$

e também que

$$\begin{aligned} x - z &= \lambda y + (1-\lambda)z - z \\ &= \lambda y + z - \lambda z - z \\ &= -\lambda(z-y). \end{aligned}$$

Assim, se

$$\frac{f(x)-f(y)}{x-y} \leq \frac{f(x)-f(z)}{x-z}, \tag{2.3}$$

vale para qualquer x, em particular para $x = \lambda y + (1-\lambda)z$, tem-se:

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

Como $z - y > 0$, então:

$$\frac{f(x)-f(y)}{(1-\lambda)} \leq -\frac{f(x)-f(z)}{\lambda},$$

Do fato que $(1-\lambda) > 0$ e $\lambda > 0$, segue que:

$$\begin{aligned}
\lambda f(x) - \lambda f(y) &\leq (\lambda - 1)(f(x) - f(z)) \\
\lambda f(x) - \lambda f(y) &\leq \lambda f(x) - \lambda f(z) - f(x) + f(z) \\
f(x) &\leq \lambda f(y) - \lambda f(z) + f(z) \\
f(\lambda y + (1 - \lambda)z) &\leq \lambda f(y) + (1 - \lambda) f(z)
\end{aligned}$$

Portanto, da última desigualdade se segue que f é convexa.

$\square$

A Proposição 2.2.1 pode ser interpretada dizendo que ao se fixar um dos argumentos, R se torna monotamente não-decrescente no outro argumento. A construção da Figura 14 está disponível em https://ggbm.at/skddbtwh.

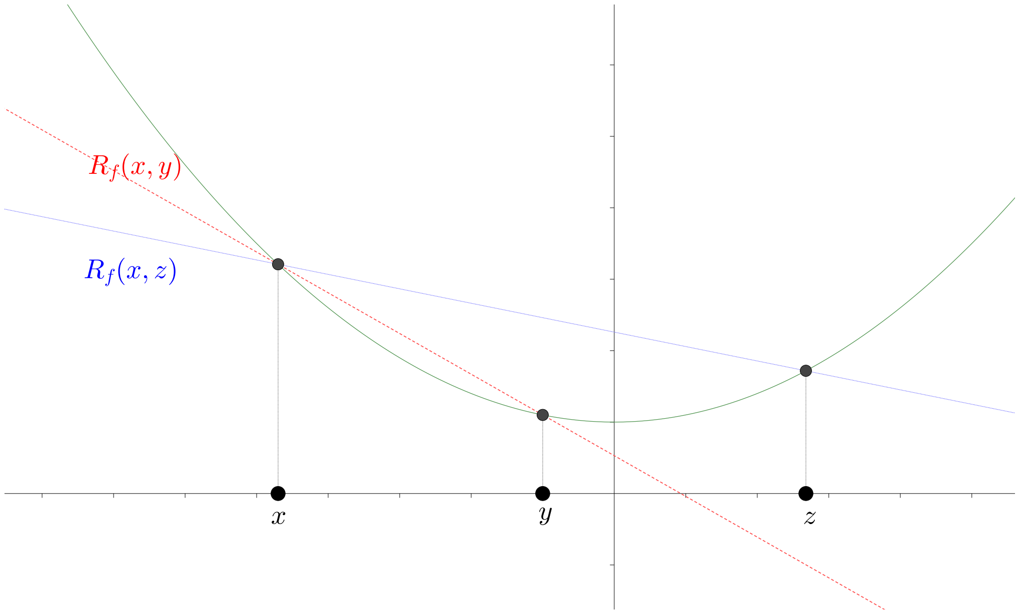

Figura 14 – Caracterização da convexidade por meio da reta secante.

É possivel observar que, fixado x, a reta secante de inclunação $R(x, y)$ (em vermelho) está abaixo da reta secante de inclinação $R(x, z)$ (em azul) para valores maiores que x. Assim, quando uma função é convexa, as retas secantes, a partir de um ponto, possuem inclinações cada vez maiores ou iguais.

2.3 CARACTERIZAÇÃO DE FUNÇÕES CONVEXAS POR MEIO DA DERIVADA

Tendo estudado as relações convexidade/reta secante e considerando que, *no limite, a reta secante é a reta tangente*, esta seção estuda a relação entre a convexidade e a derivada. Os conceitos abordados nesta seção foram estudados na obra de Boyd e Vandenberghe (2004).

O seguinte resultado afirma que se f é convexa, a reta tangente ao gráfico de f num ponto, estará sempre abaixo do gráfico de $f(x)$.

Teorema 3. *Seja $f : I \subset \mathbb{R} \to \mathbb{R}$ uma função diferenciável em um intervalo aberto I. Se f é convexa em I então,*

$$f(x) \geq f(y) + f'(y) \cdot (x - y)$$

para todo $x, y \in I$.

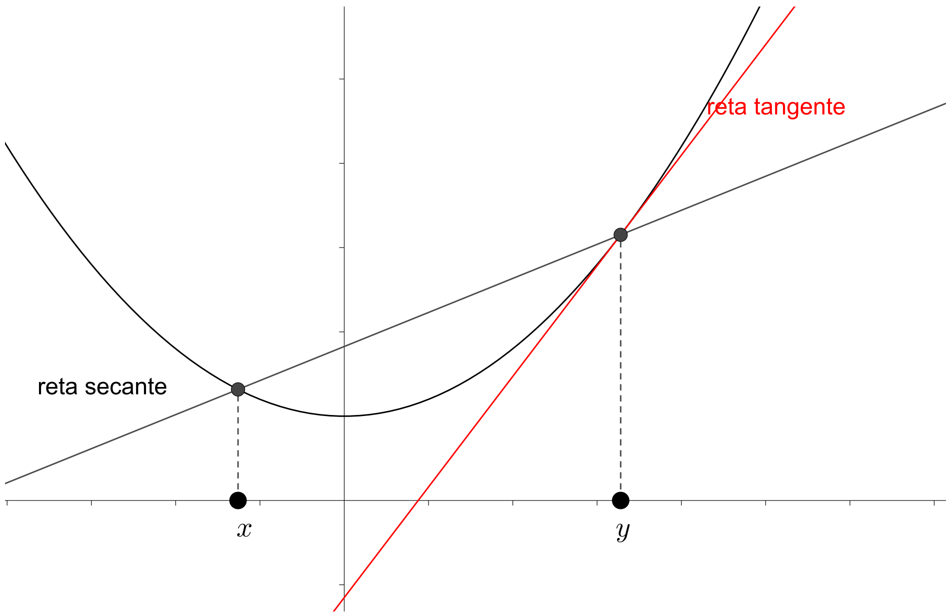

Figura 15 – Relação convexidade, reta tangente e reta secante.

Esta construção encontra-se disponível em https://ggbm.at/k9vz4frv.

Demonstração. Se $x = y$, a desigualdade se satisfaz, pois $f(x) \geq f(y) + f'(y) \cdot 0$. Se $x \neq y$, sem perda de generalidade, é possível considerar $x > y$. Se f é convexa, então, para todo $\lambda \in [0, 1]$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

segue que

$$f(\lambda x + y - \lambda y) \leq \lambda f(x) + f(y) - \lambda f(y)$$

disto

$$f(\lambda(x-y)+y) \leq \lambda(f(x)-f(y)) + f(y). \tag{2.4}$$

Tomando $0 < h < x - y$ e fazendo $\lambda = \frac{h}{x-y}$ e $h = \lambda(x-y)$ tem-se $\lambda \in [0,1]$ e por tanto, de (2.4), se segue:

$$\begin{aligned}
f(y+h) &\leq \frac{h}{x-y}(f(x)-f(y)) + f(y) \\
f(y+h) - f(y) &\leq \frac{h}{x-y}(f(x)-f(y)) \\
\frac{f(y+h)-f(y)}{h} &\leq \frac{f(x)-f(y)}{x-y}
\end{aligned}$$

Como f é diferenciável em y, fazendo h tender a zero, se tem:

$$\begin{aligned}
\lim_{h\to 0}\frac{f(y+h)-f(y)}{h} &\leq \frac{f(x)-f(y)}{x-y} \\
f'(y) &\leq \frac{f(x)-f(y)}{x-y} \\
f'(y)\cdot(x-y) &\leq f(x)-f(y) \\
f(x) &\geq f(y) + f'(y)\cdot(x-y),
\end{aligned}$$

Como se desejava.

□

O Teorema 4 apresenta uma recíproca do Teorema 3, porém fazendo uso de argumentos de convexidade via secante.

Teorema 4. *Seja $f : I \subset \mathbb{R} \to \mathbb{R}$ uma função diferenciável em intervalo aberto I. Se*

$$f(x) \geq f(y) + f'(y) \cdot (x - y)$$

para todo $x, y \in I$. Então f é convexa.

Demonstração. Será feito o uso da Proposição 2.2.1 como base desta demonstração.

