Universidade Federal de Sergipe
Centro de Ciências Exatas e Tecnologia
Programa de Pós–Graduação em Matemática
Mestrado em Matemática

Funções Convexas e as Transformadas de Legendre e Fenchel

Lucas de Melo Pontes e Silva

São Cristóvão – SE
Fevereiro de 2021

Universidade Federal de Sergipe
Centro de Ciências Exatas e Tecnologia
Programa de Pós–Graduação em Matemática
Mestrado em Matemática

Funções Convexas e as Transformadas de Legendre e Fenchel

por

Lucas de Melo Pontes e Silva

sob a orientação do

Prof. Dr. José Anderson Valença Cardoso

São Cristóvão – SE
Fevereiro de 2021

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE

Silva, Lucas de Melo Pontes e
S586f Funções convexas e as transformadas de Legendre e Fenchel / Lucas de Melo Pontes e Silva ; orientador José Anderson Valença Cardoso. – São Cristóvão, 2021.
124 f.; il.

Dissertação (mestrado em Matemática) – Universidade Federal de Sergipe, 2021.

1. Funçoes convexas. 2. Funçoes de Legendre. 3. Banach, Espaços de. 4. Hilbert, Espaço de. I. Cardoso, José anderson Valença orient. II. Título.

CDU 5

UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Dissertação submetida à aprovação pelo Programa de Pós-Graduação em Matemática da Universidade Federal de Sergipe, como parte dos requisitos para obtenção do grau de Mestre em Matemática.

Funções Convexas e as Transformadas de Legendre e Fenchel

por

Lucas de Melo Pontes e Silva

Aprovada pela banca examinadora:

José Anderson V. Cardoso

Prof. José Anderson Valença Cardoso - UFS
Orientador

Almir Rogério Silva Santos

Prof. Almir Rogerio Silva Santos - UFS
Primeiro Examinador

Uberlandio Batista Severo

Prof. Uberlandio Batista Severo - UFPB
Segundo Examinador

São Cristóvão, 05 de Fevereiro de 2021

Cidade Universitária “Prof. José Aloísio de Campos” – Av. Marechal Rondon, s/no - Jardim Rosa Elze
– Campus de São Cristóvão. Tel. (00 55 79) 3194-6887
CEP: 49100-000 - São Cristóvão – Sergipe - Brasil – E-mail: promat.ufs@gmail.com

Resumo

O presente trabalho aborda os principais elementos da análise convexa em espaços vetoriais de dimensões finita e infinita. Em dimensão finita, introduz-se conceitos básicos sobre espaços vetoriais e topologia de conjuntos para desenvolver a teoria dos conjuntos convexos. Então define-se os conjuntos convexos e suas propriedades apresentando exemplos de operações que preservam convexidade, conjuntos convexos clássicos e o importante teorema da separação por hiperplano. Em seguida, o trabalho apresenta as funções convexas e suas propriedades, das quais podemos destacar a continuidade em subconjuntos abertos e a existência da derivada direcional. O arcabouço teórico desenvolvido permite apresentar a transformada de Legendre para o caso de funções convexas de classe C^1 e a transformada de Fenchel para o caso de funções convexas não suaves. Apresenta-se aplicações da transformada de Legendre, em especial, na formulação de equações da mecânica clássica além uma tabela com funções e transformadas. Em dimensão infinita, introduz-se conceitos topológicos e propriedades de espaços métricos, continuidade, Teorema de Bolzano-Weierstrass, espaços de Hilbert e Banach e o Teorema de Hahn-Banach. O trabalho segue definindo pontos interiores, conjuntos e funções convexas em espaços de Hilbert, definindo importantes propriedades, em especial, a existência da conjugada nesse espaço. Por fim, apresenta-se aplicação da desigualdade de Jensen para resolução de problemas olímpicos do Ensino Médio.

Palavras-chave: Análise convexa, Conjuntos convexos, Funções convexas, Transformada de Legendre e Fenchel. Espaços de Banach e Hilbert.

Abstract

The present work addresses the main elements of convex analysis in vector spaces of finite and infinite dimensions. In finite-dimension, it presents fundamental concepts of norms, inner-product, and topology. Then, it defines convex sets and explores their properties. It shows operations that preserve convexity, classic convex sets, and the hyperplane separation theorem. Next, the work presents the convex functions and their properties, from which we can highlight the continuity in open subsets and the existence of the directional derivative. The theoretical framework developed allows presenting the Legendre transform when the convex functions are C^1 and the Fenchel transform for non-smooth convex functions. Among all applications of the Legendre transform, this work highlights the formulation of equations of classical mechanics. A table with selected smooth convex functions and their respective Legendre transform is shown. In infinite dimension, the work develops topological concepts and properties of metric spaces, continuity, Bolzano-Weierstrass theorem, Hilbert and Banach spaces, and Hahn-Banach theorem. Then, it defines interior points, convex sets, and convex functions in Hilbert spaces, defining main properties, especially the existence of the conjugate function in this space. Finally, it shows an application of Jensen's inequality to solve High School Olympic problems.

Keywords: Convex analysis. Convex sets. Convex functions. Legendre transform. Fenchel tranform. Banach and Hilbert spaces.

Sumário

Introdução

Na natureza são observados diversos princípios que estão relacionados com a minimização de grandezas físicas, tais como: caminhos, trabalho e energia. Em [15] os autores citam no seu prefácio que em 1650 foi postulado por Fermat que um raio de luz percorre o caminho entre dois pontos A e B, que minimiza o tempo. Problemas como esses são classificados como problemas de valores críticos ou extremos. Ainda, de acordo com [15], houve um grande desenvolvimento da análise matemática, principalmente pelos trabalhos de Newton e Leibniz com a descoberta do Teorema Fundamental do Cálculo, por volta de 1670. Em 1750, Euler estendeu os conceitos prévios estabelecidos por Fermat, sobre as condições necessárias para estabelecimento de extremos de funções do tipo:

$$J(y) = \int_a^b L(t, y(t), y'(t))dt$$

conhecidas como equação de Euler-Lagrange. A partir daí, progressos na determinação de condições necessárias para apresentação de pontos críticos para funções foram realizados por Legendre, Jacobi, Weierstrass, Volterra e Hadamard.

Mais recentemente, vem sendo estudado as condições de existência de soluções para problemas extremos, ou seja, quais as condições que devem ser estabelecidas para provar a existência de pontos críticos de funções.

A análise convexa é um importante ramo da matemática, que trabalha com conjuntos e funções convexas e suas propriedades analíticas e geométricas. Há diversas propriedades importantes relacionadas a continuidade, diferenciabilidade e também estudos de mínimos ou máximo globais de funções. Suas aplicações são diversas, abrangendo os ramos da economia, física, engenharia e computação, tendo em vista a aplicação na busca por otimizações (maximização ou minimização), bem como, a resolução de equações diferenciais parciais.

De acordo com [1], a primeira definição de convexidade (para curvas e superfícies) foi dada por Arquimedes (287-212 a.C.), que estava interessado no estudo de áreas e volumes. No trabalho em [3], descreve-se que entre as diferentes propriedades obtidas por Arquimedes sobre convexidade, destaca-se os postulados e os resultados referentes ao centro de gravidade de conjuntos planos e a descrição dos 13 poliedros semirregulares, também conhecidos como sólidos arquimedianos. Ao longo do tempo houve algumas contribuições sobre convexidade, porém apenas no final do século XIX que observa-se diversos resultados de grande importância em convexidade, como o trabalho de Minkowski sobre o estudo aprofundado de conjuntos convexos em dimensão finita. A definição de função convexa é dada por Jensen, que demonstra várias desigualdades conhecidas, decorrentes das propriedades dessas funções. Já o interesse real pela geometria convexa tem como marco inicial o trabalho de Bonnensen e Fenchel no livro *Theorie der Konvexen Körper* de 1934. Em análise funcional destaca-se algumas propriedades interessantes de conjuntos convexos em espaços vetoriais de dimensão infinita. Muitos resultados da análise não-linear estão baseados na convexidade de conjuntos e de funções.

No Capítulo 1, apresentaremos alguns resultados básicos que serão utilizados para o desenvolvimento dos capítulos posteriores com destaque para conceitos de espaço vetoriais, norma, produto interno e topologia no $\mathbb{R}^n$. O trabalho será desenvolvido no $\mathbb{R}^n$, com exceção do Capítulo 6, em que será trabalhado conceitos em espaços vetoriais de dimensão infinita.

No Capítulo 2, apresentaremos os conceitos de conjuntos afim, convexos, suas propriedades, exemplos e alguns resultados fundamentais com destaque para o teorema do hiperplano suporte e da separação por hiperplano.

No Capítulo 3, apresentaremos as funções convexas, côncavas, suas propriedades e alguns exemplos, com destaque para a caracterização das funções convexas, a propriedade de continuidade e a definição de semicontinuidade.

No Capítulo 4, apresentaremos a definição da transformada de Legendre, propriedades, exemplos de aplicações com destaque para a apresentação de uma tabela com as transformadas de Legendre para funções de classe C^1.

No Capítulo 5, apresentaremos a transformada de Fenchel, introduzindo os conceitos importantes, com destaque para o teorema de Fenchel-Moreau e aplicações.

No Capítulo 6, apresentaremos alguns conceitos topológicos e introduziremos conjuntos e funções convexas para o caso de espaços vetoriais de dimensão infinita, apresentando alguns resultados importantes da análise funcional com destaque para o teorema de Hanh-Banach.

No Capítulo 7, apresentaremos a resolução de problemas de desigualdades em nível olímpico do Ensino Médio, utilizando o conceito de funções convexas.

No Capítulo 8, consolidamos os principais resultados desse trabalho e apresentamos uma proposição de continuidade e trabalhos futuros.

Capítulo 1

Preliminares

Neste capítulo serão apresentados e desenvolvidos conceitos e ferramentas básicas que serão utilizadas ao longo do trabalho.

Por simplificação, desenvolveremos este trabalho no espaço euclidiano $\mathbb{R}^n$, sabendo que várias das propriedades aqui definidas podem ser estendidas para outros espaços vetoriais de dimensão finita ou infinita.

Os conceitos, as definições e os teoremas desse capítulo foram consultados nas referências [11, 16].

1.1 Conceitos Básicos de Espaços Vetoriais

Definição 1.1. *X é um espaço vetorial sobre um corpo $\mathbb{K}$, quando X é um conjunto munido com as operações de soma e multiplicação por escalar, fechado nessas operações e com as seguintes propriedades:*

A1. Para quaisquer $u, v \in X$, então $u + v = v + u$.

A2. Para quaisquer $u, v, w \in X$, então $u + (v + w) = (u + v) + w$.

A3. Existe $0 \in X$, tal que $0 + u = u$.

A4. Para todo $u \in X$, existe $-u \in X$, tal que $u + (-u) = 0$.

M1. $\alpha(\beta u) = (\alpha\beta)u$, para todos $\alpha, \beta \in \mathbb{K}$ e $u \in X$.

M2. $(\alpha + \beta)u = \alpha u + \beta u$, para todos $\alpha, \beta \in \mathbb{K}$ e $u \in X$.

M3. $\alpha(u + v) = \alpha u + \alpha v$, para todos $\alpha \in \mathbb{K}$ e $u, v \in X$.

M4. $1(u) = u$ para todo $u \in X$, onde 1 é o elemento neutro da multiplicação do corpo $\mathbb{K}$.

Neste trabalho utilizaremos, em grande parte, o corpo dos reais, que será a definição padrão, quando não mencionado o contrário.

Definição 1.2. *Um subconjunto $S \subseteq X$ é um subespaço vetorial de X, quando é fechado com respeito à soma e à multiplicação por escalar, ou seja, para todos $u, v \in S$ e $\lambda \in \mathbb{K}$, tem-se $u + \lambda v \in S$.*

Definição 1.3. *A norma definida num espaço vetorial X sobre o corpo dos reais é uma função $\|\cdot\| : X \to \mathbb{R}$, que satisfaz as seguintes propriedades:*

N1. $\|u\| \geq 0$, *para todo* $u \in X$, *e* $\|u\| = 0 \iff u = 0$.

N2. $\|\lambda u\| = |\lambda| \|u\|$, *para todo* $u \in X$ *e* $\lambda \in \mathbb{R}$.

N3. $\|u + v\| \leq \|u\| + \|v\|$, *para todo* u, $v \in X$.

Observação 1.1. *O par $(X, \|\cdot\|)$, que significa que o espaço vetorial X está munido da norma, é chamado de espaço vetorial normado. Dado $u = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$, a norma euclidiana é definida por:*

$$\|u\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}. \tag{1.1}$$

A norma euclidiana, por vezes, é denotada por $\|\cdot\|_2$. Entretanto, por simplificação de notação, quando não expresso de forma contrária a norma $\|\cdot\|$ será a norma euclidiana. No espaço vetorial $\mathbb{R}^n$ podemos definir as seguintes normas:

$$\|u\|_1 = |x_1| + ... + |x_n| \tag{1.2}$$

e

$$\|u\|_\infty = \max\{|x_1|, ..., |x_n|\}. \tag{1.3}$$

Verifica-se que $\|u\|_2$, $\|u\|_1$ e $\|u\|_\infty$, são de fato normas, pois satisfazem as propriedades N1, N2 e N3. Adicionalmente, observa-se que:

$$\|u\|_\infty \leq \|u\|_2 \leq \|u\|_1 \leq n \|u\|_\infty.$$

Na verdade, no $\mathbb{R}^n$ existem infinitas normas e todas elas são equivalentes.

Definição 1.4. *O produto interno definido num espaço vetorial X sobre o corpo dos reais é uma função $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$, que satisfaz as seguintes propriedades:*

PI1. $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$, *para todos* u, v *e* $w \in X$.

PI2. $\langle u, \lambda v \rangle = \lambda \langle u, v \rangle$, *para todos* u, $v \in X$ *e* $\lambda \in \mathbb{R}$.

PI3. $\langle u, v \rangle = \langle v, u \rangle$, *para todos* u, $v \in X$.

PI4. $\langle u, u \rangle \geq 0$, *para todo* $u \in X$ *e* $\langle u, u \rangle = 0 \iff u = 0$.

A definição acima pode ser generalizada ao trabalhar no espaço dos números complexos $\mathbb{C}$, ou seja, $\langle .,. \rangle : X \times X \to \mathbb{C}$, substituindo $PI3$ com $\langle u, v \rangle = \overline{\langle v, u \rangle}$. O produto interno nos complexos é denominado Hermitiano.

O produto interno canônico de vetores $u = (x_1, x_2, ..., x_n)$, $v = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$, é definido como:

$$\langle u, v \rangle = \sum_{i=1}^{n} x_i y_i. \tag{1.4}$$

Verifica-se que a definição acima satisfaz *PI1, PI2, PI3 e PI4*, sendo de fato um produto interno. Além disso, tem-se a seguinte relação entre norma e produto interno, para todo $u \in \mathbb{R}^n$: $\langle u, u \rangle = \|u\|^2 = u^T u$, onde u é representado como um vetor coluna ou como uma matriz $n \times 1$ no espaço euclidiano $\mathbb{R}^n$.

Para a seção seguinte, assumiremos conceitos básicos de conjuntos e operações com conjuntos no espaço euclidiano $\mathbb{R}^n$.

1.2 Conceitos Topológicos Básicos em Espaços Euclidianos de Dimensão Finita

Definição 1.5. *O conjunto $B \subset \mathbb{R}^n$ definido por $B = \{u \in \mathbb{R}^n \colon \|u\| < 1\}$ é conhecido como a bola aberta unitária do espaço euclidiano $\mathbb{R}^n$.*

Por vezes denominamos $B = B(0, 1)$. Caso queiramos definir uma vizinhança aberta de um elemento $u \in \mathbb{R}^n$, dado $\epsilon > 0$, tem-se que $B(u, \epsilon) = u + \epsilon B$ é a vizinhança de u representada por uma bola aberta centrada em u de raio ϵ. Adicionalmente, denotamos pela bola fechada, o conjunto $B[u, \epsilon] = \{u \in \mathbb{R}^n \colon \|u\| \leq \epsilon\}$. Por exemplo, no $\mathbb{R}^2$ as bolas unitárias com respeito às normas euclidiana, norma da soma e norma do máximo são mostradas na figura a seguir:

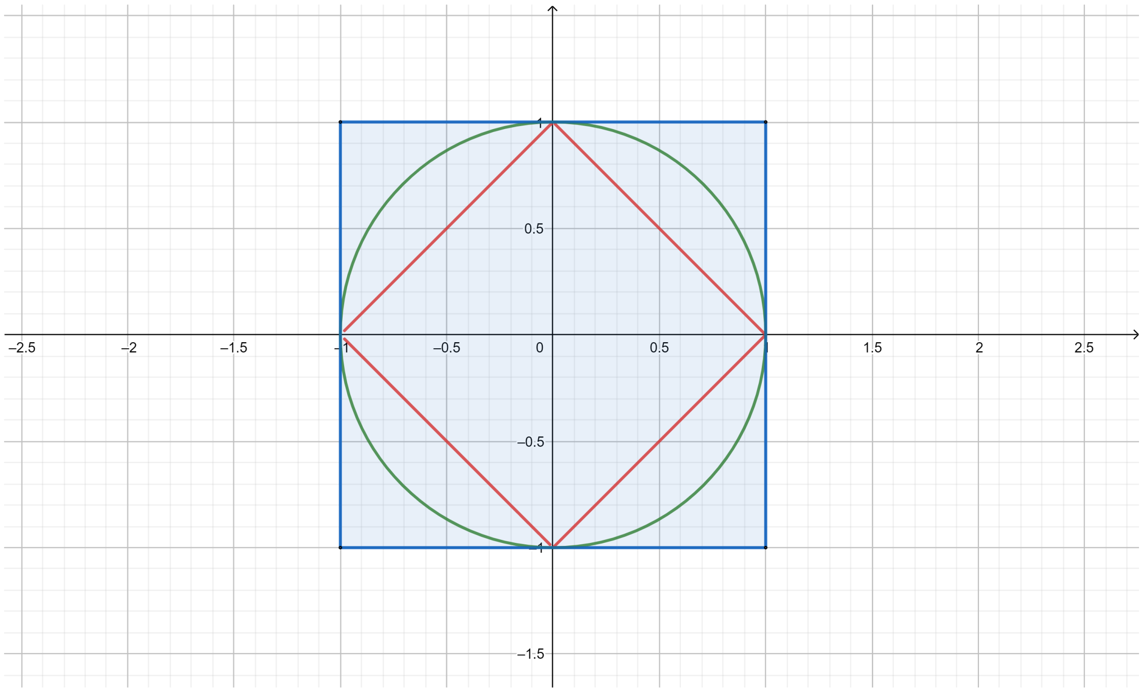

Figura 1.1: Em azul bola unitária no $\mathbb{R}^2$ utilizando a norma do máximo, em verde a bola unitária no $\mathbb{R}^2$ utilizando a norma euclidiana e em vermelho a bola unitária no $\mathbb{R}^2$, utilizando a norma da soma.