Sejam $x, y, z \in I$. Sem perda de generalidade, considere $z > x > y$.

Por um lado,

$$\begin{aligned}
f(y) &\geq f(x) + f'(x)(y - x) && \text{(hipótese)} \\
f(y) - f(x) &\geq f'(x)(y - x) && \text{(subtraíndo} f(x)) \\
\frac{f(y) - f(x)}{y - x} &\leq f'(x) && \text{(pois } \ x - y > 0)
\end{aligned}$$

Por outro lado,

$$\begin{aligned}
f(z) &\geq f(x) + f'(x)(z - x) && \text{(hipótese)} \\
f(z) - f(x) &\geq f'(x)(z - x) && \text{(subtraíndo} f(x)) \\
\frac{f(z) - f(x)}{z - x} &\geq f'(x) && \text{(pois } \ z - x > 0)
\end{aligned}$$

Segue que:

$$\frac{f(y) - f(x)}{y - x} \leq f'(x) \leq \frac{f(z) - f(x)}{z - x} \tag{2.5}$$

Os extremos das desigualdades (2.5) são, por definição, $R_f(x, y)$ e $R_f(x, z)$.

Como $R_f(x, y) \leq R_f(x, z)$, segue da Proposição 2.2.1 que f é convexa.

□

Sobre o Teorema anterior, Bortolossi (2002) afirma que numa função convexa a reta tangente, dada por $T(x) = f(y) + f'(y) \cdot (x - y)$ está sempre abaixo da curva de $f(x)$.

Para ilustrar o teorema anterior foi elaborada uma construção no GeoGebra disponível no link https://www.geogebra.org/m/g2rmF2xE.

Teorema 5. *Seja $f : I \subset \mathbb{R} \to \mathbb{R}$ contínua e diferenciável em I. f é convexa, se e somente se, f' é monótona não-descrescente, ou seja $f'(x) \leq f'(y)$ para todos $x, y \in I$ com $x \leq y$.*

Demonstração. Por um lado, supondo f convexa e, dados quaisquer quatro pontos $x, y, z, w \in I$, que cumpram a desigualdade $x < z < y < w$, pela Proposição 2.2.1, se tem:

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

Fazendo z tender a x, então

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

segue que:

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

e fazendo w tender a y, tem-se

$$f'(x) \leq \lim_{w \to y} \frac{f(w) - f(y)}{w - y},$$

então, tem-se $f'(x) \leq f'(y)$. Portanto f' é não descrescente.

Reciprocamente, sejam $x, y \in I$ e $\lambda \in [0, 1]$. Tomando $z = \lambda x + (1 - \lambda)y$ fica evidente que $x \leq z \leq y$.

Nessas condições, pelo Teorema do Valor Médio para derivadas, existem $a \in (x, z)$ e $b \in (z, y)$ de modo que

$$f'(a) = \frac{f(z) - f(x)}{z - x} \tag{2.6}$$

e

$$f'(b) = \frac{f(y) - f(z)}{y - z}. \tag{2.7}$$

Disto, segue que

$$f(z) - f(x) = f'(a)(z - x) = (1 - \lambda)(y - x)f'(a); \tag{2.8}$$

$$f(y) - f(z) = f'(b)(y - z) = \lambda(y - x)f'(b). \tag{2.9}$$

Como $a \leq b$, pela hipótese $f'(a) \leq f'(b)$. Seque que

$$\begin{aligned}
f(z) - \lambda f(x) - (1-\lambda)f(y) &= (\lambda f(z) - \lambda f(z)) + f(z) - \lambda f(x) - (1-\lambda)f(y) \\
&= \lambda f(z) - \lambda f(x) + (1-\lambda)f(z) - (1-\lambda)f(y) \\
&= \lambda(f(z) - f(x)) + (1-\lambda)(f(z) - f(y)) \\
\text{(De (2.8) e (2.9))} \quad &= \lambda(1-\lambda)(y-x)(f'(a) - f'(b)) \\
&\leq 0;
\end{aligned}$$

Portanto, $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)y$, mostrando a convexidade de f.

□

Corolário 2.3.1. *Seja $f : I \subset \mathbb{R}$ uma função contínua e não-decrescente com I um intervalo convexo. Então a função g dada por*

$$g(x) = \int_{x_0}^{x} f(t)dt$$

com $x_0 \in I$, é convexa.

Demonstração. Como f é contínua, pelo Teorema Fundamental do Cálculo, g também o é, além disso, $g'(x) = f(x)$ para todo $x \in I$ e $g'(x)$ é não decrescente, logo pelo Teorema 5, g é convexa. □

Corolário 2.3.2. *Seja $f : I \subset \mathbb{R}$ contínua e, duas vezes, diferenciável em I. A função f é convexa se e somente se $f''(x) \geq 0$ para todo $x \in I$.*

Demonstração. Como f é duas vezes diferenciável, então pelo Teorema 5, f' é não decrescente. Reciprocamente, uma função não decrescente jamais terá derivadas negativas, assim, $f''(x) \geq 0$. □

Exemplo 2.3.3. A função $f(x) = -\ln(x)$ é convexa, pois $f''(x) = \frac{1}{x^2} > 0$.

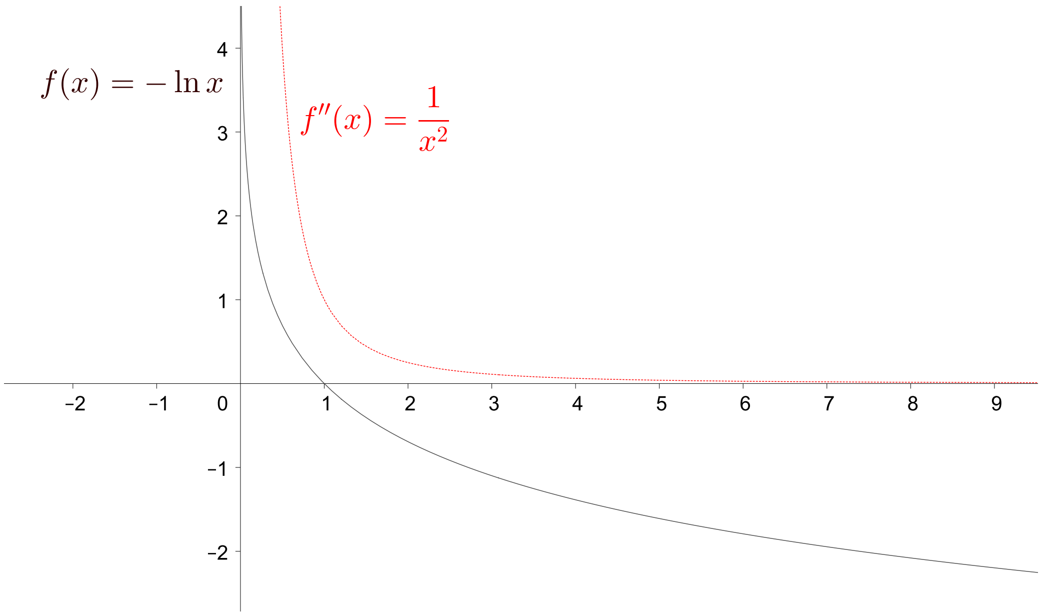

Figura 16 – Teste da segunda derivada para $f(x) = -\ln(x)$.

Exemplo 2.3.4. A função $f(x) = x^4 + x^2$ é convexa, pois $f'(x) = 4x^3 + 2x$ é não decrescente, satisfazendo o Teorema 5 e $f''(x) = 12x^2 + 2 \geq 0$ satisfazendo as condições do Corolário 2.3.2.

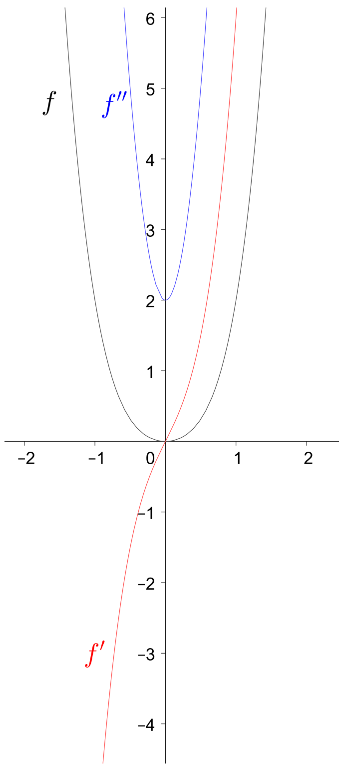

Figura 17 – Estudo da convexidade de $f(x) = x^4 + x^2$.

A função $f(x) = x^4 + x^2$ representada em preto é convexa, sua derivada $f'(x) = 4x^3 + 2x$, em vermelho, é não decrescente e $f''(x) = 12x^2 + 2$, em azul, é sempre positiva.

A figura que segue ilustra o Corolário 2.3.1 em que $f(x) = \ln x$ (em preto) representa uma função não decrescente e $g(x) = x \ln x - x$ (em vermelho), uma primitiva.

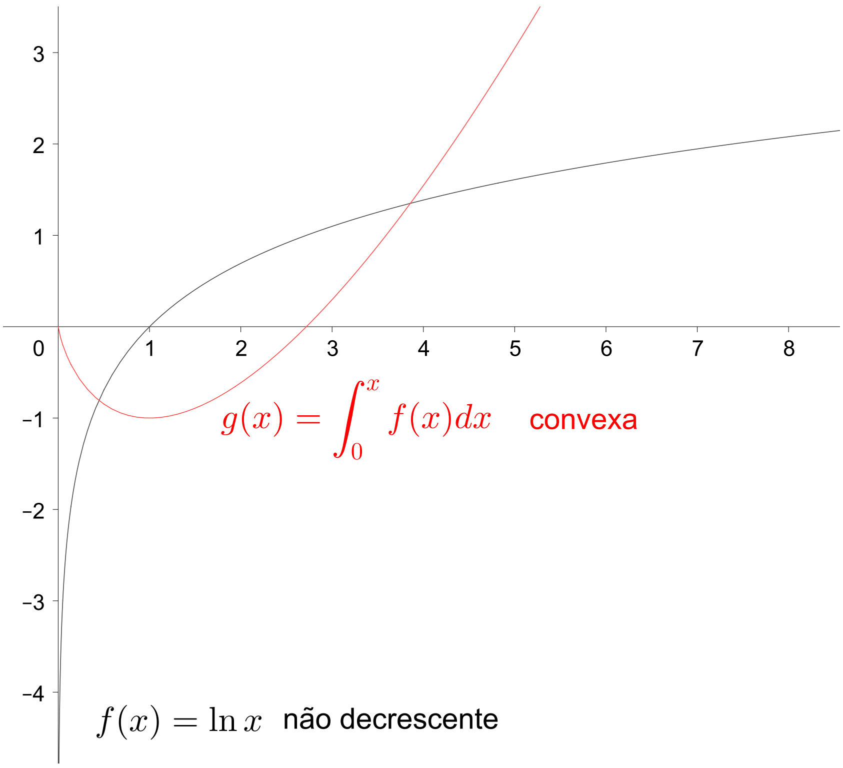

Figura 18 – Convexidade da integral de uma função não descrescente.

O resultado que segue é uma versão do Teorema 2.3.5 adicionando a hipótese de diferenciabilidade das funções. Optou-se por manter as duas versões devido a natureza diferente de suas demonstrações.

Corolário 2.3.5. *Seja $f : I \subset \mathbb{R} \to \mathbb{R}$ uma função convexa, diferenciável e invertível, então:*

(i) Se f é crescente, sua inversa é côncava;

(ii) Se f é decrescente, sua inversa é convexa.

Demonstração. Seja $f : I \subset \mathbb{R} \to \mathbb{R}$ uma função convexa e invertível, então, pelo Teorema 5, para $x_1, x_2 \in I$ com $x_1 < x_2$, se tem que

$$f'(x_1) \leq f'(x_2). \tag{2.10}$$

Seja g a inversa de f e $f(x_i) = y_i$, então pela regra da cadeia para derivadas,

$$f'(x_i) = \frac{1}{g'(y_i)}. \tag{2.11}$$

Segue de (2.10) e (2.11) que

$$\frac{1}{g'(y_1)} \leq \frac{1}{g'(y_2)}$$

Como f é crescente, g também será pois suas derivadas têm o mesmo sinal, logo, g' será positiva, disto,

$$g'(y_1) \geq g'(y_2). \tag{2.12}$$

Para mostrar o item (i), se supõe que f é crescente, ou seja, quando $x_1 < x_2$, se tem que $y_1 < y_2$, de (2.12), se conclui que, neste caso, g' é monótona não-crescente. Portanto, g é côncava.

Para mostrar o item (ii), se supõe que f é decrescente, ou seja, quando $x_1 < x_2$, se tem que $y_1 > y_2$, de (2.12), se conclui que, neste caso, g' é monótona não-decrescente. Portanto, g é convexa.

$\square$

2.3.1 CONVEXIDADE E DERIVADAS LATERAIS

Nem toda função é diferenciável, porém, isto não implica que ela não possa ser convexa, exemplo claro disso é a função $f(x) = |x|$ que não é diferenciável em $x = 0$, porém é, claramente, convexa.

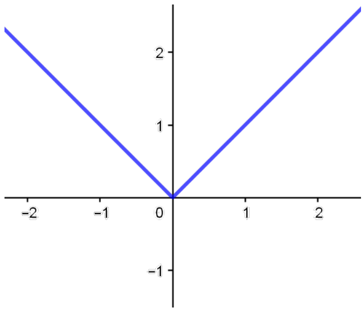

Figura 19 – $f(x) = |x|$

Embora a função não seja diferenciável em $x = 0$, as suas derivadas laterais existem.

Definição 6. *Seja $f : I \subset \mathbb{R}$. As derivadas laterais de f, denotadas por f'_+ e f'_-, são definidas como:*

$$f'_+ := \lim_{h \to 0^+} \frac{f(x+h)-f(x)}{h}, \quad e$$
$$f'_- := \lim_{h \to 0^-} \frac{f(x+h)-f(x)}{h}$$

A seguinte proposição, cuja discussão mais formal está disponível no Teorema 1.3.3 da obra de Niculescu e Persson (2004), estabelece uma relação entre convexidade e derivadas laterais.

Proposição 2.3.6. *Se $f : I \subset \mathbb{R}$ é convexa, então f é contínua em I, além disso, possui derivadas laterais em cada ponto de I e é válida a desigualdade:*

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

com $x, y \in I$ e $x < y$.

A proposição anterior afirma que f'_- e f'_+ são funções monotonamente não-decrescentes.

3 DESIGUALDADES ENTRE MÉDIAS

Este capítulo consiste em apresentar diversas demonstrações de desigualdades entre médias usando como ferramental algumas funções convexas.

As demonstrações de desigualdades como

$$\frac{x_1 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \cdots x_n},$$

em geral, são apresentadas de maneira complexa, analisando muitos casos particulares e, muitas vezes incompletas. No livro de Matemática Discreta Morgado e Carvalho (2015) da coleção do PROFMAT há uma demonstração para essa desigualdade fazendo a prova para 2 termos, comenta a demonstração para 4 termos, sugere utilizar a indução para $n = 2^k$ e termina fazendo para 3 termos usando o resultado de $n = 4$ e sugerindo uma similar estratégia para os demais termos. As demonstrações de outras desigualdades entre médias são deixadas a cargo do leitor como exercício. Neste capítulo esta e outras demonstrações serão feitas via convexidade usando a Proposição 2.1.4, que caracteriza a convexidade de uma função por n pontos.

As ideias apresentadas neste capítulo foram exploradas pelo autor, de maneira independentes, a partir dos conceitos de convexidade e médias clássicas presentes em obras do PROFMAT. Entretanto, esses tópicos podem ser estudados, de maneira mais aprofundada, na obra de Bullen (2003).

3.1 MÉDIAS CLÁSSICAS ENTRE TERMOS POSITIVOS

Na defnição que segue são apresentadas as médias clássicas.

Definição 7. *Dados $x_1, \ldots, x_n$ números reais positivos, se definem:*

- *Média aritmética*

$$\bar{X} := \frac{x_1 + \cdots + x_n}{n}$$

- *Média geométrica*

$$\bar{G} := \sqrt[n]{x_1 \cdots x_n}$$

- *Média harmônica*

$$\bar{H} := \frac{n}{\frac{1}{x_1} + \cdots + \frac{1}{x_n}}$$

- *Média quadrática*

$$\bar{Q} := \sqrt{\frac{x_1^2 + \cdots + x_n^2}{n}}$$

Ao longo desse capítulo os termos $x_1, \ldots, x_n$ serão considerados reais positivos, exceto se for explicitado algum caso especial distinto.