Definição 1.6. *O subconjunto $X \subset \mathbb{R}^n$ é limitado, quando existe uma bola aberta em $\mathbb{R}^n$ que contém X, ou seja, quando existe $r \in \mathbb{R}$, $r > 0$, tal que $X \subset rB$.*

Definição 1.7. *O subconjunto $X \subset \mathbb{R}^n$ é aberto, se para qualquer elemento, $u \in X$, existe um $\epsilon > 0$, tal que $u + \epsilon B \subset X$. O complementar de X em $\mathbb{R}^n$ é dito fechado, denotaremos por X^c.*

Exemplo 1.1. *O intervalo aberto $(0, 1) \subset \mathbb{R}$ é um conjunto aberto em $\mathbb{R}$. A bola unitária aberta $B(0, 1) \subset \mathbb{R}^n$ é um conjunto aberto em $\mathbb{R}^n$.*

Propriedade 1.1. *Seja $\{A_\alpha\}_{\alpha \in I}$ uma família de abertos do $\mathbb{R}^n$. A reunião arbitrária de conjuntos $\bigcup A_\alpha$ é aberta.*

Demonstração. Seja, $u \in \bigcup A_\alpha$. Então $u \in A_\alpha$, para algum índice $\alpha \in I$. Logo, por definição, existe um $\epsilon > 0$, tal que $u + \epsilon B \subset A_\alpha$. Portanto, $u + \epsilon B \subset \bigcup A_\alpha$. $\square$

Propriedade 1.2. *Seja $\{F_\alpha\}_{\alpha \in I}$ uma família de conjuntos fechados do $\mathbb{R}^n$. A interseção arbitrária $\bigcap F_\alpha$ é fechada.*

Demonstração. Defina $A = \mathbb{R}^n \setminus \bigcap F_\alpha$. Então, pela propriedade dos conjuntos, $A = \bigcup \mathbb{R}^n \setminus F_\alpha$. Defina $A_\alpha = \mathbb{R}^n \setminus F_\alpha$ para cada $\alpha \in I$. Logo, A_α é aberto e

$A = \bigcup A_\alpha$. Daí, pela Propriedade 1.1 concluímos que A é aberto e $\bigcap F_\alpha = \mathbb{R}^n \setminus A$ é fechado. $\square$

Definição 1.8. *Um elemento $u \in \mathbb{R}^n$ é dito ponto de aderência de um conjunto $X \subseteq \mathbb{R}^n$ quando for o limite de uma sequência de elementos de X.*

Exemplo 1.2. *A origem $0 \in \mathbb{R}$ é o ponto de aderência da sequência de números reais $(x_k)_{k\in\mathbb{N}}$, definida por:*

$$x_n = 1/n, \textit{ para todo } n \textit{ natural.}$$

Definição 1.9. *Um subconjunto $X \subseteq \mathbb{R}^n$ é compacto quando ele for fechado e limitado.*

Exemplo 1.3. *A bola fechada unitária do $\mathbb{R}^n$ definida por:*

$$B[0,1] = \{u \in \mathbb{R}^n \colon \; \|u\| \leq 1\}$$

é compacta.

A Definição 1.9 , na verdade, é uma propriedade do conjunto $\mathbb{R}^n$, tendo em vista que a definição topológica de compacidade é mais abrangente, como veremos adiante.

Definição 1.10. *O interior de um conjunto X é o conjunto de todos os elementos $u \in X$, tais que $u + \epsilon B \subset X$ para algum $\epsilon > 0$. Denotamos o interior do conjunto X como $int(X)$.*

Exemplo 1.4. *O interior da bola unitária fechada do $\mathbb{R}^n$ definida por:*

$$B[0,1] = \{u \in \mathbb{R}^n \colon \; \|u\| \leq 1\}$$

é a bola aberta:

$$B = B(0,1) = \{u \in \mathbb{R}^n \colon \; \|u\| < 1\}.$$

Definição 1.11. *O fecho de um subconjunto $X \subseteq \mathbb{R}^n$, é a intersecção de todos os conjuntos fechados que contém X. Denotamos o fecho de X por $cl(X)$.*

Observação 1.2. *Seja $X \subseteq \mathbb{R}^n$. Então*

$$X \subseteq cl(X).$$

Observação 1.3. *Seja $X \subseteq \mathbb{R}^n$. Então*

$$int(X) \subseteq X \subseteq cl(X).$$

Exemplo 1.5. *Em* $\mathbb{R}$*, o fecho do intervalo aberto* $(0,1) \in \mathbb{R}$ *é o intervalo fechado* $[0,1] \in \mathbb{R}$.

Proposição 1.1. *Seja* $X \subseteq \mathbb{R}^n$ *e considere* $\{u_n\}_{n\in\mathbb{N}}$ *uma sequência de* X *que converge para* u*. Então* $u \in cl(X)$.

Demonstração. Por definição $cl(X)$ é um conjunto fechado. Suponha que $u \notin cl(X)$. Então u pertence ao complementar de $cl(X)$, que é aberto por definição. Denotemos o complementar de $cl(X)$ por A. Naturalmente, $A \cap cl(X) = \emptyset$. Então como A é aberto, existe $\epsilon > 0$, tal que $u + \epsilon B \subset A$, onde B é a bola aberta unitária do $\mathbb{R}^n$. Absurdo, pois pela definição de convergência, toda vizinhança de u deverá conter um elemento da sequência que pertence a X. □

Definição 1.12. *A fronteira de um conjunto qualquer* $X \subseteq \mathbb{R}^n$*, é definida pelo conjunto de elementos* $u \in \mathbb{R}^n$*, tais que para qualquer* $\epsilon > 0$*, tem-se que* $(u + \epsilon B) \cap X \neq \emptyset$ *e* $(u + \epsilon B) \cap X^c \neq \emptyset$*. Este conjunto é denotado por* $bd(X)$.

Exemplo 1.6. *O conjunto* $\mathbb{R}^n$ *é a fronteira do conjunto* $\mathbb{Q}^n$*. Observe que o fecho do* $\mathbb{Q}^n$ *é o* $\mathbb{R}^n$ *e toda bola de centro num elemento em* $\mathbb{Q}^n$ *contém infinitos pontos do* $\mathbb{Q}^n$ *e do* $\mathbb{R}^n$.

Propriedade 1.3. *Pelas definições acimas, temos que para todo conjunto* $X \subseteq \mathbb{R}^n$*, tem-se*

i. $\mathbb{R}^n = int(X) \cup bd(X) \cup (\mathbb{R}^n \setminus X)$

ii. $cl(X) = int(X) \cup bd(X)$

iii. $bd(X) = cl(X) \cap cl(\mathbb{R}^n \setminus X)$

Com estas preliminares estabelecidas podemos avançar para os próximos capítulos, em que vamos explorar os conjuntos e as funções convexas.

Capítulo 2

Conjuntos Convexos

Nesta seção, apresentamos os conceitos e os teoremas fundamentais para o entendimento dos conjuntos convexos e das suas propriedades. A compreensão desses conceitos, das propriedades topológicas e o teorema da separação são essenciais para o avanço na teoria da Análise Convexa. Os conceitos, as definições e os teoremas deste capítulo foram consultados nas referências [6, 11, 16].

2.1 Conjuntos Afins

Definição 2.1. *Um conjunto $S \subseteq \mathbb{R}^n$ é denominado afim, se para quaisquer $u, v \in S$ e $\theta \in \mathbb{R}$, tem-se:*

$$(1-\theta)u + \theta v \in S.$$

A definição significa que a reta que conecta dois elementos distintos do conjunto afim, está contida neste conjunto.

Observação 2.1. *Seja $S \subseteq \mathbb{R}^n$ um subespaço vetorial. Então S é um conjunto afim que contém a origem.*

Proposição 2.1. *Seja $\{S_\alpha\}_{\alpha \in I}$ uma família de subconjuntos afins do $\mathbb{R}^n$. Então $X = \bigcap_{\alpha \in I} S_\alpha$ é afim.*

Demonstração. Dado $\theta \in \mathbb{R}$ e $u, v \in X$, então $u, v \in S_\alpha$, para todo $\alpha \in I$. Logo $(1-\theta)u + \theta v \in S_\alpha \Rightarrow (1-\theta)u + \theta v \in \bigcap_{\alpha \in I} S_\alpha$. Portanto, X é afim. $\square$

Definição 2.2. *A envoltória afim de um subconjunto $X \subseteq \mathbb{R}^n$ é definida pela interseção de todos os conjuntos afim que contém X. Denotamos por $aff(X)$ a envoltória afim do conjunto X.*

Observação 2.2. *Seja $X \subseteq \mathbb{R}^n$. O conjunto $aff(X)$ é um conjunto afim tal que $X \subseteq aff(X)$.*

Proposição 2.2. *Seja $X \subseteq \mathbb{R}^n$. O conjunto aff(X) é formado por todas as combinações afins dos elementos de X, ou seja,*

$$\mathit{aff}(X) = \left\{ \textstyle\sum_{j=1}^{k} \alpha_j u_j \colon k \in \mathbb{N},\ u_j \in X \ e\ \sum \alpha_j = 1 \right\}.$$

Demonstração. Seja Y o conjunto de todas as combinações afim de elementos de X, ou seja,

$$Y = \left\{ \textstyle\sum_{j=1}^{k} \alpha_j u_j \colon k \in \mathbb{N},\ u_j \in X \text{ e } \sum \alpha_j = 1 \right\}.$$

Vamos demonstrar essa proposição em três etapas. Primeiramente, vamos mostrar que o conjunto Y é afim e $Y \supseteq X$. Em seguida, mostraremos que qualquer conjunto afim Z, tal que $Z \supseteq X$ então $Z \supseteq Y$. Por fim, concluiremos que $Y = \mathit{aff}(X)$. Dados dois elementos quaisquer de Y $u = \sum_{j=1}^{k} \alpha_j u_j$, $v = \sum_{j=1}^{s} \lambda_j v_j$ e $\theta \in \mathbb{R}$, tem-se:

$$(1-\theta)u + \theta v = \sum_{j=1}^{k} (1-\theta)\alpha_j u_j + \sum_{j=1}^{s} \theta \lambda_j v_j, \tag{2.1}$$

em que as somas dos coeficientes é

$$\textstyle\sum_{j=1}^{k} (1-\theta)\alpha_j + \sum_{j=1}^{s} \theta \lambda_j = 1 - \theta + \theta = 1.$$

Ou seja, a equação 2.1 é uma combinação afim de elementos de X. Portanto, $(1-\theta)u + \theta v \in Y$, para quaisquer $u, v \in Y$ e $\theta \in \mathbb{R}$, ou seja, Y é afim. Se $u \in X$, então pela definição de Y, $1 \cdot u \in Y \Rightarrow u \in Y$. Daí, $Y \supseteq X$. Como Y é afim e $Y \supseteq X$ então $\mathit{aff}(X) \subseteq Y$. Agora, suponha Z um conjunto afim qualquer que contém X. Seja $u \in Y$. Se u é um elemento de X ou uma combinação afim de quaisquer dois elementos de X, então $u \in Z$, conforme mostra 2.1. Suponha, por hipótese de indução, que a combinação afim de quaisquer k elementos de X pertença a Z. Considere uma combinação afim de $k+1$ elementos de X, ou seja, $u = \sum_{j=1}^{k+1} \alpha_j u_j$, com $\sum_{j=1}^{k+1} \alpha_j = 1$. Afirmamos que existem k elementos do conjunto $\{\alpha_1, ..., \alpha_{k+1}\}$ com soma não nula. Com efeito, caso a soma de quaisquer k elementos do conjunto $\{\alpha_1, ..., \alpha_{k+1}\}$ fosse nula, teríamos que $\sum_{j=1}^{k+1} \alpha_j = 0$, absurdo, pois a soma dos elementos do conjunto é unitária. Suponha sem perda de generalidade que $\sum_{j=1}^{k} \alpha_j \neq 0$. Considere $\theta_j = \alpha_j / \sum_{l=1}^{k} \alpha_l$, para $j = 1, ..., k$. Daí,

$$u = \textstyle\sum_{j=1}^{k+1} \alpha_j u_j = (\sum_{l=1}^{k} \alpha_l)(\sum_{j=1}^{k} \theta_j u_j) + \alpha_{k+1} u_{k+1}.$$

Mas veja que como $\sum_{j=1}^{k} \theta_j = 1$, por hipótese de indução, $\sum_{j=1}^{k} \theta_j u_j \in Z$ e como a soma dos coeficientes α é 1, novamente por 2.1 temos $u \in Z$. Logo, $Y \subseteq Z$. Por fim, considerando em particular $Z = \mathit{aff}(X)$, temos $Y \subseteq \mathit{aff}(X)$. Portanto $Y = \mathit{aff}(X)$
. □

Definição 2.3 (Conjuntos Paralelos)**.** *Os conjuntos* $X, Y \subseteq \mathbb{R}^n$ *são paralelos quando existe um* $a \in \mathbb{R}^n$ *tal que* $X = Y + a = \{y + a \colon y \in Y\}$.

Proposição 2.3. *Dados* $X, Y \subseteq \mathbb{R}^n$ *paralelos. Se* Y *é afim então* X *é afim.*

Demonstração. Seja $a \in \mathbb{R}^n$ tal que $X = Y + a$. Dados $u, v \in X$ e $\theta \in \mathbb{R}$, existem $w, t \in Y$, tal que $u = w + a$ e $v = t + a$. Daí, $(1-\theta)u + \theta v = (1-\theta)(w + a) + \theta(t + a) = (1-\theta)w + \theta t + a$. Veja que, se Y é afim, então $(1-\theta)w + \theta t \in Y$. Logo, $(1-\theta)w + \theta t + a \in X$, concluindo a demonstração. □

Teorema 2.1. *Seja* $S \subseteq \mathbb{R}^n$ *um conjunto afim não-vazio. Então* S *é paralelo a um único subespaço* $L \subseteq \mathbb{R}^n$.

Demonstração. Dado um elemento $v \in S$, defina o conjunto $L = S - v$. Afirmamos que L é um subespaço. Primeiramente, observe que $0 \in L$. Adicionalmente, podemos escrever $L = S + (-v)$. Como S é afim, pela Proposição 2.3, L é afim. Logo, dados quaisquer $x, y \in L$ e $\theta \in \mathbb{R}$, temos $(1-\theta)0 + \theta x \in L$, ou seja, $\theta x \in L$. Além disso, $(1/2)x + (1/2)y \in L$ e $2((1/2)x + (1/2)y) \in L$, ou seja, $x + y \in L$. Logo, L é um subespaço vetorial. Suponha que Z seja um outro subespaço vetorial paralelo a S. Então Z é paralelo a L, ou seja, existe um $a \in \mathbb{R}^n$, tal que $L = Z + a$. Como $0 \in Z$, então $a \in L$. Como L é subespaço, então $-a \in L$. Se $u \in L$, então $u + (-a) \in L \Rightarrow (u - a) + a \in Z \Rightarrow u \in Z$, então $L \subseteq Z$. De maneira análoga, $Z \subseteq L$, logo, $L = Z$. □

Exemplo 2.1. *O conjunto* $A = \{(x, y) \in \mathbb{R}^2 \colon y = 2x + 1\}$ *é afim e paralelo a um único subespaço do* $C \subset \mathbb{R}^2$, $C = \{(x, y) \in \mathbb{R}^2 \colon y = 2x\}$.

Definição 2.4. *Seja* $L \subset \mathbb{R}^n$ *um subespaço vetorial de dimensão* $m \leq n$. *Definimos* $L^\perp = \{u \in \mathbb{R}^n \colon \langle u, v \rangle = 0, \text{ para todo } v \in L\}$ *o subespaço ortogonal de* L.

Observação 2.3. *Seja* $L \subset \mathbb{R}^n$ *um subespaço vetorial de dimensão* m. *A dimensão do subespaço* $L^\perp$ *conforme Definição 2.4 é* $n - m$. *Na verdade, podemos escrever*

$$\mathbb{R}^n = L \oplus L^\perp.$$

Definição 2.5 (Dimensão de um conjunto afim)**.** *A dimensão de um conjunto afim é definida pela dimensão do subespaço paralelo a este conjunto.*

Definição 2.6 (Dimensão de um conjunto)**.** *Seja* $X \subseteq \mathbb{R}^n$ *um conjunto qualquer. A dimensão de* X *é a dimensão da envoltória afim de* X, *ou seja a dimensão de* $\mathit{aff}(X)$ *conforme definido em 2.5.*

Observação 2.4. *O conjunto de $k+1$ elementos $\{v_0, v_1, ..., v_k\}$ é denominado de afim independente se a envoltória afim $aff(\{v_0, v_1, ..., v_k\})$ for k–dimensional. Seja $v \in aff(\{v_0, v_1, ..., v_k\})$. Do Teorema 2.1, o único subespaço paralelo à este conjunto é $L = aff(\{v_0, v_1, ..., v_k\}) - v$. Em particular, tomando $v = v_0$ podemos escrever,*

$$L = aff(\{0, v_1 - v_0, ..., v_k - v_0\}).$$

Portanto, a envoltória é afim independente se, e somente se, L for de dimensão k, ou seja, os elementos $v_1 - v_0, ..., v_k - v_0$ forem linearmente independentes.