Exemplo 3.1.1. Será demonstrada a desigualdade entre as médias geométrica e aritmética. Considerando $\lambda_i = \frac{1}{n}$ e a função $f(x) = -\ln(x)$. Da convexidade e monotonia de f, apresentada no Exemplo 2.3.3 e da Proposição 2.1.4 seguem as desigualdades:

$$\begin{aligned}
f(\lambda_1 x_1 + \cdots + \lambda_n x_n) &\leq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n) \\
-\ln\left(\frac{x_1 + \cdots + x_n}{n}\right) &\leq -\frac{1}{n}\left(\ln(x_1) + \cdots + \ln(x_n)\right) \\
\ln\left(\frac{x_1 + \cdots + x_n}{n}\right) &\geq \frac{1}{n}\ln(x_1 x_2 \cdots x_n) \\
&\geq \ln((x_1 \cdots x_n)^{\frac{1}{n}}) \\
\frac{x_1 + \cdots + x_n}{n} &\geq \sqrt[n]{x_1 \cdots x_n} \\
\bar{X} &\geq \bar{G}.
\end{aligned}$$

Exemplo 3.1.2. Será demonstrada a desigualdade entre a média harmônica e geométrica. Considerando, ainda, $f(x) = -\ln x$, tomando $\lambda_i = \frac{1}{n}$ e $u_i = \frac{1}{x_i}$, para $i = 1, 2, \ldots$, segue que:

$$
\begin{aligned}
f(\lambda_1 u_1 + \cdots + \lambda_n u_n) &\leq \lambda_1 f(u_1) + \cdots + \lambda_n f(u_n) \\
-\ln\left(\frac{1}{n}u_1 + \cdots + \frac{1}{n}u_n\right) &\leq \frac{-1}{n}\ln u_1 + \cdots + \frac{-1}{n}\ln u_n \\
-\ln\left(\frac{u_1 + \cdots + u_n}{n}\right) &\leq \ln u_1^{-1/n} + \cdots + \ln u_n^{-1/n} \\
&\leq \ln\sqrt[n]{\frac{1}{u_1}} + \cdots + \ln\sqrt[n]{\frac{1}{u_n}} \\
\ln\left(\frac{n}{u_1 + \cdots + u_n}\right) &\leq \ln\left(\sqrt[n]{\frac{1}{u_1}} \cdots \sqrt[n]{\frac{1}{u_n}}\right) \\
\frac{n}{\dfrac{1}{x_1} + \cdots + \dfrac{1}{x_n}} &\leq \sqrt[n]{x_1 \dots x_n} \\
\bar{H} \leq \bar{G}.
\end{aligned}
$$

Exemplo 3.1.3. Será demonstrada a desigualdade entre a média aritmética e a quadrática. Tomando a função convexa $f(x) = x^2$ e $\lambda_i = \frac{1}{n}$ para $i = 1, 2, \dots,$ então:

$$
\begin{aligned}
f(\lambda_1 x_1 + \cdots + \lambda_n x_n) &\leq \lambda_1 f(x_1) + \dots \lambda_n f(x_n) \\
\left(\frac{x_1 + \cdots + x_n}{n}\right)^2 &\leq \frac{x_1^2 + \cdots + x_n^2}{n} \\
\frac{x_1 + \cdots + x_n}{n} &\leq \sqrt{\frac{x_1^2 + \cdots + x_n^2}{n}} \\
\bar{X} \leq \bar{Q}
\end{aligned}
$$

Dos exemplos anteriores se conclui a desigualdade

$$
\bar{H} \leq \bar{G} \leq \bar{X} \leq \bar{Q}.
$$

3.2 SEQUÊNCIAS DE p-MÉDIAS

Algumas das médias anteriormente citadas podem ser escritas de maneira mais geral usando a seguinte definição:

Definição 8. *Sejam $x_1, \dots, x_n$ números reais positivos e $p \in \mathbb{Z}, p \neq 0$. A p-média deles, denotada por X_p, é definida como:*

$$X_p(x_1, \dots, x_n) := \left(\frac{x_1^p + \cdots + x_n^p}{n}\right)^{\frac{1}{p}}.$$

Essas médias, em inglês, recebem o nome **power mean**, conforme consta da obra de Bullen (2003).

As médias harmônica, aritmética e quadrática são casos particulares da p-média, considerando $p = -1$, $p = 1$ e $p = 2$, respectivamente.

Proposição 3.2.1. *A p-média, com $p > 1$ é maior do que a média aritmética, ou seja, $X_1 \leq X_p$.*

Demonstração. Tomando $f(x) = x^p$, com $p \in \mathbb{N}$ e $\lambda_i = \frac{1}{n}$ com $i = 1, 2, \dots$, como x^p é convexa em $[0, +\infty)$, então então:

$$\begin{aligned} f(\lambda_1 x_1 + \cdots + \lambda_n x_n) &\leq \lambda_1 f(x_1) + \dots \lambda_n f(x_n) \\ \left(\frac{x_1 + \cdots + x_n}{n}\right)^p &\leq \frac{x_1^p + \cdots + x_n^p}{n} \\ \frac{x_1 + \cdots + x_n}{n} &\leq \sqrt[p]{\frac{x_1^p + \cdots + x_n^p}{n}}. \end{aligned}$$

Pontanto, $X_1 \leq X_p$. □

As médias X_1 e X_p serão iguais se $p = 1$ ou quando $x_1 = x_2 = \cdots = x_n$.

Com base nas desigualdades mostradas na primeira seção desse capítulo e na Proposição 3.2.1, é possível se indagar se a seguinte sequência de desigualdades é válida:

$$\bar{X_{-1}} \leq \bar{G} \leq X_1 \leq X_2 \leq \cdots \leq X_p.$$

Parte desta desigualdade é mostrada na proposição que segue.

Proposição 3.2.2. *A sequência de médias $X_1, X_2, \dots, X_n$ é monótona não decrescente, ou seja, $X_p \leq X_{p+1}$ para todo $p = 2, 3, \dots$.*

Demonstração. Seja a função $f(x) = x^{\frac{p+1}{p}}$, com $p \in \mathbb{R}^+$. Como $f''(x) = \frac{p+1}{p^2}x^{\frac{1-p}{p}} \geq 0$ então f é convexa. Tomando $\lambda_i = \frac{1}{n}$ e $u_i = x_i^p$ para $i = 1, 2, \ldots,$ então:

$$
\begin{aligned}
\left(\frac{u_1 + \cdots + u_n}{n}\right)^{\frac{p+1}{p}} &\leq \frac{1}{n}u_1^{\frac{p+1}{p}} + \cdots + \frac{1}{n}u_n^{\frac{p+1}{p}} \\
\left(\frac{u_1 + \cdots + u_n}{n}\right)^{\frac{1}{p}} &\leq \left(\frac{1}{n}u_1^{\frac{p+1}{p}} + \cdots + \frac{1}{n}u_n^{\frac{p+1}{p}}\right)^{\frac{1}{p+1}} \\
\sqrt[p]{\frac{x_1^p + \cdots + x_n^p}{n}} &\leq \left(\frac{x_1^{p+1} + \cdots + x_n^{p+1}}{n}\right)^{\frac{1}{P+1}} \\
X_p &\leq X_{p+1}
\end{aligned}
$$

Como se desejava mostrar. □

Proposição 3.2.3. *A sequência $X_1, X_2, \ldots, X_p$ é limitada superiormente por $x_M := \max\{x_1, \ldots, x_n\}$.*

Demonstração. Por definição, $x_i \leq x_M$ para $i = 1, 2, \ldots,$ então $x_i^p \leq x_M^p$, disto

$$
\begin{aligned}
x_1^p + \cdots + x_n^p &\leq \underbrace{x_M^p + \cdots + x_M^p}_{n \text{ vezes}} \\
\frac{x_1^p + \cdots + x_n^p}{n} &\leq \frac{x_M^p + \cdots + x_M^p}{n} \\
\sqrt[p]{\frac{x_1^p + \cdots + x_n^p}{n}} &\leq \sqrt[p]{\frac{x_M^p + \cdots + x_M^p}{n}}.
\end{aligned}
$$

No entanto,

$$
\sqrt[p]{\frac{x_M^p + \cdots + x_M^p}{n}} = \sqrt[p]{\frac{nx_M^p}{n}} = x_M.
$$

Portanto, $X_p \leq x_M$ □

Note que se todos os termos são iguais, a sequência converge para x_M, caso sejam diferentes, a Proposição 3.2.3 garante que a sequência de médias é limitada por x_M. Além disso, a sequência de médias é não decrescente. Assim, se conclui que essa sequência é convergente pelo Teorema 4 do Capítulo 3 da obra Lima (2016). No entanto, ainda não há garantia de que essa sequência esteja, de fato, convergindo para x_M. Essa convergência é demonstrada na Proposição 3.2.5. Para auxiliar nessa demonstração se faz necessário o Lema 3.2.4.

Lema 3.2.4. *Seja* $f : \mathbb{R}^+ \to \mathbb{R}^+$*, tal que* $\lim_{p\to\infty} f(p) = 0$ *e seja* $k > 0$*, então*

$$\lim_{p\to\infty} (k + f(p))^{1/p} = 1.$$

Demonstração. De fato, da continuidade de e^x e $\ln(x)$ segue que

$$\begin{aligned}
\lim_{p\to\infty} (k + f(p))^{\frac{1}{p}} &= \lim_{p\to\infty} e^{\ln(k + f(p))^{\frac{1}{p}}} \\
&= \lim_{p\to\infty} e^{\frac{1}{p}\ln(k + f(p))} \\
&= e^{\lim_{p\to\infty} \frac{1}{p} \cdot \lim_{p\to\infty} \ln(k + f(p))} \\
&= e^{0 \cdot \ln(k + \lim_{p\to\infty} f(p))} \\
&= e^{0 \cdot \ln k} = 1.
\end{aligned}$$

□

Proposição 3.2.5. *A sequência de médias* $X_1, X_2, \ldots, X_p$ *converge para* x_M*, ou seja,*

$$\lim_{p\to\infty} X_p = x_M.$$

Demonstração. Essa prova consiste em verificar que

$$\lim_{p\to\infty} \frac{X_p}{x_M} = \frac{x_M}{x_M} = 1$$

ou seja, mostrar que vale o limite

$$\lim_{p\to\infty} \left(\frac{x_1^p + \cdots + x_n^p}{n x_M^p} \right)^{\frac{1}{p}} = 1$$

De fato, sem perda de generalidade, como há pelo menos um $x_i = x_M$, pode se considerar os primeiros k termos iguais a x_M, tem-se que

$$\left(\frac{x_1^p+\cdots+x_n^p}{nx_M^p}\right)^{\frac{1}{p}} = \left(\frac{k\cdot x_M^p+x_{k+1}^p+\cdots+x_n^p}{nx_M^p}\right)^{\frac{1}{p}} \tag{3.1}$$

$$= \left(\frac{1}{n}\right)^{\frac{1}{p}}\left(\frac{kx_M^p}{x_M^p}+\left(\frac{x_{k+1}}{x_M}\right)^p+\cdots+\left(\frac{x_n}{x_M}\right)^p\right)^{\frac{1}{p}} \tag{3.2}$$

$$= \left(\frac{1}{n}\right)^{\frac{1}{p}}\left(k+\left(\frac{x_{k+1}}{x_M}\right)^p+\cdots+\left(\frac{x_n}{x_M}\right)^p\right)^{\frac{1}{p}}. \tag{3.3}$$

Como $0 < x_i < x_M$ para $i > k$, então segue que $0 < \frac{x_i}{x_M} < 1$, logo $\lim_{p\to\infty}(\frac{x_i}{x_M})^p = 0$. Tomando $f(p) = \left(\frac{x_{k+1}}{x_M}\right)^p + \cdots + \left(\frac{x_n}{x_M}\right)^p$ segue que

$$\lim_{p\to\infty} f(p) = 0.$$

De (3.3) tem-se que

$$\lim_{p\to\infty}\left(\frac{x_1^p+\cdots+x_n^p}{nx_M^p}\right)^{\frac{1}{p}} = \lim_{p\to\infty}\left(\frac{1}{n}\right)^{\frac{1}{p}}(k+f(p))^{\frac{1}{p}},$$

pelo fato de que $\lim_{p\to\infty}\left(\frac{1}{n}\right)^{\frac{1}{p}=1}$ e pelo Lema 3.2.4, então

$$\lim_{p\to\infty}\left(\frac{x_1^p+\cdots+x_n^p}{nx_M^p}\right)^{\frac{1}{p}} = \lim_{p\to\infty}\left(\frac{1}{n}\right)^{\frac{1}{p}}\cdot 1 = 1.$$

como se desejava demonstrar. □

Em resumo, temos o seguinte resultado:

Corolário 3.2.6. *Sejam $x_1,\ldots,x_n \in \mathbb{R}^+$, então*

$$\bar{X_{-1}} \leq \bar{G} \leq X_1 \leq X_2 \leq \cdots \leq X_p \leq \cdots \leq x_M;\ \ e$$

$$\lim_{p\to\infty} X_p = x_M.$$

O resultado que segue expande o Corolário 3.2.6 para $p = -1, -2, \ldots$

Proposição 3.2.7. *Sejam $x_1,\ldots,x_n \in \mathbb{R}^+$, então*

$$x_m \leq \cdots \leq X_{-p} \cdots \leq X_{-1} \leq \bar{G} \leq X_1 \leq \cdots \leq X_p \leq \cdots \leq x_M; \textit{ e}$$

$$\lim_{p \to -\infty} X_p = x_m; \lim_{p \to \infty} X_p = x_M.$$

Demonstração. O Corolário 3.2.6 demonstra para $p = 1, 2, \ldots$ as desigualdades desejadas.

Agora, considerando a sequência de médias $X_{-1}, X_{-2}, \ldots, X_{-p} \ldots$, tem-se que

$$X_{-p} = \left(\frac{x_1^{-p} + \cdots + x_n^{-p}}{n} \right)^{\frac{1}{-p}} = \frac{1}{\left(\frac{u_1^p + \cdots + u_n^p}{n} \right)^{\frac{1}{p}}} = \frac{1}{U_p},$$

com $u_i = \frac{1}{x_i}$ para $i = 1, 2, \ldots, n$.

Note que U_p é uma sequência de p-médias de termos positivos, logo, pelo Corolário 3.2.6 para $p = 1, 2, \ldots, n$ se satisfaz que:

- $U_p \leq U_{p+1}$;
- $U_p \leq u_M$ e
- $\lim_{p \to \infty} U_p = u_M$.

Como $X_{-p} = \frac{1}{U_p}$ e $u_M = \frac{1}{x_m}$ de modo que $x_m := \min\{x_1, \ldots, x_n\}$, tem-se que, para $p = 1, 2, 3, \ldots, n$

- $X_{-p} \geq X_{-(p+1)}$;
- $X_{-p} \geq x_m$ e
- $\lim_{p \to \infty} X_{-p} = x_m$.

Portanto, tem-se a sequência de desigualdades:

$$x_m \leq \cdots \leq X_{-p} \cdots \leq X_{-1} \leq \bar{G} \leq X_1 \leq \cdots \leq X_p \leq \cdots \leq x_M \qquad (3.4)$$

obtendo, assim, o resultado desejado.

□

Uma conjectura plausível seria supor que na desiguldade (3.4), $\bar{G} = X_0$, entretanto X_p não é definida para $p = 0$, assim uma possibilidade seria que

$$\lim_{p\to 0}\left(\frac{x_1^p + \cdots + x_n^p}{n}\right)^{\frac{1}{p}} = \sqrt[n]{x_1 \cdots x_n}.$$

Esse fato é demonstrado na seguinte proposição:

Proposição 3.2.8. *Seja* $x_1, \ldots, x_n \in \mathbb{R}^+$, *então*

$$\lim_{p\to 0}\left(\frac{x_1^p + \cdots + x_n^p}{n}\right)^{\frac{1}{p}} = \sqrt[n]{x_1 \cdots x_n}.$$

Demonstração. Considerando $g(p) = \ln \frac{x_1^p + \cdots + x_n^p}{n}$, calcularemos $g'(0)$ de duas maneiras diferentes, uma pelas regras operacionais de derivação e outra por definição via limite.

De fato, pelas regras de derivação

$$g'(p) = \frac{n}{x_1^p + \cdots + x_n^p}\left(\frac{x_1^p \ln x_1 + \cdots + x_n^p \ln x_n}{n}\right),$$

aplicando em $p = 0$ temos

$$g'(0) = \frac{n}{n}\left(\frac{\ln x_1 + \cdots + \ln x_n}{n}\right) = \frac{1}{n}\ln(x_1 \cdots x_n) = \ln \sqrt[n]{x_1 \cdots x_n}. \tag{3.5}$$

Por outro lado, aplicando a definição de derivada em $p = 0$ tem-se que

$$g'(0) = \lim_{p\to 0}\frac{g(p) - g(0)}{p - 0} \tag{3.6}$$

$$= \lim_{p\to 0}\frac{\ln\left(\frac{x_1^p + \cdots + x_n^p}{n}\right) - \ln\left(\frac{x_1^0 + \cdots + x_n^0}{n}\right)}{p} \tag{3.7}$$

$$g'(0) = \lim_{p\to 0}\frac{1}{p}\ln\left(\frac{x_1^p + \cdots + x_n^p}{n}\right) \tag{3.8}$$

Das equações (3.5) e (3.8) tem-se que

$$
\begin{aligned}
\lim_{p\to 0}\frac{1}{p}\ln\left(\frac{x_1^p+\cdots+x_n^p}{n}\right) &= \ln\sqrt[n]{x_1\cdots x_n}\\
e^{\lim\limits_{p\to 0}\frac{1}{p}\ln\left(\frac{x_1^p+\cdots+x_n^p}{n}\right)} &= e^{\ln\sqrt[n]{x_1\cdots x_n}}\\
\lim_{p\to 0} e^{\ln\left(\frac{x_1^p+\cdots+x_n^p}{n}\right)^{\frac{1}{p}}} &= e^{\ln\sqrt[n]{x_1\cdots x_n}}\\
\lim_{p\to 0}\left(\frac{x_1^p+\cdots+x_n^p}{n}\right)^{\frac{1}{p}} &= \sqrt[n]{x_1\cdots x_n}.
\end{aligned}
$$