Exemplo 2.2. *O subconjunto $A = \{(0,0,1), (0,1,1), (1,1,1)\}$ do $\mathbb{R}^3$ é afim independente, pois a envoltória afim $aff(A)$, que é paralela ao subespaço*

$$C = aff(\{(0,0,0), (0,1,0), (1,0,0)\}),$$

tem dimensão 2. Na verdade C é uma cópia do $\mathbb{R}^2$ no espaço tri-dimensional $\mathbb{R}^3$.

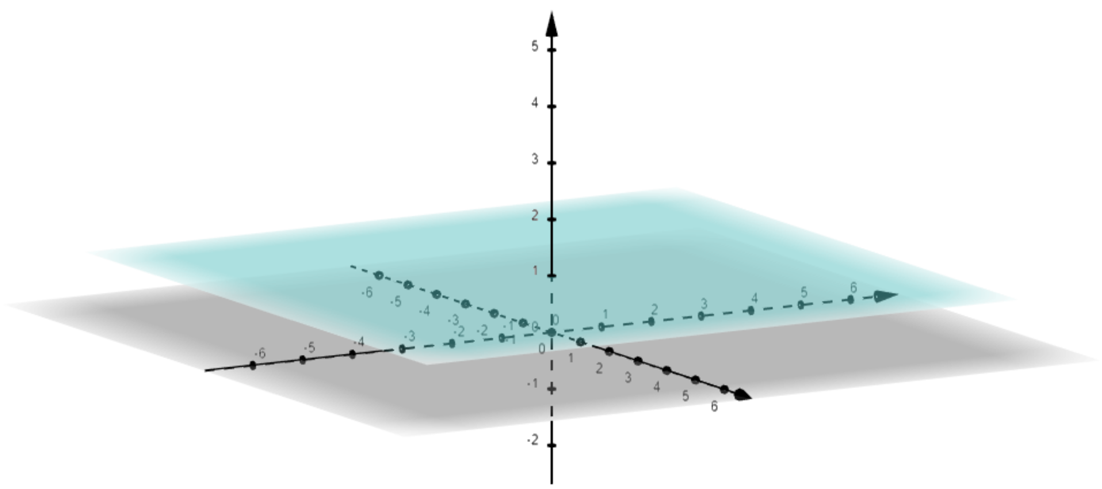

Figura 2.1: Representação do plano $aff(A) \subset \mathbb{R}^3$ que é a envoltória afim do conjunto A. Veja que $aff(A)$ é o plano $z = 1$ paralelo ao plano $z = 0$ do $\mathbb{R}^3$.

Definição 2.7. *Dizemos que $T : \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear se satisfaz as seguintes condições:*

i. $T(u+v) = T(u) + T(v)$ *para quaisquer* $u, v \in \mathbb{R}^n$.

ii. $T(\alpha u) = \alpha T(u)$ *para quaisquer* $u \in \mathbb{R}^n$ *e* $\alpha \in \mathbb{R}$.

Observação 2.5. *Alguns autores utilizam a notação* Tu *em vez de* $T(u)$*, para denotar a transformação linear de* u*.*

Exemplo 2.3. *A projeção* $\pi : \mathbb{R}^2 \to \mathbb{R}$*, definida por* $\pi((x, y)) = x$ *é uma transformação linear. De fato, dados* $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ *e* $\alpha \in \mathbb{R}$*, temos que:*

$$\pi((x_1, y_1) + (x_2, y_2)) = \pi((x_1 + x_2, y_1 + y_2)) = x_1 + x_2 = \pi((x_1, y_1)) + \pi((x_2, y_2))$$

e

$$\pi(\alpha(x_1, y_1)) = \pi((\alpha x_1, \alpha y_1)) = \alpha x_1 = \alpha\pi((x_1, y_1)).$$

Teorema 2.2. *Sejam* $\{v_0, ..., v_m\}$ *e* $\{w_0, ..., w_m\}$ *conjuntos afim independentes em* $\mathbb{R}^n$*, com* $m \leq n$*. Existe uma transformação linear* $T : \mathbb{R}^n \to \mathbb{R}^n$ *tal que*

$$T(v_j - v_0) = w_j - w_0,$$

para $j = 1, ..., m$*.*

Demonstração. Os conjuntos $\{v_0, ..., v_m\}$ e $\{w_0, ..., w_m\}$ são afim independentes. Então, da Observação 2.4, os elementos dos conjuntos $V = \{v_1 - v_0, ..., v_m - v_0\}$ e $W = \{w_1 - w_0, ..., w_m - w_0\}$ são linearmente independentes. Como $m \leq n$, podemos completar os conjuntos V e W com $n - m$ vetores do $\mathbb{R}^n$ cada, formando os conjuntos $V' = \{v'_1, ..., v'_n\}$ e $W' = \{w'_1, ..., w'_n\}$ bases do $\mathbb{R}^n$, onde $v'_j = v_j - v_0$ e $w'_j = w_j - w_0$ para $j = 1, ..., m$. Seja $u \in \mathbb{R}^n$, existem únicos $\alpha_1, ..., \alpha_n \in \mathbb{R}$ tal que:

$$u = \textstyle\sum_{j=1}^n \alpha_j v'_j.$$

Defina $T : \mathbb{R}^n \to \mathbb{R}^n$ como sendo

$$T(u) = \textstyle\sum_{j=1}^n \alpha_j w'_j.$$

Afirmamos que T é uma transformação linear. Com efeito, dados $u, v \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, existem únicos $\alpha_1, ..., \alpha_n \in \mathbb{R}$ e $\beta_1, ..., \beta_n \in \mathbb{R}$, tal que

$$u = \textstyle\sum_{j=1}^n \alpha_j v'_j \text{ e } v = \sum_{j=1}^n \beta_j v'_j.$$

Então

$$T(u + v) = T(\textstyle\sum_{j=1}^n \alpha_j v'_j + \sum_{j=1}^n \beta_j v'_j) = T(\sum_{j=1}^n (\alpha_j + \beta_j) v'_j) = \sum_{j=1}^n (\alpha_j + \beta_j) w'_j.$$

Mas,

$$\textstyle\sum_{j=1}^n (\alpha_j + \beta_j) w'_j = \sum_{j=1}^n \alpha_j w'_j + \sum_{j=1}^n \beta_j w'_j = T(u) + T(v).$$

Além disso,

$$T(\lambda u) = T(\lambda \textstyle\sum_{j=1}^n \alpha_j v'_j) = T(\sum_{j=1}^n \lambda\alpha_j v'_j) = \sum_{j=1}^n \lambda\alpha_j w'_j = \lambda \sum_{j=1}^n \alpha_j w'_j = \lambda T(u).$$

Veja que $T(v_j') = T(1.v_j') = 1.w_j' = w_j'$, para todo $j = 1, ..., n$. Em particular para $j = 1, ..., m$, temos $T(v_j - v_0) = w_j - w_0$.

□

Exemplo 2.4. *Os subconjuntos $A, B \subset \mathbb{R}^3$ definidos por:*

$$A = \{(0,0,1), (0,1,1), (1,0,1)\} \text{ e } B = \{(1,0,2), (2,0,2), (1,1,2)\}$$

são afim independentes paralelos ao mesmo subespaço L gerado pelos vetores do conjunto $\{(0,1,0), (1,0,0)\}$. Então a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$, tal que $T((x,y,z)) = (y,x,z)$ satisfaz as condições do problema anterior. Ao plotar a envoltória afim dos conjuntos A e B no $\mathbb{R}^3$, observa-se planos paralelos, isto é, aff(A) pode ser obtida de aff(B) através de uma translação.

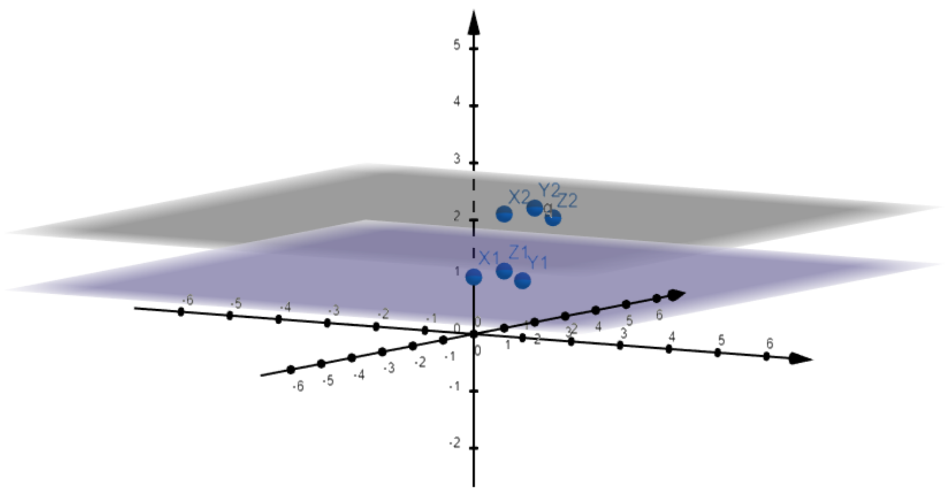

Figura 2.2: Os planos *aff*(A) e *aff*(B) representados pela envoltória afim dos pontos dos conjuntos $A = \{X_1, Y_1, Z_1\}$ e $B = \{X_2, Y_2, Z_2\}$, onde $X_1 = (0,0,1), Y_1 = (0,1,1), Z_1 = (1,0,1), X_2 = (1,0,2), Y_2 = (2,0,2), Z_2 = (1,1,2)$.

Definição 2.8. *O interior relativo de um conjunto $X \subseteq \mathbb{R}^n$ é definido por*

$$ri(X) = \{v \in \mathit{aff}(X) \colon (v + rB) \cap \mathit{aff}(X) \subseteq X, \text{ para algum } r > 0\}.$$

Onde B é a bola unitária aberta conforme Definição 1.5.

Exemplo 2.5. *Considere o subconjunto do $\mathbb{R}^3$*

$$C = \{(x, y, 0) \in \mathbb{R}^3 \colon x^2 + y^2 \leq 1\}.$$

Observe que conjunto aff(C) é o plano $z = 0$. O interior do conjunto C é vazio, entretanto, o conjunto ri(C) = $\{(x, y, 0) \in \mathbb{R}^3 \colon x^2 + y^2 < 1\}$. A figura a seguir ilustra o conceito de interior relativo.

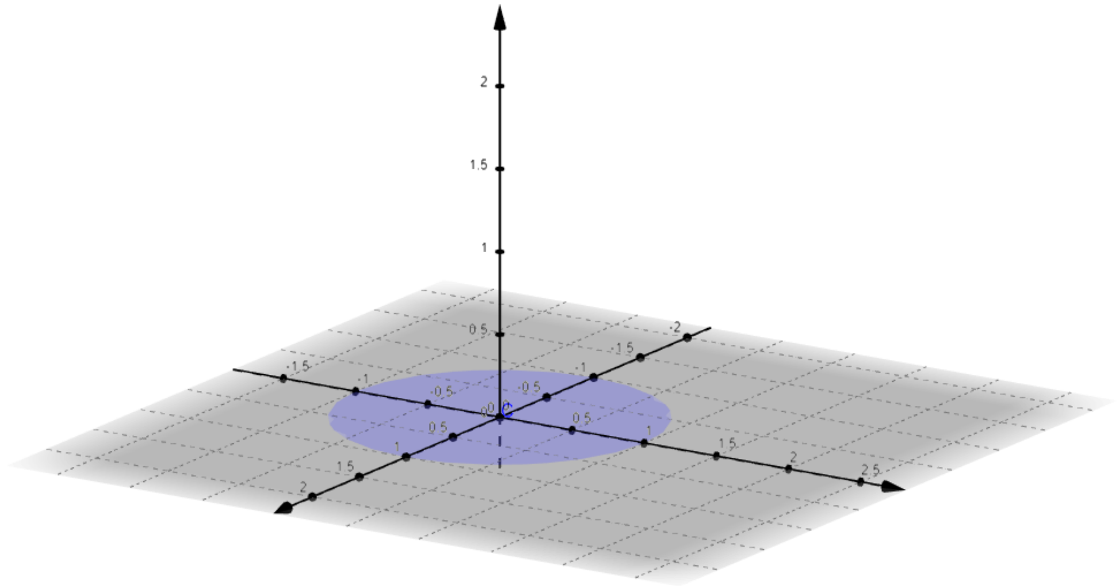

Figura 2.3: Representação do interior relativo de um conjunto $C \subset \mathbb{R}^3$ que é a bola aberta de dimensão 2 contida no plano $z = 0$.

2.2 Hiperplanos

Definição 2.9. *Um hiperplano $H \subseteq \mathbb{R}^n$ é um conjunto afim $(n-1)$-dimensional.*

Observação 2.6. *Pelo Teorema 2.1, H é paralelo à um único subespaço L de dimensão $n - 1$. Escrevemos $H = L + a$ para algum $a \in \mathbb{R}^n$. Existe um elemento $b \in \mathbb{R}^n$, que pertence ao subespaço ortogonal $L^{\perp}$, tal que para todo $v \in L$, $\langle v, b\rangle = 0$. Então*

$$H = \{v + a \colon v \in L\} = \{v + a \colon \langle v, b\rangle = 0, v \in \mathbb{R}^n\} = \{w \colon \langle w - a, b\rangle = 0, w \in \mathbb{R}^n\} = \{w \colon \langle w, b\rangle = c, w \in \mathbb{R}^n\},$$

onde $c = \langle a, b\rangle$. O elemento $b \in \mathbb{R}^n$ é chamado de vetor normal ao hiperplano H. Observe que b não é único. De fato, para todo $\lambda \in \mathbb{R}$ se b é normal ao hiperplano, então λb também é normal.

Exemplo 2.6. *O conjunto* $H = \{(x,y) \in \mathbb{R}^2 \colon y = 2x\}$ *é um hiperplano no* $\mathbb{R}^2$*, representado pela reta de inclinação* 2 *que passa pela origem. Um vetor normal à reta é o* $(-2,1)$.

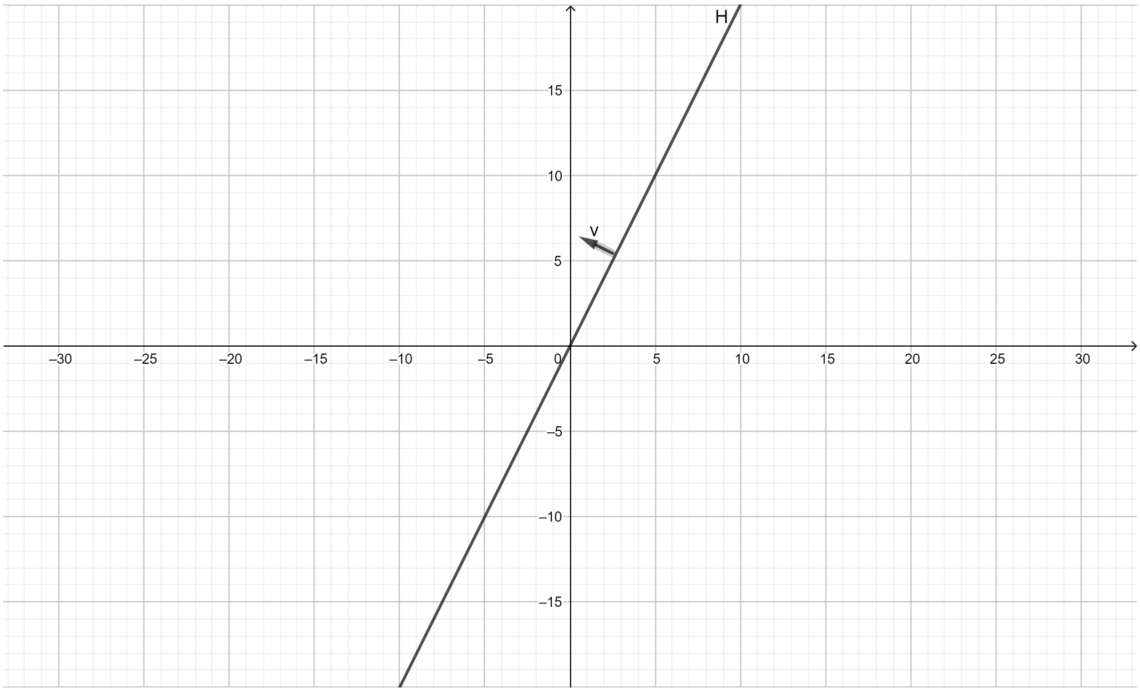

Figura 2.4: Representação de um hiperplano H e seu vetor normal v no $\mathbb{R}^2$. No caso do $\mathbb{R}^2$ um hiperplano é uma reta, ou seja, um conjunto afim unidimensional.