□

Pelo resultado anterior, podemos definir

$$X_0 := \lim_{p\to 0} X_p = \bar{G}.$$

Definindo $X_{-\infty}$ e X_{∞} como

$$X_{-\infty} := \lim_{p\to -\infty} X_p = x_m \ \text{ e}$$

$$X_{\infty} := \lim_{p\to \infty} X_p = x_M.$$

Os resultados deste capítulo são resumidos no seguinte teorema:

Teorema 6. *Sejam* $x_1,\ldots,x_n\in\mathbb{R}^+$ *então*

$$X_{-\infty}\le\cdots\le X_{-p}\cdots\le X_{-1}\le X_0\le X_1\le\cdots\le X_p\le\cdots\le X_{\infty}.$$

4 DECOMPOSIÇÃO DE FUNÇÕES COMO DIFERENÇA DE FUNÇÕES CONVEXAS

Como foi comentado no segundo capítulo, as funções convexas não formam um espaço vetorial, no entanto, pode ser obtido um espaço vetorial a partir de funções convexas e côncavas. Surge a pergunta: Quais funções estão nesse espaço vetorial? De maneira equivalente: Para quais funções $f : [a, b] \to \mathbb{R}$, é possível determinar g, h convexas em $[a, b]$, de modo que $f = g - h$?

Essa questão é abordada de maneira mais precisa e formal na obra de Robertz e Varberg (1973) (Teorema 14A), usando o conceito de variação total de uma função. Segue uma versão desse teorema, adaptada para o contexto deste trabalho .

Teorema 7. *Seja $f : [a, b] \to \mathbb{R}$, então f é a diferença entre duas funções convexas definidas em $[a, b]$ se e somente se,*

$$f(x) - f(a) = \int_a^x r(t)dt$$

para alguma função r de variação total limitada em $[a, b]$.

Derivando a expressão do Teorema 7, obtem-se $f'(x) = r(x)$, isto é, uma função pode ser escrita como diferença entre duas funções convexas quando sua derivada é uma função de variação total limitada.

Em seguida é feita uma breve discussão a respeito da variação total de uma função em um intervalo e uma exploração do problema de decomposição de funções em funções convexas.

4.1 VARIAÇÃO TOTAL DE UMA FUNÇÃO

Intuitivamente a variação total de uma função num intervalo, pode ser interpretada como o *acúmulo* de todas as variações verticais (em y) conforme x varia num intervalo $[a, b]$.

Definição 9. *Seja a função $f : I \to \mathbb{R}$ com $[a, b] \subset I$. A* ***variação total*** *de f de a até b é definida como:*

$$V_a^b(f) := \sup_{\mathcal{P}} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$$

em que $\mathcal{P}$ é o conjunto de todas as partições $\{x_0, x_1, \ldots, x_n\}$ do intervalo $[a, b]$.

Uma interpretação da definição: a partir da partição do intervalo $[a, b]$ em diversos pequenos intervalos $[x_i, x_{i+1}]$, o termo $|f(x_i) - f(x_{i-1})|$ nos dá, em módulo, a variação que f teve em $[x_i, x_{i+1}]$. Caso f seja monótona em $[a, b]$ sua variação total é $V_a^b = |f(b) - f(a)|$. Em contrapartida, caso f tenha um comportamento não monótono em algum intervalo, é possível particioná-lo de modo que em cada um desses sub-intervalos f seja monótona.

O propósito do supremo na definição é, justamente, garantir que estejamos sempre tomando a variação total.

Em geral essa variação pode ser obtida com o auxílio dos pontos críticos de f. Nos intervalos em que a função é crescente, o acréssimo da função coincide com o acréssimo na variação, em contrapartida, nos intervalos em que a função é descrescente, os decréscimos da função representam um acréssimo na variação.

O seguinte teorema, demonstrado em Stein e Shakarchi (2005) (p. 137), serve de instrumento para calcular a variação total de funções diferenciáveis em algum intervalo.

Teorema 8. *Seja a função $f : I \to \mathbb{R}$ com $[a, b] \subset I$, f' contínua em $[a, b]$. A* ***variação total*** *de f pode ser calculada por*

$$V_a^x(f) = \int_a^x |f'(t)|dt$$

Note que $V_a^x(f)$ é não decrescente pois $(V_a^x(f))' = |f'(x)| \leq 0$. Integrais similares à do Teorema 8 são utilizadas em Cálculo Vetorial, por exemplo, para calcular comprimentos de arcos, entretanto, neste contexto, apenas uma variável é considerada no cálculo da variação total.

Supondo $r(t) = (0, f(t))$ uma curva parametrizada. O comprimento, S, do arco r com $t \in [a, x]$ é obtido a partir da integral:

$$S = \int_a^x |r'(t)|dt = \int_a^x \sqrt{0^2 + (f'(t))^2}dt = \int_a^x |f'(t)|dt = V_a^x(f(x))$$

Exemplo 4.1.1. A variação total de $f(x) = x^2$ em $[-1, 2]$ pode ser calculada da seguinte forma:

$$V_{-1}^2(x^2) = \int_{-1}^2 |2x|dx = \int_{-1}^0 -2x\ dx + \int_0^2 2x\ dx = -x^2|_{-1}^0 + x^2|_0^2 = 1 + 4 = 5.$$

Uma contrução dinâmica desenvolvida no software GeoGebra sobre este fato pode ser encontrada em <https://ggbm.at/tw72vtev> e ilustrada na figura seguinte

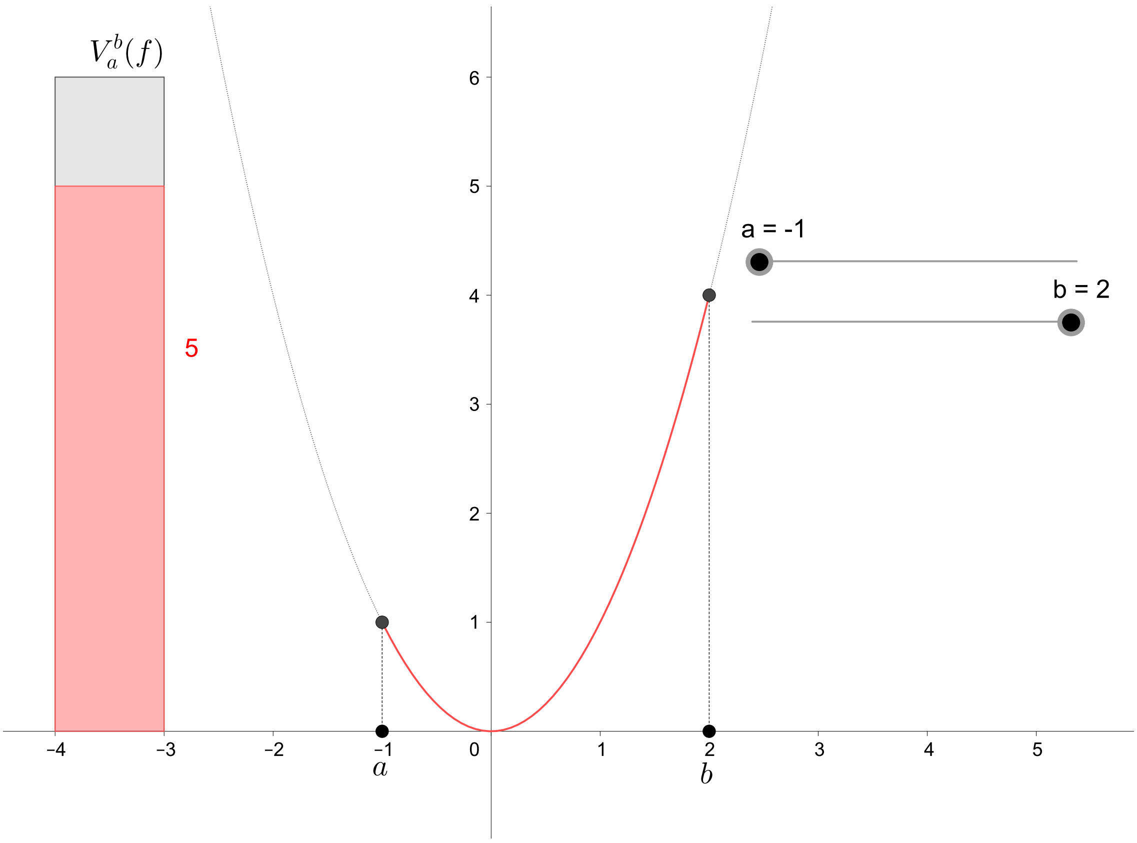

Figura 20 – Ilustração no GeoGebra da variação total de $f(x) = x^2$ em $[-1, 2]$.

A barra vertical presente na Figura 20 representa um "indicador"de variação total de $f(x)$ à medida que variam os controles deslizantes a e b.

Exemplo 4.1.2. Variação total de $f(x) = x^3 + 2x^2$ em $[-2, 1]$. Ora, f é crescente em $[-2, \frac{-4}{3}]$ e em $[0, 1]$ e descrescente em $[-\frac{4}{3}, 0]$.

$$
\begin{aligned}
V_{-2}^{1}(f) &= \int_{-2}^{1} |f'(x)|dx = \int_{-2}^{1} |3x^2+4x|dx \\
&= \int_{-2}^{-\frac{4}{3}} 3x^2+4xdx + \int_{-\frac{4}{3}}^{0} -3x^2-4xdx + \int_{0}^{1} 3x^2+4xdx \\
&\approx 5,37.
\end{aligned}
$$

Na Figura 21 o contador de variação total à esquerda se estrutura do mesmo modo do contador da Figura 20.

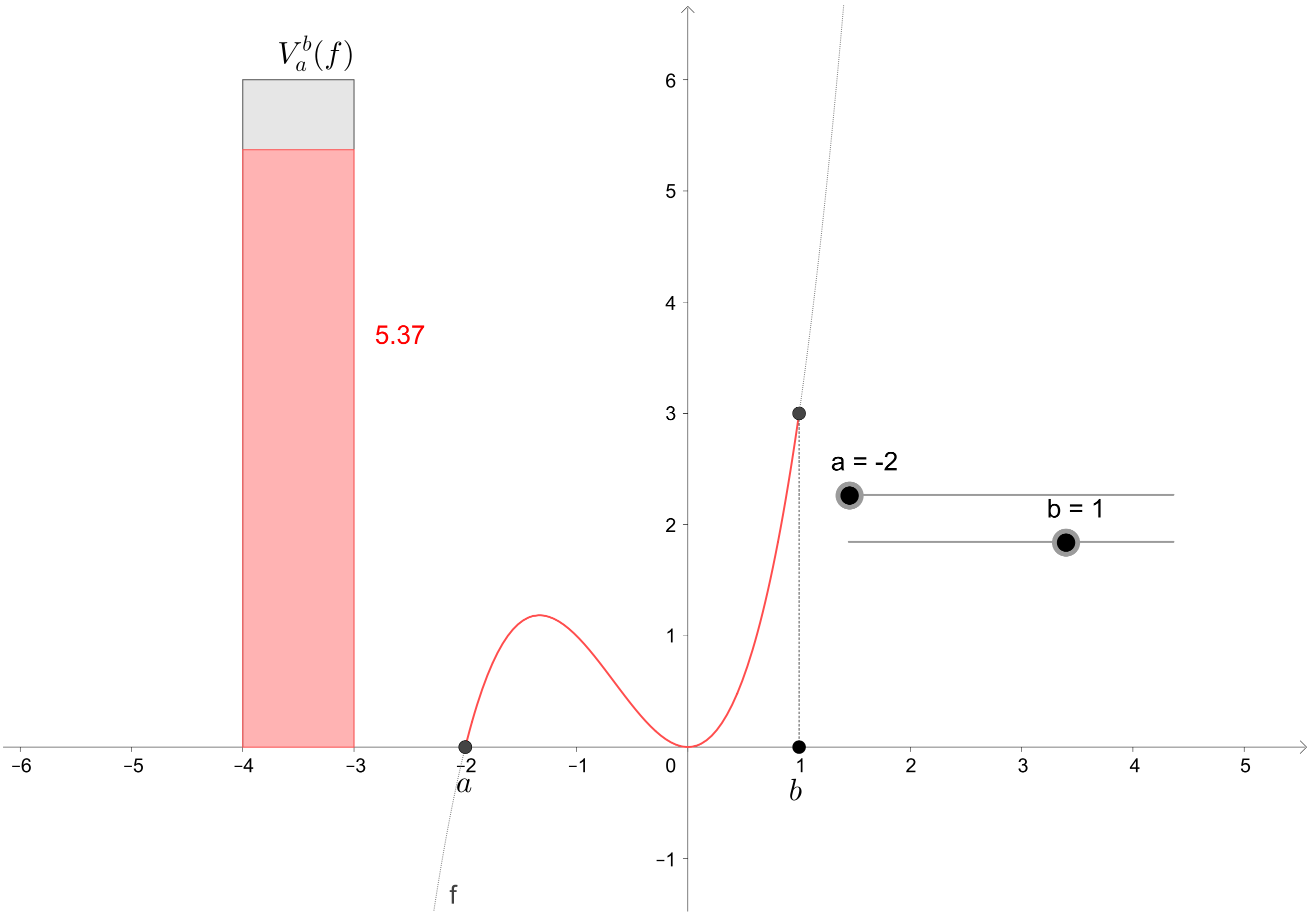

Figura 21 – Ilustração no GeoGebra da variação total de $f(x) = x^3 + 2x^2$ em $[-2, 1]$.

Exemplo 4.1.3. A variação total de $f(x) = \cos x$ em $[0, 2\pi]$ é dada por:

$$\int_{0}^{2\pi} |-\sin x|dx = \int_{0}^{\pi} \sin xdx - \int_{\pi}^{2\pi} \sin xdx = 4$$

Caso queira se obter a variação total de $\cos x$ em $[0, x]$ com $0 \leq x \leq 2\pi$, é preciso analisar os casos:

Se $0 \leq x \leq \pi$, tem-se que

$$V_0^x(f(x)) = \int_0^x |-\sin t| dt = -\cos t|_0^x = 1 - \cos x.$$

Se $\pi \leq x \leq 2\pi$, tem-se

$$V_0^x(f(x)) = \int_0^x |-\sin t| dt = \int_0^\pi \sin t dt + \int_\pi^x -\sin(t) dt = 2 + \cos t|_\pi^x = 3 + \cos x.$$

Em resumo, se $f(x) = \cos x$ e $x \in [0, 2\pi]$, então:

$$V_0^x(f(x)) = \begin{cases} 1 - \cos x & \text{se } 0 \leq x \leq \pi \\ 3 + \cos x & \text{se } \pi \leq x \leq 2\pi. \end{cases}$$

Exemplo 4.1.4. Algumas funções não possuem variação total limitada, é o caso de

$$f(x) = \begin{cases} \sin(\frac{1}{x}) & \text{se } x \neq 0 \\ 0 & \text{se } x = 0. \end{cases}$$

A função não tem variação total limitada no intervalo $[-1, 1]$ por exemplo, pois próximo de zero as oscilações da função faz a variação total tender ao infinito.

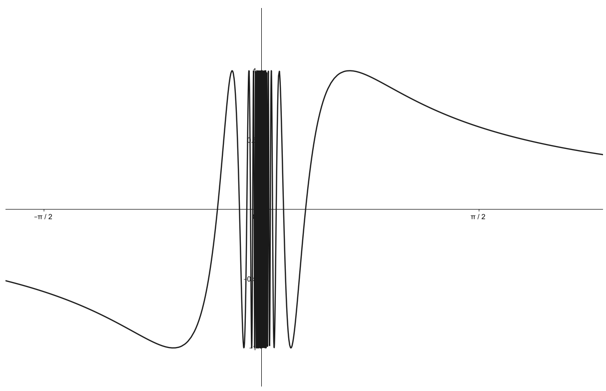

Figura 22 – $f(x) = \sin(\frac{1}{x})$

4.2 DECOMPOSIÇÃO DE UMA FUNÇÃO COMO DIFERENÇA DE FUNÇÕES CONVEXAS

Como mencionado no início do capítulo, uma função pode ser escrita como diferença de funções convexas se sua derivada é de variação total limitada. Nesta seção serão consideradas $f : [a,b] \to \mathbb{R}$ uma função duas vezes diferenciável, f' de variação total limitada e f'' contínua.

Note que

$$f' = V_a^x(f') - [V_a^x(f') - f']. \tag{4.1}$$

A função $V_a^x(f') = \int_a^x |f''(t)|dt$ é não decrescente. A função $V_a^x(f') - f'$ também é não decrescente, pois

$(V_a^x(f') - f')' = |f''(x)| - f''(x) \geq 0$

Pelo Corolário (2.3.1), são convexas em $[a,b]$ as funções

$$g(x) = \int_a^x V_a^t(f')dt \ \text{ e}$$

$$h(x) = \int_a^x [V_a^t(f') - f']dt.$$

Integrando a expressão 4.1 em ambos os lados, tem-se que $f = g - h$, ou seja f pode ser decomposta como a diferença entre duas funções convexas.

Desta forma, é possível enumerar 5 passos para determinar esta decomposição:

1. Calcular a derivada de f em relação a x;
2. Calcular $V_a^x(f')$;
3. Integrar $V_a^x(f')$ em relação a x. Essa será uma das funções convexas desejadas e será denotada por g;
4. Calcular $h = g - f$;
5. Verificar que h é convexa.