Teorema 2.3. *Dados os elementos* $b \in \mathbb{R}^m$ *e* B *uma matriz real* $m \times n$*, então o conjunto* $S = \{v \in \mathbb{R}^n \colon Bv = b\}$ *é um conjunto afim. Adicionalmente, todo conjunto afim pode ser representado dessa forma.*

Demonstração. Considere $u, v \in S$. Então, dados $\alpha, \beta \in \mathbb{R}$, com $\alpha + \beta = 1$, tem-se que $B(\alpha u + \beta v) = \alpha Bu + \beta Bv = \alpha b + \beta b = b$. Seja $S \subseteq \mathbb{R}^n$ um conjunto afim de dimensão $m \leq n$. Então, pelo Teorema 2.1, S é paralelo a um único subespaço $L \subset \mathbb{R}^n$ de dimensão $n - m$, ou seja, existe $a \in \mathbb{R}^n$, tal que, $S = L + a$. Seja $\{v_1, ..., v_m\}$ uma base para $L^\perp$. Então

$$L = (L^\perp)^\perp = \{u \in \mathbb{R}^n \colon \langle u, v_j \rangle = 0\} = \{u \in \mathbb{R}^n \colon Bu = 0\}.$$

Onde B é a matriz $m \times n$ em que as linhas são os vetores $\{v_1, ..., v_m\}$. Dado $v \in S$, podemos escrever $v = u + a$, com $u \in L$. Então

$$Bv = Bu + Ba = Ba = b.$$

□

Observação 2.7. *O conjunto S definido no Teorema 2.3 é uma interseção finita de m hiperplanos. Com efeito, sejam $l_1, ..., l_m$ os vetores representando as linhas da matriz B, e seja $H_i = \{v \in \mathbb{R}^n : \langle l_i, v\rangle = b_i\}$ para $i = 1, ..., m$, então $S = \bigcap_{1\le i\le m} H_i$.*

Corolário 2.4. *Todo conjunto afim do $\mathbb{R}^n$ é uma interseção finita de hiperplanos.*

Demonstração. De acordo com Teorema 2.3, todo conjunto afim pode ser representado na forma do conjunto $S = \{v \in \mathbb{R}^n : Bv = b\}$, em que S pode ser representado pela intersecção finita de hiperplanos. $\square$

Definição 2.10. *O semiespaço fechado é definido pelo conjunto*

$$S = \{v \in \mathbb{R}^n : \langle v, b\rangle \le c\}, \text{ com } b \neq 0.$$

A fronteira do semiespaço é o hiperplano e o interior do semiespaço é um semiespaço aberto do $\mathbb{R}^n$.

Observação 2.8. *Um hiperplano divide o $\mathbb{R}^n$ em dois semiespaços.*

Exemplo 2.7. *Seja $H = \{(x, y, z) \in \mathbb{R}^3 : x + 2y + 3z = 0\}$ um hiperplano no $\mathbb{R}^3$. H divide o espaço $\mathbb{R}^3$ em duas regiões denominadas semiespaços abertos do $\mathbb{R}^3$ representadas por*

$$HS1 = \{(x, y, z) \in \mathbb{R}^3 : x + 2y + 3z > 0\},$$
$$HS2 = \{(x, y, z) \in \mathbb{R}^3 : x + 2y + 3z < 0\}.$$

O vetor normal ao plano H é $v = (1, 2, 3)$, conforme figura a seguir.

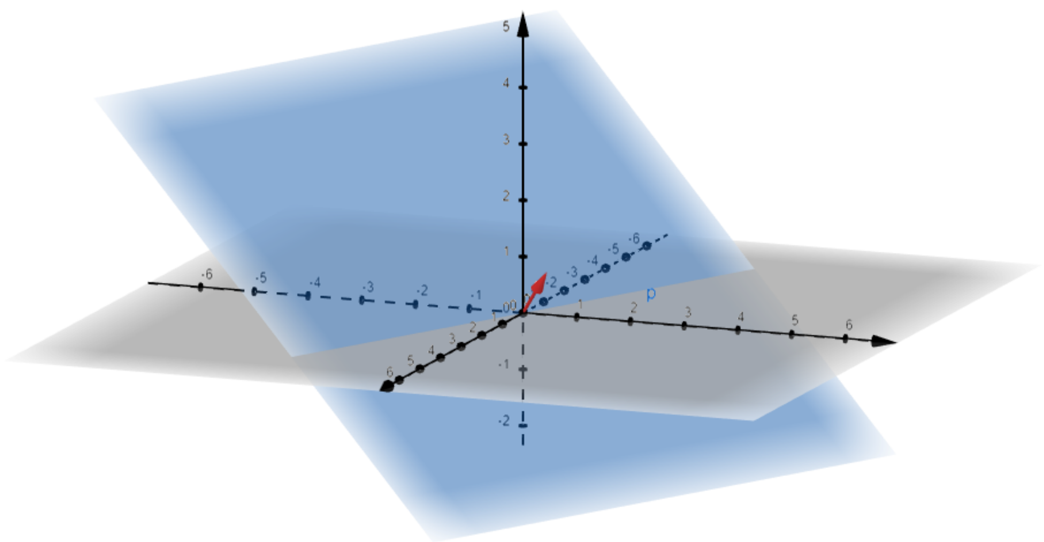

Figura 2.5: Representação de um Hiperplano H e seu vetor normal v no $\mathbb{R}^3$. No caso do $\mathbb{R}^3$ um Hiperplano é um plano, ou seja, um conjunto afim bidimensional.

2.3 Conjuntos Convexos

Definição 2.11. *Um conjunto $S \subseteq \mathbb{R}^n$ é denominado convexo, se para quaisquer $u, v \in S$ e $0 \leq \theta \leq 1$, tem-se $(1-\theta)u + \theta v \in S$. Ou seja, todo segmento de reta que conecta dois pontos distintos do conjunto convexo, está contido neste conjunto.*

Observação 2.9. *Todo conjunto afim é um conjunto convexo.*

Exemplo 2.8. *O conjunto $A = \{(x, y) \in \mathbb{R}^2 \colon x+y = 0\}$ é afim e convexo, entretanto o conjunto $B = (0, 1) \in \mathbb{R}$ é convexo, porém não é afim.*

Exemplo 2.9. *Um n-ágono regular é um conjunto convexo.*

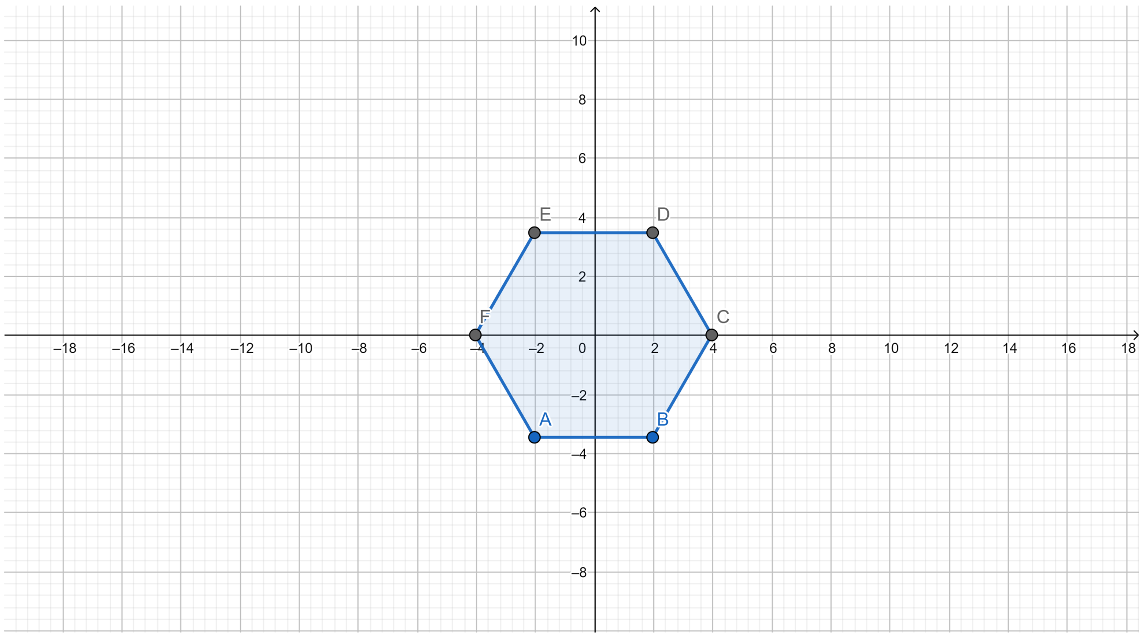

Figura 2.6: Representação de um polígono regular de 6 lados no $\mathbb{R}^2$ que é um exemplo de um conjunto convexo.

Teorema 2.5. *Considere $\{S_\alpha \colon \alpha \in I\}$ uma família de conjuntos convexos do $\mathbb{R}^n$. Então $X = \bigcap S_\alpha$ é convexo.*

Demonstração. Dados dois elementos $v, v \in X$, então $u, v \in S_\alpha$, para todo $\alpha \in I$. Seja $\lambda \in [0, 1]$. Pela convexidade dos conjuntos da família, $(1-\lambda)u + \lambda v \in S_\alpha$, para todo $\alpha \in I$. Portanto, $(1-\lambda)u + \lambda v \in X$. □

Definição 2.12. *Seja $S \subseteq \mathbb{R}^n$. A combinação convexa dos elementos de $v_1, v_2, ..., v_k \in S$ é definida pela soma:*

$$\theta_1 v_1 + \theta_2 v_2 + ... + \theta_k v_k, \ \textit{onde}\ \theta_1, \theta_2, ..., \theta_k \in \mathbb{R}_+ \ \textit{e}\ \theta_1 + \theta_2 + ... + \theta_k = 1.$$

Posto isso, temos o seguinte teorema:

Teorema 2.6. *Um conjunto $S \subseteq \mathbb{R}^n$ é convexo se, e somente se, contém toda combinação convexa de seus elementos.*

Demonstração. A prova é feita por indução, sendo similar à demonstração da Proposição 2.2. Para maiores detalhes, ver [16], Capítulo 2, Teorema 2.2. $\square$

Definição 2.13 (Envoltória Convexa)**.** *A envoltória convexa de um subconjunto $S \subseteq \mathbb{R}^n$ denominada conv(S) é definida como a interseção de todos os conjuntos convexos que contém S.*

Observação 2.10. *De acordo com Teorema 2.5, conv(S) é o menor conjunto convexo que contém S, ou seja, se X é convexo, tal que, $S \subseteq X$, então conv$(S) \subseteq X$.*

Exemplo 2.10. *Considere o conjunto $Q \subset \mathbb{R}^2$, definido por:*

$$Q = \{(x, y) \in \mathbb{R}^2 \colon -3 \leq x \leq -1 \text{ e } -1 \leq y < 1\} \cup \{(-3, 1), (-1, 1)\}.$$

O conjunto Q não é convexo pois $(-3, 1), (-1, 1) \in Q$, porém

$$(-2, 1) = 1/2(-3, 1) + 1/2(-1, 1) \notin Q.$$

A envoltória convexa de Q é definida por:

$$conv(Q) = \{(x, y) \in \mathbb{R}^2 \colon -3 \leq x \leq -1 \text{ e } -1 \leq y \leq 1\}.$$

A figura a seguir ilustra este exemplo, onde para melhor visualização representamos os conjuntos Q e conv$(Q) + (4, 0)$.

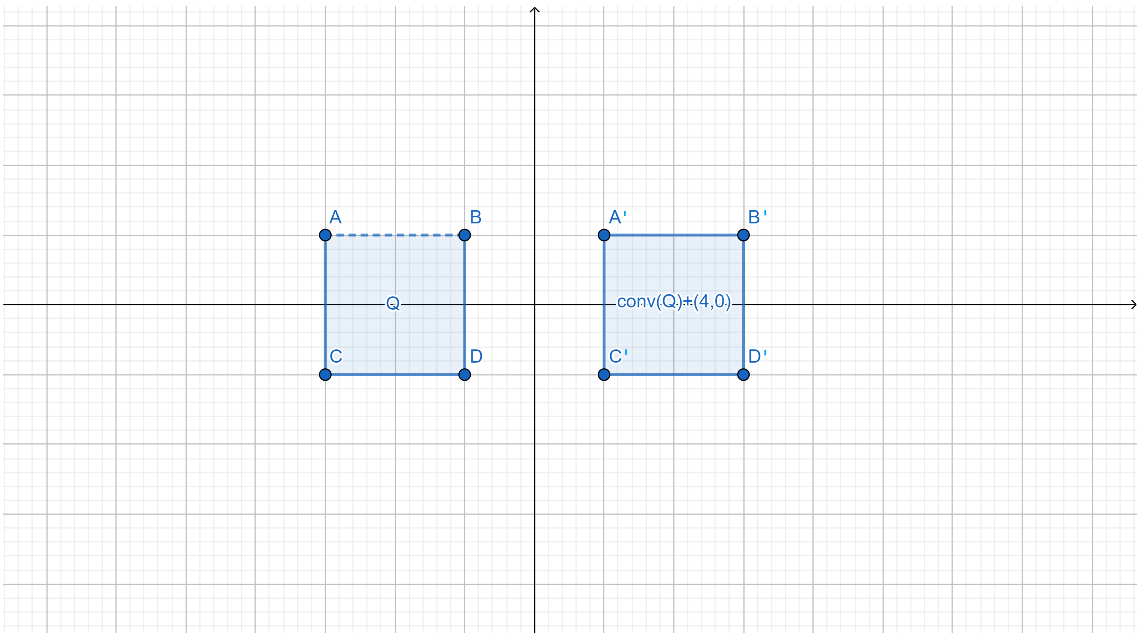

Figura 2.7: Representação do conjunto Q à esquerda e sua envoltória convexa transladada pelo vetor $(4, 0) \in \mathbb{R}^2$ à direita.

Teorema 2.7. *Considere um subconjunto $X \subseteq \mathbb{R}^n$. Então*

$$conv(X) = \{\textstyle\sum_{j=1}^{k} \alpha_j u_j \colon \alpha_j \geq 0,\ \sum_{j=1}^{k} \alpha_j = 1,\ u_j \in X, k \in \mathbb{N}\}.$$

Demonstração. Defina o conjunto

$$Y = \{\textstyle\sum_{j=1}^{k} \alpha_j u_j \colon \alpha_j \geq 0,\ \sum_{j=1}^{k} \alpha_j = 1,\ u_j \in X, k \in \mathbb{N}\}.$$

Veja que Y é convexo, pois dados $u, v \in Y$ e $\theta \in [0, 1]$, podemos escrever

$$u = \textstyle\sum_{j=1}^{k} \alpha_j u_j \text{ e } v = \sum_{j=1}^{l} \lambda_j v_j.$$

Daí,

$$(1-\theta)u + \theta v = \textstyle\sum_{j=1}^{k}(1-\theta)\alpha_j u_j + \sum_{j=1}^{l} \theta\lambda_j v_j \in Y, \text{ pois}$$
$$\textstyle\sum_{j=1}^{k}(1-\theta)\alpha_j + \sum_{j=1}^{l} \theta\lambda_j = (1-\theta)\sum_{j=1}^{k} \alpha_j + \theta \sum_{j=1}^{l} \lambda_j = 1 - \theta + \theta = 1.$$

Do Teorema 2.6, se Z é convexo tal que $Z \supseteq X$, então Z contém todas as combinações convexas dos elementos de X. Logo $Z \supseteq Y$. Portanto, Y é o menor conjunto convexo que contém X, daí $Y = conv(X)$. □

Observação 2.11. *Um conjunto definido pela envoltória convexa de um número finito de pontos é denominado poliedro convexo.*

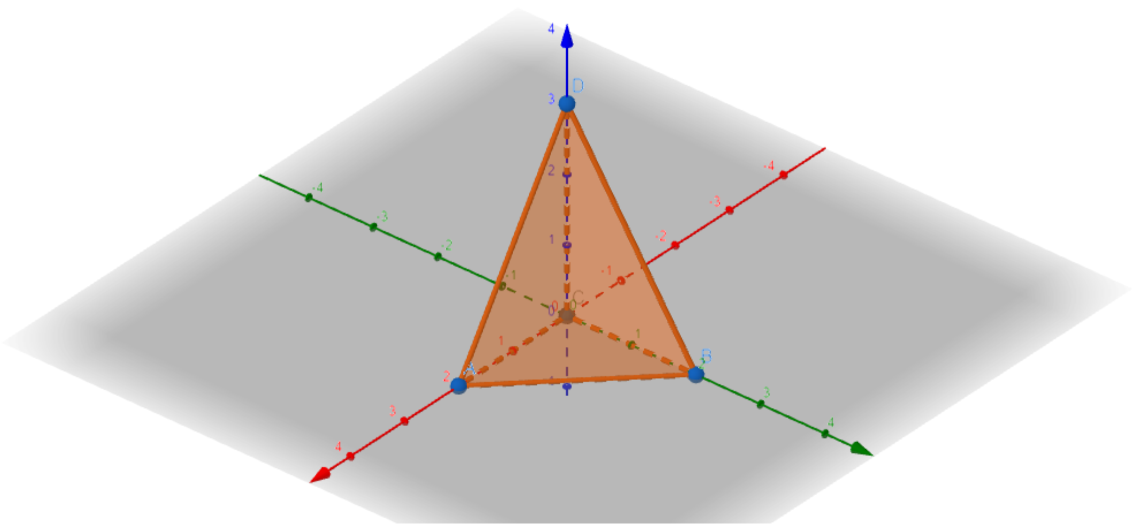

Figura 2.8: A envoltória convexa de 4 pontos do $\mathbb{R}^3$ definindo uma piramide de base triangular que é um conjunto convexo.

Definição 2.14. *O conjunto $L = \{v_0, v_1, ..., v_k\} \subset \mathbb{R}^n$ é denominado afim-independente se $\{v_1 - v_0, v_2 - v_0 ..., v_k - v_0\}$ for um conjunto linearmente independente.*

Definição 2.15. *A dimensão de um conjunto convexo $S \subseteq \mathbb{R}^n$ é a dimensão da sua envoltória afim, conforme Definição 2.5.*

Observação 2.12. *A envoltória convexa do conjunto afim independente conforme Definição 2.14 é denominada de um simplex k-dimensional, onde $v_0, v_1, ..., v_k$ são os vértices deste simplex. Por exemplo, para $k = 0, 1, 2$ ou 3, a envoltória convexa é um ponto, um segmento de reta, um triângulo ou um tetraedro, respectivamente.*

Observação 2.13. *Os semiespaços abertos ou fechados conforme Definição 2.10 são exemplos de conjuntos convexos. Com efeito, considere o seguinte semiespaço*

$$HS = \{v \in \mathbb{R}^n; \langle v, b\rangle \geq c\}, \text{ com } b \neq 0.$$

Dados $v, w \in HS$ e $0 \leq \theta \leq 1$, tem-se

$$\langle((1-\theta)v + \theta w), b\rangle = (1-\theta)\langle v, b\rangle + \theta\langle w, b\rangle \geq (1-\theta)c + \theta c = c.$$

Portanto, $(1-\theta)v + \theta w \in HS$.