Exemplo 4.2.1. Seguindo os passos acima, é obtida a decomposição de $f(x) = \sin x$ com $x \in [0, 2\pi]$ como a diferença entre duas funções convexas.

Sabe-se que $f'(x) = \cos x$ e, além disso, como visto no Exemplo 4.1.3,

$$V_a^x(f') = \begin{cases} 1 - \cos x & 0 \leq x < \pi \\ 3 + \cos x & \pi \leq x < 2\pi. \end{cases}$$

Como $V_a^x(f')$ é não-decrescente, $g(x) = \int_0^x V_a^t(f')dt$ é convexa.

$$g(x) = \begin{cases} x - \sin x & 0 \leq x < \pi \\ 3x + \sin x - 2\pi & \pi \leq x < 2\pi. \end{cases}$$

A constante de integração 2π, no segundo caso de g, foi determinada para garantir a continuidade da função.

A figura que segue representa graficamente a função g. Em vermelho o caso em que $g(x) = x - \sin x$ e em laranjado o caso em que $g(x) = 3x + \sin x - 2\pi$.

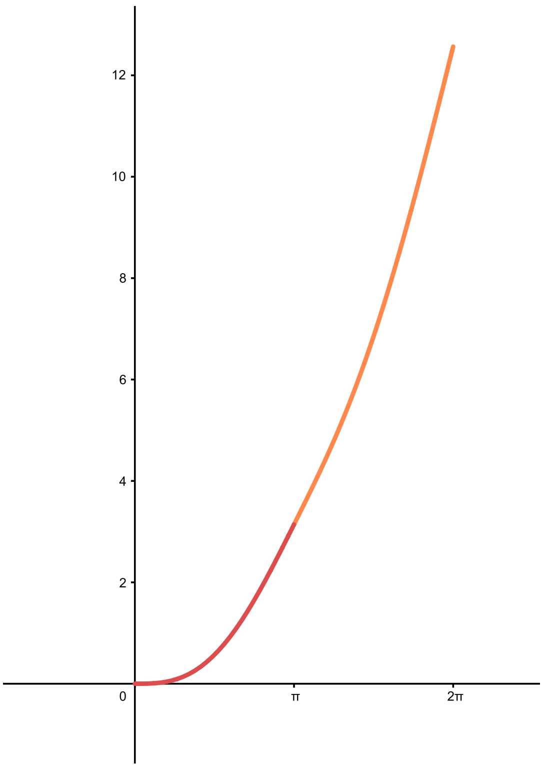

Figura 23 – Representação gráfica da função $g(x)$.

No passo que segue se calcula h, de forma que $f = g - h$.

$$h(x) = \begin{cases} x - 2\sin x & 0 \leq x < \pi \\ 3x - 2\pi & \pi \leq x < 2\pi. \end{cases}$$

Por fim, basta verificar via derivação que h é, de fato, convexa.

$$h'(x) = \begin{cases} 1 - 2\cos x & 0 \leq x < \pi \\ 3 & \pi \leq x < 2\pi, \end{cases}$$

Disto, como $\cos x$ é descrescente em $[0, \pi]$ então h' é não decrescente. Portanto h é convexa.

A figura que segue apresenta $g(x)$ em vermelho e $h(x)$ em azul.

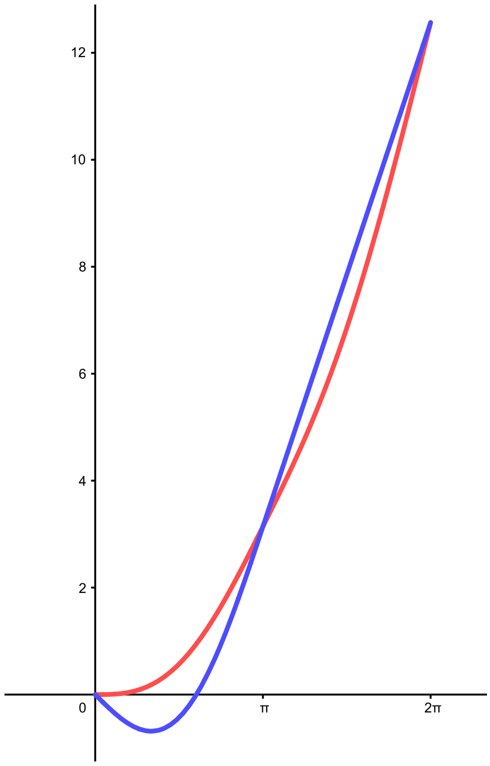

Figura 24 – Representação gráfica de das funções $g(x)$ e $h(x)$

Uma construção dinâmica está disponível no link https://ggbm.at/znchyw7r.

Esse procedimento não funciona sempre, por exemplo quando a variação total da função não pode ser calculada ou não é bem definida por problemas de diferenciabilidade.

5 CONSIDERAÇÕES FINAIS

Esse trabalho permitiu desenvolver um estudo introdutório aos conjuntos e funções convexas. Foram estudados diversos conceitos, tais como combinações convexas e envoltório convexo. Foram estudados alguns resultados referentes às funções convexas, caracterizando-as via reta secante e por meio da sua derivada.

Foram discutidos argumentos, via convexidade, para demonstrar algumas desigualdades entre médias, presentes em livros da coleção do PROFMAT. Em seguida foram estudadas as p-Médias e apresentadas algumas desigualdades entre elas utilizando o conceito de convexidade.

Por fim, foi explorado o problema da decomposição de funções reais com diferença entre convexas.

O desenvolvimento desse trabalho possibilitou um crescimento profissional, visto que as discussões feitas em cada etapa estiveram fora da minha zona de conforto. As reflexões realizadas no decorrer do processo, serviram no sentido de questionar mais as informações presentes nos livros, pois pude perceber que, um mesmo enunciado pode ser escrito de diversas maneiras e, dependendo do contexto, alguns são mais adequados que outros.

O conteúdos das funções está presente em praticamente todo o ensino médio e, muitas vezes, é trabalhado sem o devido cuidado com elementos básicos como domínio e imagem, ficando, muitas vezes, restrito o estudo da lei de formação e a interpretação gráfica, apenas. Ao longo deste trabalho pude fazer algumas reflexões sobre esses elementos e sua importancia. Estas reflexões terão um impacto direto nas minhas aulas visto que, independentemente do formalismo exigido em cada nível de ensino, o fato de conhecer um conceito em um nível mais avançado, permite discutir esse mesmo conceito de forma adequada para diferentes níveis de complexidades.

Futuramente, pretende-se avançar nos estudos de convexidade desenvolvendo representações que ilustrem os resultados de funções com mais variáveis e noutros contextos. Almeja-se, também, estudar problemas de minimização na Programação Convexa.

REFERÊNCIAS

AMORIN, R. G. de. **Introdução à Análise Convexa**. [S.l.]: Dissertação de Mestrado. 79p. Instituto de Matemática e Estatística, Universidade Federal de Goiás, 2013.

BARATA, J. C. A. **Notas para um Curso de Física-Matemática**. [S.l.]: Departamento de Física Matemática - USP, Disponível em: <http://denebola.if.usp.br/ jbarata/Notas_de_aula/capitulos.html> Último acesso 23/11/2018, 2018.

BORTOLOSSI, H. J. **Cálculo Diferencial a Várias Variáveis: Uma Introdução à Teoria da Otimização**. [S.l.]: PUC Loyola, 2002.

BOYD, S.; VANDENBERGHE, L. **Convex Optimization**. 2. ed. [S.l.]: CAMBRIDGE UNIVERSITY PRESS, 2004.

BULLEN, P. S. **Handbook of Means and Their Inequalities**. 2. ed. [S.l.]: Springer-Science+Business Media, BV, 2003.

KRANTZ, S. G. **Convex Analysis**. 6000 Broken Sound Parkway NW, Suite 300: [s.n.], 2015.

LIMA, E. L. **Funções de Uma Variável**. Estrada Dona Castorina, 110 Jardim Botânico | CEP 22460-320 Rio de Janeiro, RJ - Brasil: IMPA, 2016. v. 1. (1, v. 1).

MORGADO, A. C.; CARVALHO, P. C. P. **Matemática Discreta**. 2. ed. [S.l.]: SBM, 2015.

NICULESCU, C. P.; PERSSON, L.-E. **Convex Functions and Their Applications: A contemporary approach**. [S.l.]: Monograph, 2004.

ROBERTZ, A. W.; VARBERG, D. E. **Convex Functions**. [S.l.]: Academic Press, 1973.

STEIN, E. M.; SHAKARCHI, R. **Real Análisis**. 1. ed. Princeton, New Jersey 08540: PRINCETON UNIVERSITY PRESS, 2005.