Nas figuras a seguir, observe alguns exemplos de conjuntos convexos, não-convexos e envoltória convexa.

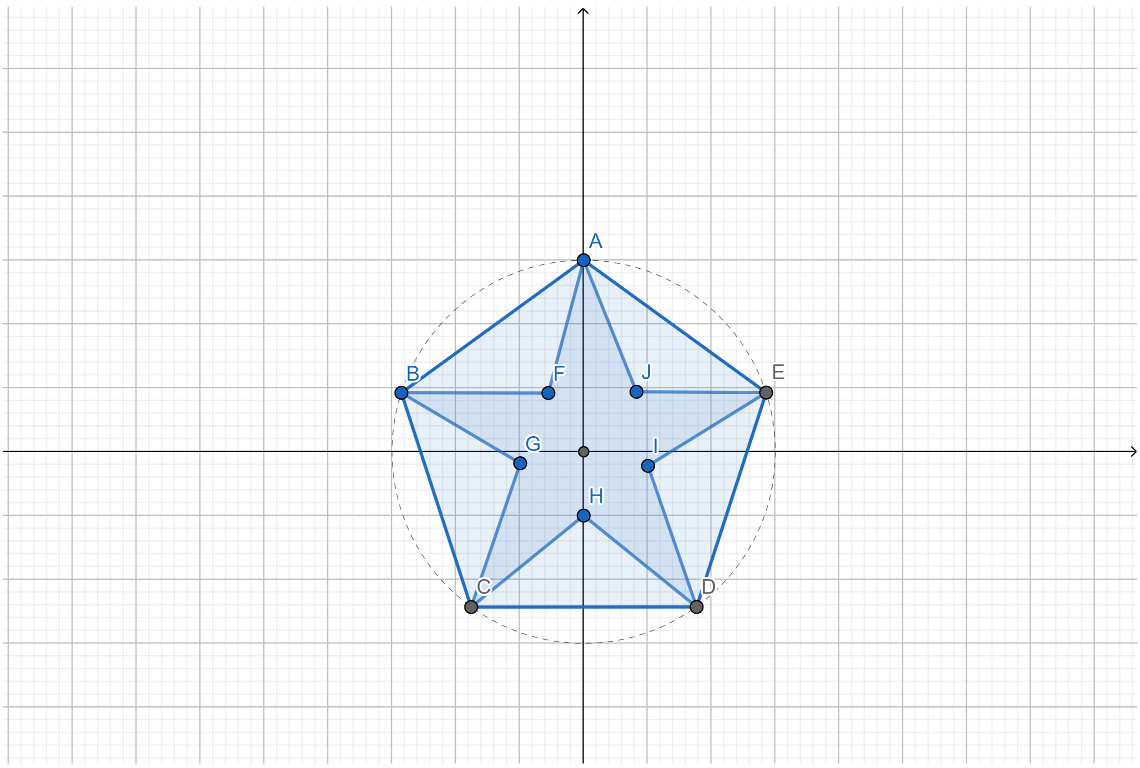

Figura 2.9: A figura representa um estrela no $\mathbb{R}^2$ com vértices A,F,B,G,C,H,D,I,E,J que é um conjunto não convexo. Sua envoltória convexa representada pelo pentágono A,B,C,D,E incluindo seus vértices e arestas.

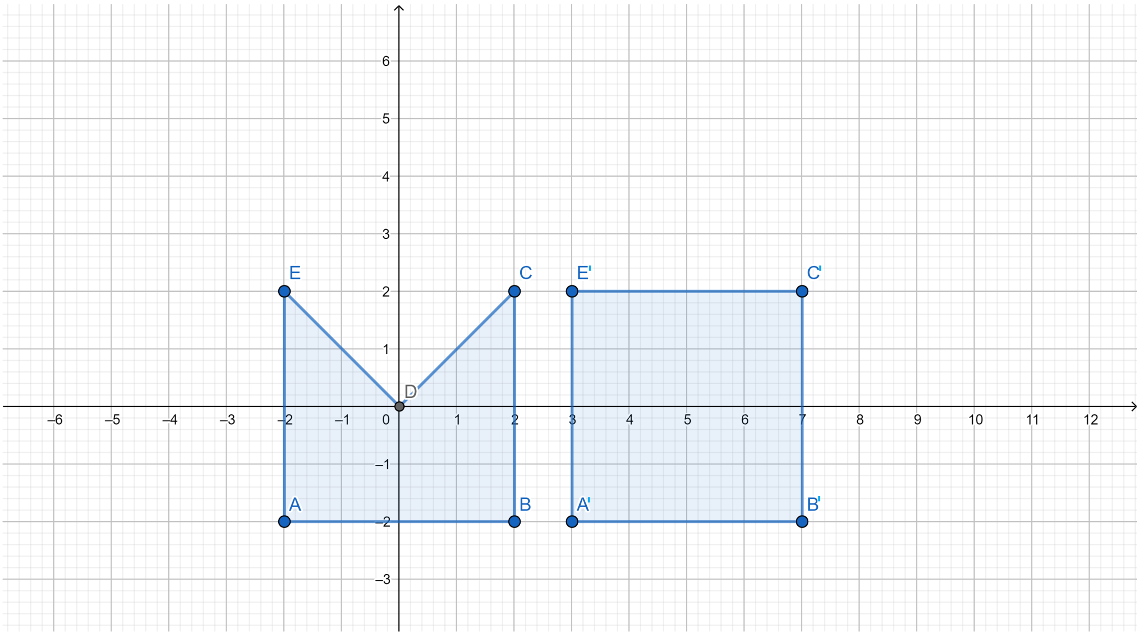

Figura 2.10: A figura representa um polígono não convexo no $\mathbb{R}^2$ com vértices A,B,C,D,E. Sua envoltória convexa representada pelo quadrado A',B',C',E' incluindo seus vértices e arestas transladado pelo vetor $(5, 0)$ é mostrado à direita.

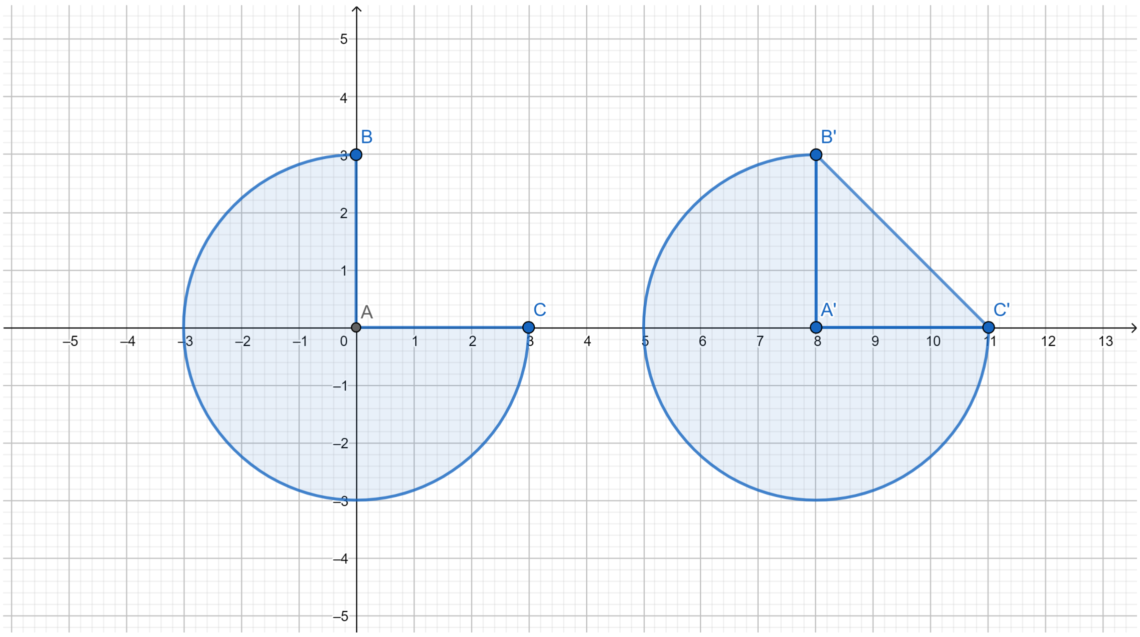

Figura 2.11: Temos a representação de um setor circular no $\mathbb{R}^2$ que é um conjunto não convexo e à direita a sua envoltória convexa transladada pelo vetor $(8, 0)$.

2.4 Operações que Preservam Convexidade

A seguir listamos algumas operações em conjuntos convexos que preservam a convexidade. A verificação da preservação da convexidade pode ser obtida diretamente da definição:

i. Se $S \subseteq \mathbb{R}^n$ é um conjunto convexo, então $-S$ também é convexo.

ii. Se $S \subseteq \mathbb{R}^n$ é um conjunto convexo, então os conjuntos $ri(S)$ e $cl(S)$ são convexos.

iii. Dados $\{S_\alpha\}_{\alpha \in I}$ uma coleção infinita de conjuntos convexos, então sua intersecção é convexa, ou seja $\bigcap_{\alpha \in I} S_\alpha$ é convexa. Por exemplo, um poliedro é a intersecção de semi-espaços, que são convexos, logo, é convexo.

iv. A imagem de um conjunto convexo através de uma transformação linear é convexa. Ou seja, se $S \subseteq \mathbb{R}^n$ é convexo e $T : \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear o conjunto $T(S) \subseteq \mathbb{R}^m$ é convexo.

v. A convexidade é preservada pela soma e multiplicação por escalar. Noutras palavras, dados $S_1, S_2 \subset \mathbb{R}^n$ conjuntos convexos e $\lambda \in \mathbb{R}$, a soma

$$S_1 \pm S_2 = \{u \in \mathbb{R}^n \colon u = v \pm w, \text{ onde } v \in S_1 \text{ e } w \in S_2\}.$$

e o produto

$$\lambda S_1 = \{\lambda u \in \mathbb{R}^n \colon u \in S_1\}.$$

são conjuntos convexos. Portanto, a combinação linear de conjuntos convexos é um conjunto convexo.

vi. O produto cartesiano de conjuntos convexos é convexo. Noutras palavras, dados $S_1, S_2 \subset \mathbb{R}^n$ conjuntos convexos, o produto cartesiano

$$S_1 \times S_2 = \{u \in \mathbb{R}^{2n} \colon u = (v, w), \text{ onde } v \in S_1 \text{ e } w \in S_2\}.$$

é um conjunto convexo.

2.5 Exemplos de Conjuntos Convexos

Nesta seção apresentaremos alguns exemplos de conjuntos convexos para ilustrar o conceito. Algumas convenções a serem utilizadas nos exemplos estão definidas a seguir.

Definição 2.16. *Representaremos por convenção um elemento $v \in \mathbb{R}^n$ como um vetor coluna que é na verdade uma matriz $n \times 1$.*

Definição 2.17. *Uma matriz quadrada real $n \times n$ P é simétrica quando $P = P^T$. A matriz quadrada P é positiva definida quando para todo $v \in \mathbb{R}^n$, tem-se $v^T P v > 0$. Quando $v^T P v \geq 0$ para todo $v \in \mathbb{R}^n$, dizemos que P é positiva semi-definida.*

2.5.1 Elipsóide

Um elipsóide de centro na origem é definido pelo conjunto

$$\Sigma = \{v \in \mathbb{R}^n \colon v^T P^{-1} v \leq 1\},$$

onde P é uma matriz real invertível $n \times n$ simétrica e positiva definida. O elipsóide é um conjunto convexo. Com efeito, seja $A = P^{-1} = (a_{ij})_{i,j \in \mathbb{N}}$ e dados $u = (u_1, ..., u_n), v = (v_1, ..., v_n) \in \Sigma$ e $t \in [0, 1]$ então:

$$\begin{aligned}
((1-t)u + tv)^T A((1-t)u + tv) &= \\
(1-t)^2 u^T A u + t^2 v^T A v + 2t(1-t) u^T A v &\leq \\
(1-t)^2 + t^2 + (2t - 2t^2) \sum_{1 \leq i,j \leq n} u_i v_j a_{ij} &\leq \\
1 - 2t + 2t^2 + (2t - 2t^2) \sum_{1 \leq i,j \leq n} (u_i^2 + v_j^2) a_{ij}/2 &\leq \\
1 - 2t + 2t^2 + (2t - 2t^2)(u^T A u + v^T A v)/2 &\leq \\
1 - 2t + 2t^2 + (2t - 2t^2)(1+1)/2 = 1.&
\end{aligned}$$

Os comprimentos dos semieixos são os valores $\sqrt{\lambda_j}$, onde λ_j para $j = 1, ..., n$ são os autovalores de P. A figura a seguir ilustra um elipsoide no $\mathbb{R}^3$, onde

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

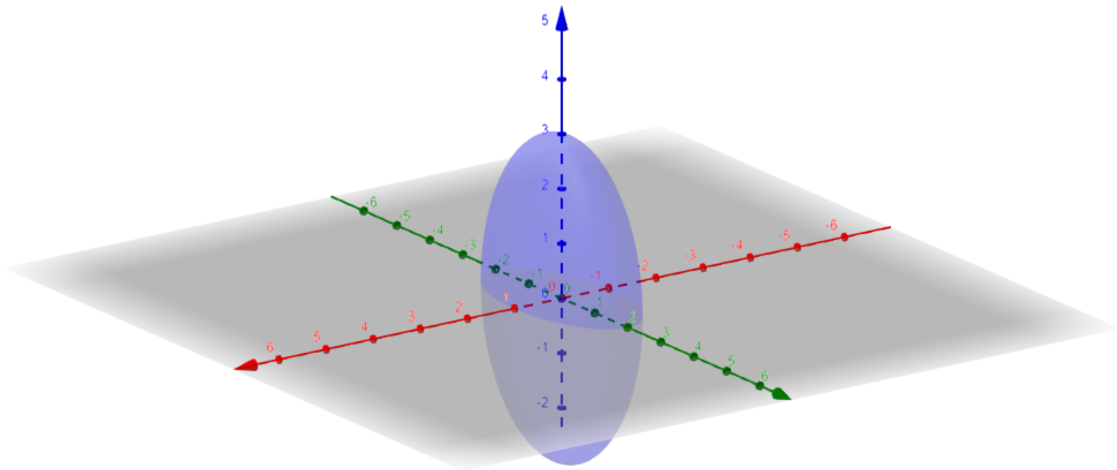

Figura 2.12: A figura acima ilustra um elipsoide no $\mathbb{R}^3$ representado pelo conjunto $E = \{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2/4 + z^2/9 \leq 1\} = \{v \in \mathbb{R}^3 \colon v^T P^{-1} v \leq 1\}$.

2.5.2 Cone Normado

Um cone normado é o conjunto definido pelo conjunto

$$C = \{(v, t) \in \mathbb{R}^{n+1} \colon \ \|v\| \leq t\},$$

onde $v \in \mathbb{R}^n$. Afirmamos que o cone normado é um conjunto convexo. Com efeito, se (v_1, t_1) e $(v_2, t_2) \in C$ e $\theta \in [0, 1]$, então $\|(1-\theta)v_1 + \theta v_2\| \leq (1-\theta)\, \|v_1\| + \theta\, \|v_2\| \leq (1-\theta)t_1 + \theta t_2 \Rightarrow (1-\theta)(v_1, t_1) + \theta(v_2, t_2) \in C$. A figura a seguir ilustra o cone normado no $\mathbb{R}^3$.

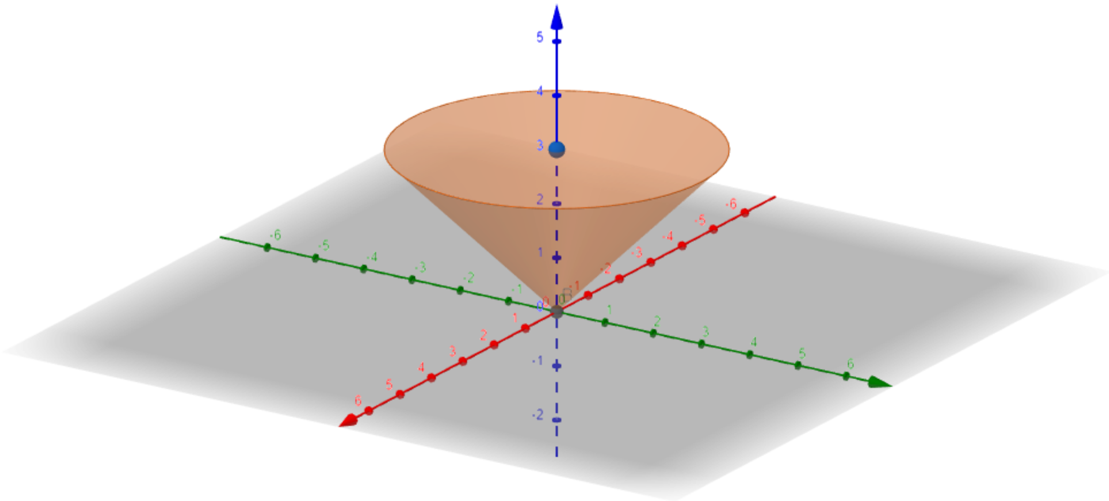

Figura 2.13: A figura acima ilustra um cone normado no $\mathbb{R}^3$ definido pelo conjunto $C = \{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 \leq z, z \leq 3\}$.

2.6 Cones

Definição 2.18. *Um conjunto $K \subseteq \mathbb{R}^n$ é denominado cone, se para quaisquer $v \in K$ e $\theta \in [0, \infty)$, tem-se $\theta v \in K$.*

Definição 2.19. *O conjunto $K \subseteq \mathbb{R}^n$ é um cone convexo se para quaisquer $v, w \in K$ e $\theta_1, \theta_2 \in [0, \infty)$, tem-se $\theta_1 v + \theta_2 w \in K$.*

Definição 2.20 (Combinação Cônica)**.** *A combinação cônica dos elementos*

$$v_1, v_2, ..., v_k \in K$$

é definida pela soma

$$\theta_1 v_1 + \theta_2 v_2 + ... + \theta_k v_k,$$

onde $\theta_1, \theta_2, ..., \theta_k \in [0, \infty)$.

Proposição 2.4. *Seja $\{K_\alpha \colon \alpha \in I\}$ uma família de cones do $\mathbb{R}^n$. O conjunto*

$$C = \bigcap_{\alpha \in I} K_\alpha$$

é um cone.

Demonstração. Seja $u \in C$ e $\theta \geq 0$, temos que $u \in K_\alpha$ para todo $\alpha \in I$. Então, $\theta u \in K_\alpha$ para todo $\alpha \in I \Rightarrow \theta u \in C$. □

Definição 2.21 (Envoltória Cônica)**.** *Dado $X \subseteq \mathbb{R}^n$, a envoltória cônica de X, denotado por $cone(X)$ é a interseção de todos os cones que contém X.*

Proposição 2.5. *Seja $X \subseteq \mathbb{R}^n$. A envoltória cônica de X é o conjunto*

$$cone(X) = \left\{ \sum_{j=1}^{k} \theta_j u_j \colon k \in \mathbb{N}, \theta_j \geq 0, u_j \in X \right\}.$$

Demonstração. A demonstração é feita em 3 passos utilizando o princípio da indução similar à demonstração da Proposição 2.2 e do Teorema 2.3. □

Teorema 2.8. *O conjunto $K \subseteq \mathbb{R}^n$ é um cone convexo se, e somente se, é fechado sobre a adição e sobre a multiplicação por um escalar não negativo.*

Demonstração. Se K é um cone convexo, então pela Definição 2.19, dados $u, v \in K$ e $\alpha, \theta \geq 0$, tem-se $\alpha u + \theta v \in K$. Em particular, se $\alpha = \theta = 1$ tem-se $u + v \in K$. Similarmente, se $\alpha = 1$ e $\theta = 0$ tem-se $\alpha u \in K$. Por outro lado, se K é fechado sobre a adição e a multiplicação por escalar não negativo, então dados $u, v \in K$ e $\alpha, \theta \geq 0$ quaisquer, $\alpha u \in K$ e $\theta v \in K \Rightarrow \alpha u + \theta v \in K$, logo, de acordo com a Definição 2.19, K é um cone convexo. □

Exemplo 2.11. *No $\mathbb{R}^2$ temos o exemplo do cone definido pelo quadrante dos valores de coordenadas positivos, ou seja,*

$$K = \{(x, y) \in \mathbb{R}^2 \ ; \ x \geq 0 \ \text{e} \ y \geq 0\}.$$

K é a envoltória cônica do conjunto

$$C = \{(x, y) \in \mathbb{R}^2 \colon (x - 1)^2 + (y - 1)^2 \leq 1\}.$$

.

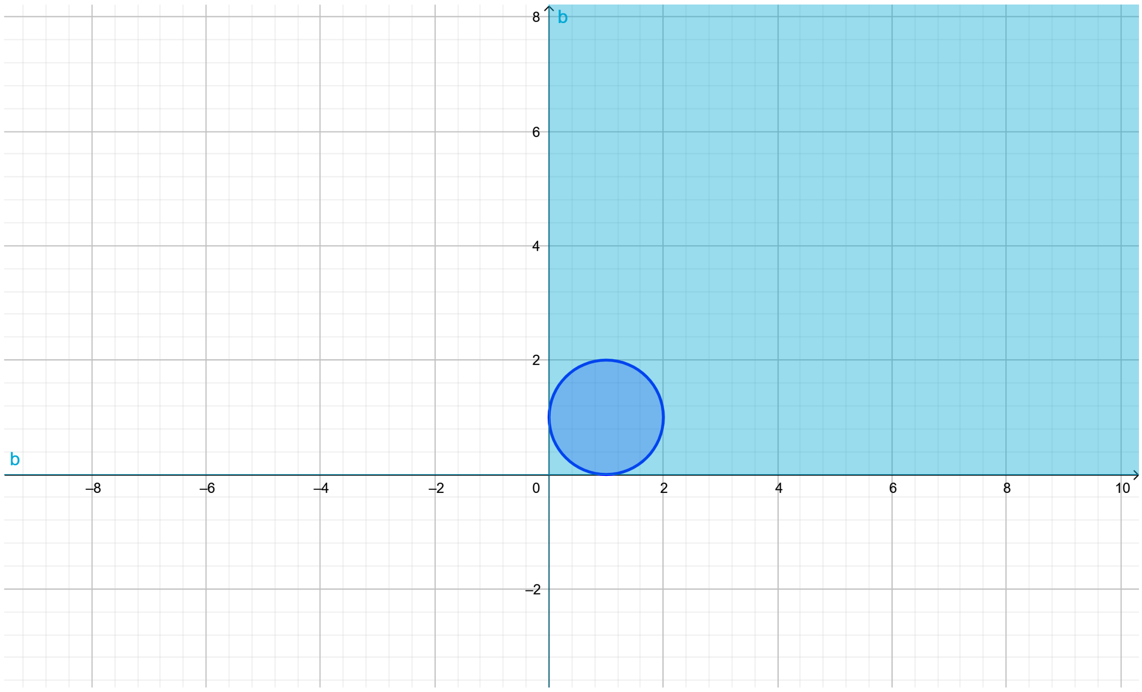

Figura 2.14: A figura acima ilustra o círculo unitário de centro $(1, 1)$ do $\mathbb{R}^2$ e sua envoltória cônica que é o primeiro quadrante conforme Exemplo 2.11.

Definição 2.22 (Cones Próprios)**.** *Um cone $K \subseteq \mathbb{R}^n$ é dito próprio, quando satisfaz as seguintes propriedades:*

i. *K é convexo e fechado.*

ii. *$int(K) \neq \emptyset$.*

iii. *K é direcional, ou seja, se $u \in K$ e $(-u) \in K$, então $u = 0$.*

Quando o cone não é próprio, então ele é denominado impróprio.

Definição 2.23 (Relação de Ordem)**.** *Na teoria de conjuntos dizemos que $\leq$ é uma relação de ordem parcial de um conjunto A se satisfaz as seguintes condições:*

i. *$u \leq u$ para todo $u \in A$.*

ii. *Se para $u, v \in A$ tem-se $u \leq v$ e $v \leq u$, então $u = v$.*

iii. *Se para $u, v, w \in A$ tem-se $u \leq v$ e $v \leq w$, então $u \leq w$.*

A relação de ordem é total, se para quaisquer $u, v \in A$, então tem-se $u \leq v$ ou $v \leq u$, ou seja, para quaisquer dois elementos em A são comparáveis.

Definição 2.24. *Seja $K \subseteq \mathbb{R}^n$ um cone próprio. Definimos uma relação dos elementos do $\mathbb{R}^n$ a operação $\preceq_K$ tal que dados $u, v \in \mathbb{R}^n$:*

$$u \preceq_K v \iff v - u \in K.$$

Proposição 2.6. *A relação da Definição 2.24 é uma relação de ordem parcial.*

Demonstração. Primeiramente observe que dado $u \in \mathbb{R}^n$, temos que $u \preceq_K u$, pois $u - u = 0 \in K$. Em seguida, observe que, se para $u, v \in \mathbb{R}^n$ tem-se $u \preceq_K v$ e $v \preceq_K u$, então, $(u - v) \in K$ e $(v - u) \in K$. Como K é próprio, pela Definição 2.22, tem-se $u - v = 0$. Por fim, se para $u, v, w \in \mathbb{R}^n$ tem-se $u \preceq_K v$ e $v \preceq_K w$, então $v - u \in K$ e $w - v \in K$. Pelas Definições 2.19 e 2.22 tem-se que

$$(v - u) + (w - v) \in K \Rightarrow (w - u) \in K.$$

Então, $u \preceq_K w$, demonstrando que a relação é uma ordem parcial. □

Exemplo 2.12. *Considere o conjunto $K = \mathbb{R}^n_+ = \{u = (u_1, ..., u_n) \in \mathbb{R}^n \colon u_j \geq 0$ para todo $1 \leq j \leq n\}$. Observe que K é um cone próprio. Da relação de ordem $\preceq_K$, tem-se dados $u, v \in \mathbb{R}^n$:*

$$u \preceq_K v \iff u_j \leq v_j, \text{ para todo } 1 \leq j \leq n.$$

2.7 Topologia em Conjuntos Convexos

Com respeito à topologia dos conjuntos convexos, algumas propriedades e teoremas são de fundamental importância para o desenvolvimento de resultados relevantes desta seção, conforme será demonstrado nas proposições e teoremas a seguir.

Definição 2.25. *A dimensão de um conjunto convexo $X \subseteq \mathbb{R}^n$ é a dimensão da sua envoltória afim $\mathit{aff}(X)$ conforme Definição 2.5.*

Definição 2.26. *Dado $z \in \mathbb{R}^n$ a bola aberta de centro z e raio $r > 0$ é o conjunto*

$$B(z, r) = \{x \in \mathbb{R}^n \colon \|x - z\| < r\}.$$

Proposição 2.7. *Dados $z \in \mathbb{R}^n$ e $r > 0$, a bola aberta $B(z, r)$ é um conjunto convexo.*

Demonstração. Sejam $x, y \in B(z, r)$ e $\lambda \in [0, 1]$. Então

$$\|(1 - \lambda)x + \lambda y - z\| = \|(1 - \lambda)x + \lambda y - ((1 - \lambda)z + \lambda z)\| =$$
$$\|(1 - \lambda)(x - z) + \lambda(y - z)\| \leq \|(1 - \lambda)(x - z)\| + \|\lambda(y - z)\| =$$
$$|(1 - \lambda)| \, \|x - z\| + |\lambda| \, \|y - z\| < (1 - \lambda)r + \lambda r = r.$$

Portanto, $(1-\lambda)x + \lambda y \in B(z, r)$. □

Proposição 2.8. *Sejam $X, Y \subseteq \mathbb{R}^n$ conjuntos convexos e $\alpha \in \mathbb{R}$. Então o conjunto*

$$X + \alpha Y = \{x + \alpha y \colon x \in X, y \in Y\}$$

é convexo.

Demonstração. Sejam $u, v \in (X + \alpha Y)$ e $\lambda \in [0, 1]$. Podemos escrever

$$\begin{aligned} u &= x_1 + \alpha y_1 \\ v &= x_2 + \alpha y_2. \end{aligned}$$

Para $x_1, x_2 \in X$ e $y_1, y_2 \in Y$. Daí

$(1-\lambda)u+\lambda v = (1-\lambda)(x_1+\alpha y_1)+\lambda(x_2+\alpha y_2) = ((1-\lambda)x_1+\lambda x_2)+\alpha((1-\lambda)y_1+\lambda y_2).$

Como X e Y são conjuntos convexos, então

$$\begin{aligned} (1-\lambda)x_1 + \lambda x_2 \in X. \\ (1-\lambda)y_1 + \lambda y_2 \in Y. \end{aligned}$$

Logo, $(1-\lambda)u + \lambda v \in X + \alpha Y$. □

Proposição 2.9. *Seja $X \subseteq \mathbb{R}^n$ um conjunto convexo. Então o fecho de X, $cl(X)$ é convexo.*

Demonstração. Dados $x, y \in cl(X)$ e $\lambda \in [0, 1]$. Então existem sequências (x_k) e (y_k) de X, tal que $x_k \to x$ e $y_k \to y$. Defina a sequência (z_k) tal que

$$z_n = (1-\lambda)x_n + \lambda y_n,$$

para todo $n \in \mathbb{N}$. Como X é convexo, então (z_k) é uma sequência de elementos de X. Afirmamos que a sequência (z_k) converge e seu limite é $(1-\lambda)x + \lambda y$. Com efeito, dado $\epsilon > 0$

existe $n_1 \in \mathbb{N}$, tal que para todo $n > n_1$ tem-se $\|x_n - x\| < \epsilon$.
Existe $n_2 \in \mathbb{N}$, tal que para todo $n > n_2$ tem-se $\|y_n - y\| < \epsilon$.

Seja $n_0 = \max\{n_1, n_2\}$. Então, para todo $n > n_0$ tem-se

$$\begin{aligned} \|z_n - ((1-\lambda)x + \lambda y)\| = \|(1-\lambda)(x_n - x) + \lambda(y_n - y)\| \leq \\ |1-\lambda| \, \|x_n - x\| + |\lambda| \, \|y_n - y\| < (1-\lambda)\epsilon + \lambda\epsilon = \epsilon. \end{aligned}$$

Portanto $(1-\lambda)x + \lambda y \in cl(X)$, pois é limite de uma sequência em X. □

Corolário 2.9. *A bola fechada unitária do $\mathbb{R}^n$ é convexa.*

Proposição 2.10. *Seja $X \subseteq \mathbb{R}^n$ um conjunto convexo de dimensão n. Então o interior de X, $int(X)$ é convexo.*

Demonstração. Sejam $x, y \in int(X)$ e $\lambda \in [0,1]$. Então

$$\begin{aligned}\text{existe } r_1 > 0 \text{ tal que } B(x, r_1) \subset int(X).\\ \text{Existe } r_2 > 0 \text{ tal que } B(y, r_2) \subset int(X).\end{aligned}$$

Seja $r = \min\{r_1, r_2\}$. Afirmamos que $x, y \in int(X) - B(0,r)$. Com efeito, o conjunto

$$int(X) - B(0,r) = \{z - w \colon z \in int(X), \|w\| < r\} = \{l \colon z \in int(X), \|z - l\| < r\}.$$

Mas, esse é o conjunto dos l, tal que $B(l,r) \subset int(X)$, em particular, x e y são elementos desse conjunto. Pela Proposição 2.8, o conjunto $int(X) - B(0,r)$ é convexo, portanto, $(1-\lambda)x + \lambda y \in int(X) - B(0,r)$. Por fim, afirmamos que,

$$int(X) - B(0,r) \subset int(X).$$

De fato, se $l \in int(X) - B(0,r)$, então $B(l,r) \subset int(X)$. Caso $l \notin int(X)$ o conjunto $B(l,r)$ teria elementos que não estão em X, absurdo. Logo,

$$(1-\lambda)x + \lambda y \in int(X).$$

□

Teorema 2.10. *Seja $X \subseteq \mathbb{R}^n$ convexo n-dimensional. Dados $x \in int(X)$, $y \in cl(X)$ e $0 \leq \lambda < 1$, tem-se*

$$(1-\lambda)x + \lambda y \in int(X).$$

Demonstração. Dado um $t > 0$ considere o conjunto

$$X(t) = X + B(0,t) = \{z + w \colon z \in X, w \in B(0,t)\}.$$

Afirmamos que $cl(X) \subset X(t)$. Suponha que exista $z \in cl(X)$, tal que $z \notin X(t)$. Existe uma sequência (z_k) em X, tal que $z_k \to z$. Daí, dado $\epsilon = t/2$

$$\text{existe } n_0 \in \mathbb{N} \text{ tal que para todo } n \geq n_0 \text{ tem-se } \|z_n - z\| < \epsilon = t/2.$$

Assim, $z - z_{n_0} \in B(0,t)$ e como $z_{n_0} \in X$, teríamos $z_{n_0} + z - z_{n_0} \in X(t) \Rightarrow z \in X(t)$, absurdo.

Como $x \in int(X)$, dado $0 \leq \lambda < 1$ qualquer, existe um $r > 0$, tal que

$$B(x, r') \subset X, \text{ onde } r' = (1+\lambda)r/(1-\lambda).$$

Afirmamos, que $B((1-\lambda)x + \lambda y, r) \subset X$. Sabemos, que $y \in X(r)$, então $y = z + w$, onde $z \in X$ e $w \in B(0,r)$. Daí,

$$(1-\lambda)x + \lambda y + w = (1-\lambda)x + \lambda(z+w) + w = (1-\lambda)(x + (1+\lambda)w/(1-\lambda)) + \lambda z.$$

Como $\|w\| < r$, então

$$\|x + (1+\lambda)w/(1-\lambda) - x\| = \|(1+\lambda)w/(1-\lambda)\| = (1+\lambda)\|w\|/(1-\lambda) < r'.$$

Portanto, $(x+(1+\lambda)w/(1-\lambda)) \in B(x,r') \subset X$. Da convexidade em X, concluímos que

$$(1-\lambda)x + \lambda y + w \in X, \text{ com } w \in B(0,r).$$

Mas, a equação acima é equivalente a $B((1-\lambda)x+\lambda y, r) \subset X$. Portanto concluímos que

$$(1-\lambda)x + \lambda y \in int(X).$$

□

Teorema 2.11 (Distância e projeção). *Seja $X \subset \mathbb{R}^n$ um conjunto convexo, fechado e não vazio. Suponha que exista $a \in \mathbb{R}^n$, tal que $a \notin X$. Defina a distância de a a X como sendo*

$$d(a,X) = \inf\{\|a-z\| : z \in X\}.$$

Então existe um único $x \in X$, tal que

$$d(a,X) = \|a-x\|.$$

Demonstração. Seja $d = d(a,X)$. Observe que $d > 0$, pois $a \in \mathbb{R}^n \setminus X$ que é aberto. Pela propriedade de ínfimo, podemos construir uma sequência (x_k) em X, tal que $\|a-x_n\|^2 < d^2 + 1/n$. Afirmamos que essa sequência é de Cauchy. Com efeito, dados $m, n \in \mathbb{N}$, temos da regra do paralelogramo nos elementos $a - x_n$ e $a - x_m$:

$$\|x_n - x_m\|^2/4 = \|a-x_n\|^2/2 + \|a-x_m\|^2/2 - \|2a-(x_n+x_m)\|^2/4 =$$
$$\|a-x_n\|^2/2 + \|a-x_m\|^2/2 - \|a-(x_n+x_m)/2\|^2.$$

Como X é convexo, então $(x_n+x_m)/2 \in X$, logo

$$\|a-(x_n+x_m)/2\|^2 > d^2.$$

Portanto,

$$\|x_n - x_m\|^2/4 < d^2/2 + 1/(2n) + d^2/2 + 1/(2m) - d^2 = 1/(2n) + 1/(2m).$$

Ou seja, para qualquer $\epsilon > 0$, tome $N \in \mathbb{N}$ tal que $N > 4/\epsilon^2$. Dados $n, m \geq N$, tem-se

$$\|x_n - x_m\|^2/4 < 1/(2n) + 1/(2m) < \epsilon^2/8 + \epsilon^2/8 = \epsilon^2/4 \Rightarrow \|x_n - x_m\| < \epsilon.$$

Logo, (x_k) é uma sequência de Cauchy de um subconjunto do $\mathbb{R}^n$ que é completo, portanto converge para x. Como X é fechado, então $x \in X$. Suponha que exista um outro elemento $y \in X$ que também minimiza a distância. Novamente, aplicando a regra do paralelogramo em $a - x$ e $a - y$, temos

$$\|x-y\|^2/4 = \|a-x\|^2/2 + \|a-y\|^2/2 - \|a-(x+y)/2\|^2 =$$
$$d^2/2 + d^2/2 - \|a-(x+y)/2\|^2.$$

Como X é convexo, então, $(x+y)/2 \in X$, portanto

$$\|a-(x+y)/2\|^2 > d^2.$$

Daí,

$$\|x-y\|^2/4 = d^2/2 + d^2/2 - \|a-(x+y)/2\|^2 < d^2 - d^2 = 0.$$

Absurdo! Logo x é único. □

Observação 2.14. *Caso no Teorema 2.11 X não fosse um conjunto fechado, ainda assim teríamos a existência de um único ponto em $cl(X)$ que minimiza a distância de um ponto $a \in \mathbb{R}^n \setminus X$ ao conjunto convexo X.*

Lema 2.1. *Seja $X \subseteq \mathbb{R}^n$ um conjunto convexo, fechado, não vazio e seja $a \in \mathbb{R}^n \setminus X$. Do Teorema 2.11 existe um único $x \in X$ tal que*

$$d(a, X) = \|a-x\|.$$

Então, para todo $y \in X$, tem-se

$$\langle a-x, x-y \rangle \geq 0.$$

Demonstração. Suponha que $y = x$. Daí, tem-se

$$\langle a-x, x-y \rangle = \langle a-x, x-x \rangle = 0.$$

Considere $y \neq x$. Suponha por absurdo que

$$\langle a-x, x-y \rangle = -c < 0.$$

Tome $\lambda \in (0, 1]$, tal que

$$\lambda < 2c/\|x-y\|^2.$$

Como X é convexo o elemento

$$x + \lambda(y-x) \in X.$$

Portanto da definição de $d(a, X)$

$$\langle a-(x+\lambda(y-x)), a-(x+\lambda(y-x)) \rangle \geq d(a, X)^2 = \langle a-x, a-x \rangle.$$

Desenvolvendo a inequação, temos

$$2\lambda\langle a-x, x-y \rangle + \lambda^2\langle y-x, y-x \rangle \geq 0 \Rightarrow -2c + \lambda\|x-y\|^2 \geq 0 \Rightarrow \lambda \geq 2c/\|x-y\|^2,$$

absurdo!

□

2.8 Separação por Hiperplano

Definição 2.27. *Sejam X_1 e X_2 subconjuntos do $\mathbb{R}^n$. Dizemos que o hiperplano H separa os conjuntos X_1 e X_2 quando X_1 está contido no fecho de um dos semiespaços definido por H e X_2 está contido no fecho do outro semiespaço definido por H. O hiperplano H separa os conjuntos propriamente, quando os dois conjuntos não estão ambos contidos em H. O hiperplano H separa os conjuntos estritamente, quando existe um $\epsilon > 0$, tal que $X_1 + \epsilon B$ está contido em um dos semiespaços abertos definido por H e $X_2 + \epsilon B$ está contido no outro semiespaço aberto.*

Exemplo 2.13. *Sejam $H = \{(x, y) \in \mathbb{R}^2 \colon y = 2x\}$, $C1 = \{(x, y) \in \mathbb{R}^2 \colon (x + 3)^2 + (y - 3)^2 \leq 9\}$ e $C2 = \{(x, y) \in \mathbb{R}^2 \colon (x - 3)^2 + (y + 3)^2 \leq 9\}$ conjuntos $\mathbb{R}^2$. Então dizemos que H separa os conjuntos $C1$ e $C2$ estritamente, pois temos*

$$y \geq 0 \geq x \text{ em } C1 \text{ e}$$
$$y \leq 0 \leq x \text{ em } C2,$$

não ocorrendo $x = y = 0$ em nenhum dos casos.

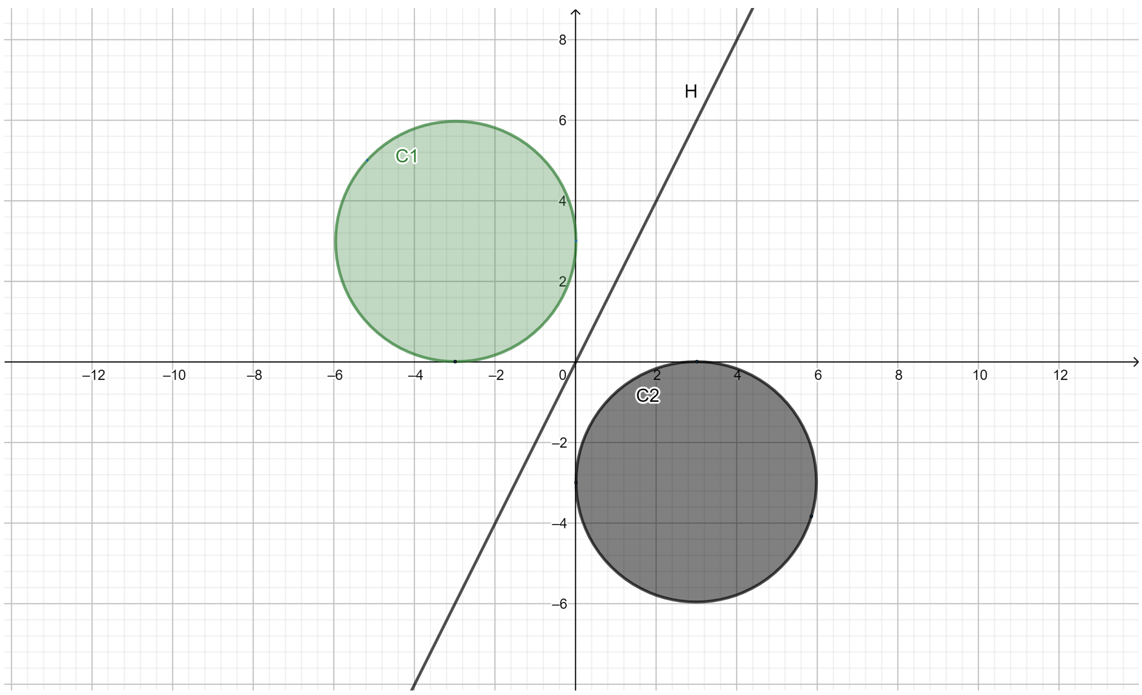

Figura 2.15: A figura acima ilustra o Exemplo 2.13 onde observa-se que a reta H separa estritamente os conjuntos $C1$ e $C2$.

Teorema 2.12. *Sejam X_1 e X_2 conjuntos não vazios do $\mathbb{R}^n$. Então existe um hiperplano separando X_1 e X_2 propriamente se, e somente se, existe um $b \in \mathbb{R}^n$, tal que:*

i. $\inf\{\langle x,b\rangle : x \in X_1\} \geq \sup\{\langle x,b\rangle : x \in X_2\}$.

ii. $\sup\{\langle x,b\rangle : x \in X_1\} > \inf\{\langle x,b\rangle : x \in X_2\}$.

Demonstração. Considere o seguinte hiperplano que separa os conjuntos X_1 e X_2 propriamente:

$$H = \{x \in \mathbb{R}^n : \langle x,b\rangle = c\} \text{ onde } b \in \mathbb{R}^n \text{ e } c \in \mathbb{R}.$$

Então de acordo com a Definição 2.27, X_1 está contido em um dos semiespaços definidos por H, digamos

$$X_1 \subset \{x \in \mathbb{R}^n : \langle x,b\rangle \geq c\}.$$

Note que sempre podemos considerar o semiespaço acima contendo X_1, pois caso fosse o outro semiespaço, bastaria utilizar os elementos $b' = -b$ e $c' = -c$ em substituição de b e c. Daí, X_2 está contido no outro semiespaço, ou seja

$$X_2 \subset \{x \in \mathbb{R}^n : \langle x,b\rangle \leq c\}.$$

Das definições acima, temos que

$$\inf\{\langle x,b\rangle : x \in X_1\} \geq c \geq \sup\{\langle x,b\rangle : x \in X_2\}.$$

Suponha sem perda de generalidade que X_1 não esteja contido em H, então existe um $x_0 \in X_1$, tal que

$$\langle x_0,b\rangle > c.$$

Portanto,

$$\sup\{\langle x,b\rangle : x \in X_1\} \geq \langle x_0,b\rangle > c \geq \sup\{\langle x,b\rangle : x \in X_2\} \geq \inf\{\langle x,b\rangle : x \in X_2\}.$$

Agora, suponha que vale as condições *i.* e *ii.*. Da condição *i.* tome um $c \in \mathbb{R}$, tal que

$$\inf\{\langle x,b\rangle : x \in X_1\} \geq c \geq \sup\{\langle x,b\rangle : x \in X_2\}.$$

Considere o hiperplano

$$H = \{x \in \mathbb{R}^n : \langle x,b\rangle = c\} \text{ onde } b \in \mathbb{R}^n \text{ e } c \in \mathbb{R}.$$

Então, para $x \in X_1$, temos:

$$\langle x,b\rangle \geq \inf\{\langle x,b\rangle : x \in X_1\} \geq c.$$

E, para $x \in X_2$, temos:

$$\langle x,b\rangle \leq \sup\{\langle x,b\rangle : x \in X_2\} \leq c.$$

Mostrando que os conjuntos X_1 e X_2 estão contidos em semiespaços opostos definidos por H. Da condição *ii.* tome um c', tal que

$$\sup\{\langle x, b\rangle : x \in X_1\} > c' > \inf\{\langle x, b\rangle : x \in X_2\}.$$

Suponha sem perda de generalidade que $c' \geq c$. Então existe um $x_0 \in X_1$, tal que

$$\langle x_0, b\rangle > c' \geq c.$$

Logo, X_1 não está contido em H, pois $x_0 \notin H$. Portanto, H separa os conjuntos propriamente. $\square$

Antes de enunciar o próximo teorema vamos precisar da seguinte definição.

Definição 2.28. *Seja $X \subseteq \mathbb{R}^n$. Dizemos que o conjunto $A \subset X$ é aberto em X se dado $a \in A$ existe um $r > 0$, tal que $B(a; r) \cap X \subset A$. O conjunto A é relativamente aberto se ele é aberto em aff(A).*

Teorema 2.13 (Existência do Hiperplano)**.** *Sejam $X \subset \mathbb{R}^n$ um conjunto não vazio, convexo, relativamente aberto e $M \subseteq \mathbb{R}^n$ um conjunto afim, tal que $M \cap X = \emptyset$. Então existe um Hiperplano H contendo M, onde um dos semiespaços abertos associados com H contém X.*

Demonstração. Se M for um conjunto $n-1$-dimensional, M será um hiperplano no $\mathbb{R}^n$, logo existem $b \in \mathbb{R}^n$ e $c \in \mathbb{R}$, tal que

$$M = \{x \in \mathbb{R}^n : \langle x, b\rangle = c\}.$$

Como $M \cap X = \emptyset$, então

$$\langle x, b\rangle \neq c \text{ para todo } x \in X.$$

Suponha que existam $y, z \in X$, tal que

$$\langle y, b\rangle > c.$$
$$\langle z, b\rangle < c.$$

Daí, tome $\lambda = (\langle y, b\rangle - c)/(\langle y, b\rangle - \langle z, b\rangle) \in (0, 1)$, então

$$\langle (1-\lambda)y + \lambda z, b\rangle = c.$$

Mas, como $y, z \in X$ e X é convexo, então $(1-\lambda)y + \lambda z \in X$, absurdo. Portanto X está contido em um dos semiespaços abertos definidos por M. Agora, considere o caso em que a dimensão de M seja $m < n-1$. Vamos demonstrar como construir um espaço afim $M1$, tal que a dimensão de $M1$ é $m+1$, $M \subset M1$ e $X \cap M1 = \emptyset$. De acordo com o Teorema 2.1, M é paralelo a um único subespaço de dimensão m. Suponha sem perda de generalidade que $0 \in M$, de tal sorte que M é um subespaço. Veja que $M^\perp$ é um subespaço de dimensão $n - m \geq 2$. Logo, existe um subespaço bidimensional $P \subset M^\perp$. Considere o conjunto

$$X - M = \{x - y \colon x \in X, y \in M\}.$$

Para melhorar a visualização dos passos a seguir, podemos enxergar P como o $\mathbb{R}^2$ em $\mathbb{R}^n$. Observe que $0 \notin X - M$ e $X \subset X - M$. Adicionalmente, de acordo com a Proposição 2.8 $X - M$ é um conjunto convexo. Defina $X' = P \cap (X - M)$. Logo, X' é um conjunto convexo e relativamente aberto, pois $X - M$ é relativamente aberto. Veja que $0 \notin X'$. Queremos achar um conjunto unidimensional $L \subset P$

$$L = \{\alpha x \colon \alpha \in \mathbb{R}\} \text{ para algum } x \in P.$$

com $L \cap X' = \emptyset$. Assim, teríamos $L \cap (X - M) = \emptyset \Rightarrow X \cap (M + L) = \emptyset$. Daí, o subespaço $M1 = M + L$ seria de dimensão $m + 1$, com $M \subset M1$ e $M1 \cap X = \emptyset$. Primeiramente, se X' é vazio tome $x_0 \neq 0$ em P, então o conjunto

$$L = \{\alpha x_0 \colon \alpha \in \mathbb{R}\}$$

satisfaz a condição pretendida. Se X' possui um único elemento, digamos z, escolha um $x_0 \in P$, tal que $x_0 \notin \{\alpha z \colon \alpha \in \mathbb{R}\}$. Portanto, o conjunto

$$L = \{\alpha x_0 \colon \alpha \in \mathbb{R}\}$$

satisfaz a condição pretendida. Se X' é unidimensional, temos duas situações. Caso $\mathit{aff}(X')$ é uma reta, que não passa pela origem, basta escolher L, a reta paralela a esta que passa pela origem. Caso $\mathit{aff}(X')$ é uma reta que passa pela origem, basta escolher L, a reta perpendicular a esta que passa pela origem. Por fim, caso X' tenha dimensão 2 considere a envoltória cônica de X', ou seja

$$\mathit{cone}(X') = \{\lambda x \colon x \in X', \lambda > 0\}.$$

Como $X' \subset \mathbb{R}^2$ é convexo e relativamente aberto, então $\mathit{cone}(X')$ é um setor do plano, em que as semirretas que definem as extremidades do conjunto possuem ângulo não superior a π. Basta tomar L como a envoltória afim de uma das retas da extremidade desse setor. Neste caso, teríamos

$$L \cap X' = \emptyset \Rightarrow L \cap (X - M) = \emptyset \Rightarrow X \cap (M + L) = \emptyset.$$

Portanto, $M1 = M + L$ será um conjunto afim que satisfaz as condições do problema.

A repetição de $n - m - 1$ passos acima nos produz um conjunto afim $n-1$-dimensional, que satisfaz as condições do problema, conforme primeira parte da demonstração.

□

Teorema 2.14 (Hiperplano Tangente)**.** *Seja $X \subseteq \mathbb{R}^n$ um conjunto convexo, fechado, não vazio e seja $x \in bd(X)$. Então existe uma hiperplano H, tal que $x \in H$ e X está contido em um dos semiespaços fechados definidos por H.*

Demonstração. Se $x \in bd(X)$, então x é um ponto de aderência de $\mathbb{R}^n \setminus X$. Então existe uma sequência de elementos (x_k) em $\mathbb{R}^n \setminus X$ que converge para x. Pelo Teorema 2.11, para cada $n \in \mathbb{N}$, existe $y_n \in X$, tal que

$$\|x_n - y_n\| = \inf\{\|x_n - z\| : z \in X\}.$$

Defina a sequência (h_k) em $\mathbb{R}^n$ como

$$h_n = (x_n - y_n)/ \|x_n - y_n\|.$$

Afirmamos que para qualquer $z \in X$, temos

$$\langle h_n, z\rangle \leq \langle h_n, y_n\rangle < \langle h_n, x_n\rangle.$$

Com efeito, a primeira desigualdade é uma aplicação direta do Lema 2.1 e a segunda desigualdade é obtida a partir de $\|x_n - y_n\| > 0$, pois:

$$\langle h_n, x_n - y_n\rangle = \langle x_n - y_n, x_n - y_n\rangle / \|x_n - y_n\| = \|x_n - y_n\| > 0.$$

Veja que $\|h_n\| = 1$ para todo $n \in \mathbb{N}$, portanto a sequência (h_k) pertence à esfera do $\mathbb{R}^n$ que é compacta. Pelo Teorema de Weierstrass (h_k) possui uma subsequência convergente. Portanto, temos:

$$\langle h_n, z\rangle \to \langle h, z\rangle \text{ em } \mathbb{N}_1 \subset \mathbb{N} \text{ e}$$
$$\langle h_n, x_n\rangle \to \langle h, x\rangle \text{ em } \mathbb{N}_1 \subset \mathbb{N}.$$

Chamamos $\langle h, x\rangle = c$, então o conjunto convexo X está contido em um dos semi-espaços fechados definido pelo hiperplano

$$H = \{w \in \mathbb{R}^n \colon \langle h, w\rangle = c\}.$$

□

Corolário 2.15. *Seja $X \subseteq \mathbb{R}^n$ um conjunto convexo, fechado, não vazio e seja $x \in \mathbb{R}^n \setminus X$. Então existe um hiperplano H, tal que $x \in H$ e X está contido em um dos semiespaços abertos definidos por H.*

Demonstração. Do Teorema 2.11 existe um único $y \in X$ tal que

$$d(x, X) = \|x - y\| = d.$$

Considere o conjunto

$$C = X + B[0, d].$$

Da Propriedade 2.8 C é convexo. Observe que $x \in bd(C)$, então do Teorema 2.14, existe um hiperplano H, tal que $x \in H$ e C está em um dos semiespaços fechados definidos por H. Como $X \subset C$, X está contido nesse semiespaço. Afirmamos que X está contido no interior desse semiespaço. Com efeito, existem $b \in \mathbb{R}^n$ não nulo e $c \in \mathbb{R}$ tal que

$$H = \{w \in \mathbb{R}^n \colon \langle w, b\rangle = c\} \text{ e}$$
$$\langle z, b\rangle \leq c,$$

para todo $z \in C$. Suponha que exista $w \in X$, tal que $\langle w, b\rangle = c$. Podemos escrever $z = w + l$, com $l \in B[0, d]$ qualquer. Daí,

$$\langle w + l, b\rangle \leq c \Rightarrow \langle w, b\rangle + \langle l, b\rangle \leq c \Rightarrow \langle l, b\rangle \leq 0, \text{ para todo } l \in B[0, d].$$

Absurdo, pois se $l \in B[0, d]$ então $-l \in B[0, d]$ e teríamos $\langle l, b\rangle = 0$, para todo $l \in B[0, d]$, mas $b \neq 0$. Portanto $\langle z, b\rangle < c$, para todo $z \in X$. □

Observação 2.15. *No caso do Teorema 2.14, considere u contido no outro semiespaço aberto que não contém X. Então o hiperplano paralelo ao hiperplano tangente que passa por u é tal que X está contido em um dos semiespaços abertos definidos por esse hiperplano.*

Com base nos resultados apresentados acima, podemos enunciar o teorema da separação por hiperplano, que é um importante resultado utilizado, principalmente, na resolução de problemas de existência.

Teorema 2.16. *Sejam $X_1, X_2 \subset \mathbb{R}^n$ conjuntos convexos, fechados, não vazios e disjuntos. Existe um hiperplano H que separa os conjuntos X_1 e X_2. No caso em que $X_2 - X_1$ seja fechado, o hiperplano separa estritamente os conjuntos X_1 e X_2.*

Demonstração. Considere $C = X_2 - X_1$. Observe que C é convexo, $0 \notin C$ e $cl(C)$ é convexo. Portanto $0 \in bd(C)$ ou $0 \in \mathbb{R}^n \setminus cl(C)$. Em qualquer caso, o Teorema 2.14 e o Corolário 2.8 garantem a existência de $b \in \mathbb{R}^n$ não nulo e $c \in \mathbb{R}$ tal que

$$H = \{w \colon \langle w, b\rangle = c\} \text{ e}$$
$$0 \in H \Rightarrow c = 0 \text{ e}$$
$$\langle z, b\rangle \leq 0,$$

para todo $z \in cl(C)$. Em particular, para $z \in C$, $\langle z, b\rangle \leq 0$, daí para quaisquer $x \in X_1$ e $y \in X_2$

$$\langle y - x, b\rangle \leq 0 \Rightarrow \langle y, b\rangle \leq \langle x, b\rangle.$$

Portanto,

$$\sup\{\langle y, b\rangle : y \in X_2\} \leq \inf\{\langle x, b\rangle : x \in X_1\}.$$

Tomando um valor d, qualquer, entre o supremo e o ínfimo da inequação anterior, garante a existência do hiperplano

$$H' = \{w : \langle w, b\rangle = d\}.$$

separando os conjuntos X_1 e X_2. No caso em que C seja fechado e $0 \notin C$, o Teorema 2.11 garante a existência de um único elemento em $z \in C$ com norma mínima. Seja $z = x_2 - x_1$. Defina

$$\begin{aligned} b &= (x_2 - x_1)/2 \neq 0 \text{ e} \\ c &= \langle (x_2 + x_1)/2, b\rangle \text{ e} \\ H &= \{w : \langle w, b\rangle = c\}. \end{aligned}$$

Afirmamos que H separa os conjuntos estritamente. Note que $d((x_2 + x_1)/2, X_1)$ é atingida de acordo com Teorema 2.11 em $x_1 \in X_1$, pois $\|x_2 - x_1\|$ é mínima em C. Então, do Lema 2.1, temos que para todo $x \in X_1$

$$\langle (x_2 + x_1)/2 - x_1, x_1 - x\rangle \geq 0.$$

Daí

$$\langle b, x_1 - x\rangle \geq 0 \Rightarrow \langle x, b\rangle \leq \langle x_1, b\rangle = \langle (x_1 - x_2)/2, b\rangle + \langle (x_1 + x_2)/2, b\rangle = c - \|b\|^2 < c.$$

Analogamente, para todo $y \in X_2$, temos $\langle y, b\rangle > c$. Assim, concluímos que H separa estritamente X_1 e X_2. □

Observação 2.16. *A condição do Teorema 2.16 que $X_2 - X_1$ seja fechado ocorre quando pelo menos um dos conjuntos for limitado e, portanto, compacto no $\mathbb{R}^n$.*

Exemplo 2.14. *Considere os conjuntos convexos do $\mathbb{R}^2$:*

$$\begin{aligned} C1 &= \{(x, y) \in \mathbb{R}^2 : x \leq 0\} \text{ e} \\ C2 &= \{(x, y) \in \mathbb{R}^2 : xy \geq 1, x > 0\}. \end{aligned}$$

É fácil verificar que ambos conjuntos são convexos e fechados, porém não existe um hiperplano que separa estritamente os conjuntos. O hiperplano $x = 0$ separa os conjuntos de forma não estrita.

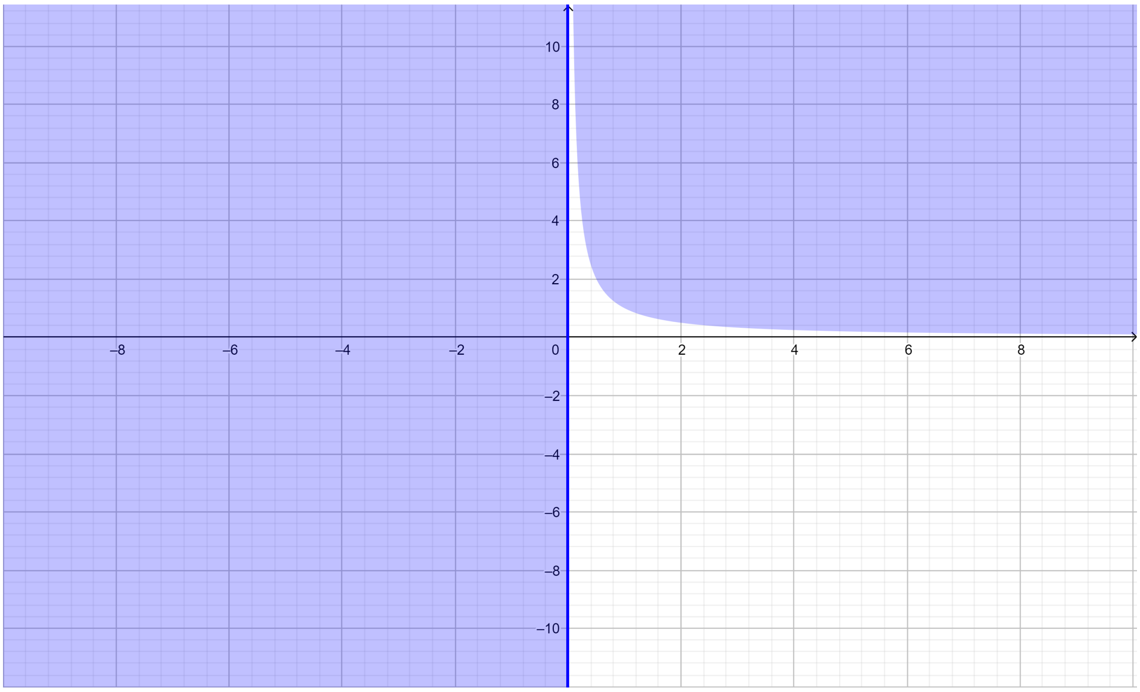

Figura 2.16: A figura acima ilustra o Exemplo 2.14 em que o hiperplano $x = 0$ não separa estritamente os conjuntos convexos fechados $C1$ e $C2$.

No teorema a seguir, apresentamos uma caracterização do conjunto convexo fechado como interseção dos semiespaços que contém o mesmo.

Teorema 2.17. *Seja $X \subset \mathbb{R}^n$ um conjunto convexo e fechado em $\mathbb{R}^n$. Então X é a interseção dos semi-espaços fechados que contém X.*

Demonstração. Para os casos extremos ($X = \emptyset$ ou $X = \mathbb{R}^n$) o teorema segue da própria definição desses conjuntos. Para os outros casos, seja $v \in \mathbb{R}^n \setminus X$. Como X é fechado, $\mathbb{R}^n \setminus X$ é aberto, então existe um $\epsilon > 0$, tal que $B(v, \epsilon) \subset \mathbb{R}^n \setminus X$. Então, de acordo com Teorema 2.16, existe um hiperplano que separa os conjuntos $B(v, \epsilon)$ e X. Então X está contido em um dos semi-espaços fechado desse hiperplano, que não contém v, válido para cada $v \in \mathbb{R}^n \setminus X$. Portanto, a intersecção de todos os semi-espaços, acima construídos, contém todos os pontos de X, porém, nenhum ponto em $\mathbb{R}^n \setminus X$, sendo portanto o conjunto X. $\square$

Capítulo 3

Funções Convexas

Nesta seção, estudaremos sobre as funções convexas e suas propriedades, que será fundamental para o entendimento dos capítulos subsequentes. Os conceitos, as definições e os teoremas desse capítulo foram consultados nas referências [6, 11, 15, 16].

3.1 A Reta Real Estendida

Em matemática, representamos com símbolo ∞ para grandezas que assumam valores tão grande quanto se queira, ou seja, quando os valores são ilimitados. Por exemplo, seja f uma função real, definida por:

$$f : \mathbb{R} \setminus \{1\} \to \mathbb{R}, \quad f(x) = 1/(x-1).$$

Dizemos que o $\lim_{x \to 1^+} f(x) = +\infty$, pois

$$\text{Dado } M > 0, \text{ existe um } \delta > 0 \text{ tal que se } 0 < x - 1 < \delta, \text{ então } f(x) > M.$$

Na linguagem formal, a utilização de $+\infty$ e $-\infty$ envolve um trabalho de tratamento desses casos e muitas vezes torna uma demonstração complicada e extensa. Portanto, existe uma forte motivação para tratar esses elementos como números observando os cuidados no tratamento das operações com os mesmos conforme será mostrado a seguir.

Definição 3.1. *A reta real estendida é definida pela inclusão de $+\infty$ e $-\infty$ como valores, ou seja,*

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\} = [-\infty, +\infty].$$

Definição 3.2. *A função $sgn : \mathbb{R} \to \{-1, 0, 1\}$ é definida como:*

$$sgn(x) = \begin{cases} +1, & se\ x < 0, \\ 0, & se\ x = 0, \\ -1, & se\ x < 1 \end{cases}$$

Propriedade 3.1. *Segue abaixo uma lista de propriedades válidas quando trabalhamos com a reta real estendida.*

i. *Dizemos que* $\infty = +\infty$.

ii. *A ordem* $-\infty < \infty$ *é válida.*

iii. $\infty + \infty = \infty \times \infty = (-\infty) \times (-\infty) = \infty$.

iv. $-\infty - \infty = (-\infty) \times \infty = \infty \times (-\infty) = -\infty$.

v. $\infty + x = x + \infty = \infty$, *para todo* $x \in \mathbb{R}$.

vi. $-\infty + x = x - \infty = -\infty$, *para todo* $x \in \mathbb{R}$.

vii. $x/\infty = x/-\infty = 0$, *para todo* $x \in \mathbb{R}$.

viii. $x \times \infty = \infty \times x = sgn(x) \times \infty$ *para todo* $x \neq 0$ *real.*

ix. $x \times (-\infty) = (-\infty) \times x = sgn(-x) \times \infty$ *para todo* $x \neq 0$ *real.*

Entretanto algumas operações são indefinidas, das quais podemos citar

i. $\infty - \infty$, $-\infty + \infty$.

ii. $0 \times \infty$, $\infty \times 0$.

iii. ∞/∞.

Observação 3.1. *Conforme mencionado, na reta real estendida podemos atribuir os valores* ∞ *ou* $-\infty$. *Portanto, podemos dizer que os conjuntos reais que não são limitados superiormente possuem supremo igual a* ∞. *Analogamente, conjuntos reais não limitados inferiormente possuem ínfimo igual a* $-\infty$.

3.2 Funções Convexas

Definição 3.3 (Domínio de uma função)**.** *Seja* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. *O domínio de* f *é definido pelo conjunto*

$$dom(f) = \{x \in \mathbb{R}^n \colon f(x) \text{ é finito.}\}.$$

Definição 3.4 (Epigráfico)**.** *Seja* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, *com domínio* $X \subset \mathbb{R}^n$. *Definimos o epigráfico de* f *o conjunto*

$$epi(f) = \{(x, \mu) \in X \times \mathbb{R} \colon f(x) \leq \mu\}.$$

Exemplo 3.1. *Seja* $f : \mathbb{R} \to \mathbb{R}$ *uma função real definida por* $f(x) = x^2/10$. *Observe que o domínio da função* $X = \mathbb{R}$ *e que o epigráfico da função é o subconjunto do* $\mathbb{R}^2$

$$epi(f) = \{(x, y) \in \mathbb{R}^2 \colon y \geq x^2/10\}.$$

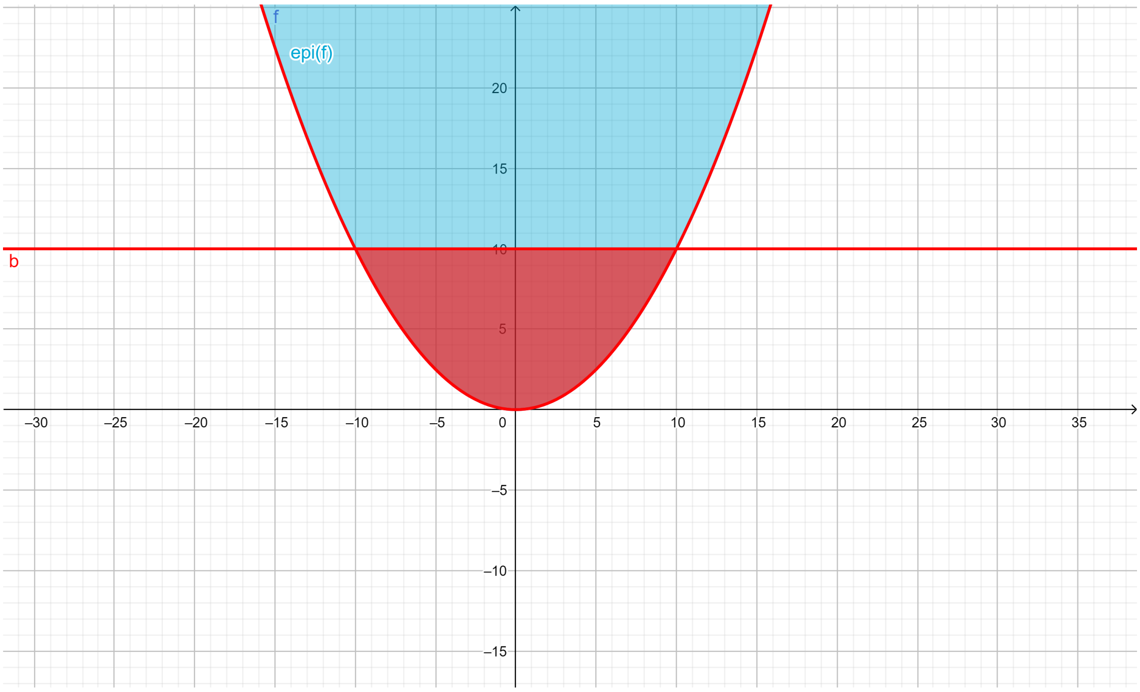

Figura 3.1: A figura acima ilustra o Exemplo 3.1 representando o epigráfico de uma função real de uma variável e sua interseção com o semiespaço $\{(x, y) \colon y \leq 10\}$. Note que tanto o epigráfico quanto a interseção são subconjuntos convexo do $\mathbb{R}^2$.

Exemplo 3.2. *Seja* $f : [-\pi, \pi] \to [-1, 1]$ *uma função real definida por*

$$f(x) = sin(x).$$

Observe que o domínio da função $X = [-\pi, \pi]$ *e que o epigráfico da função é o subconjunto do* $\mathbb{R}^2$

$$epi(f) = \{(x, y) \in X \times \mathbb{R} \colon y \geq sin(x)\}.$$

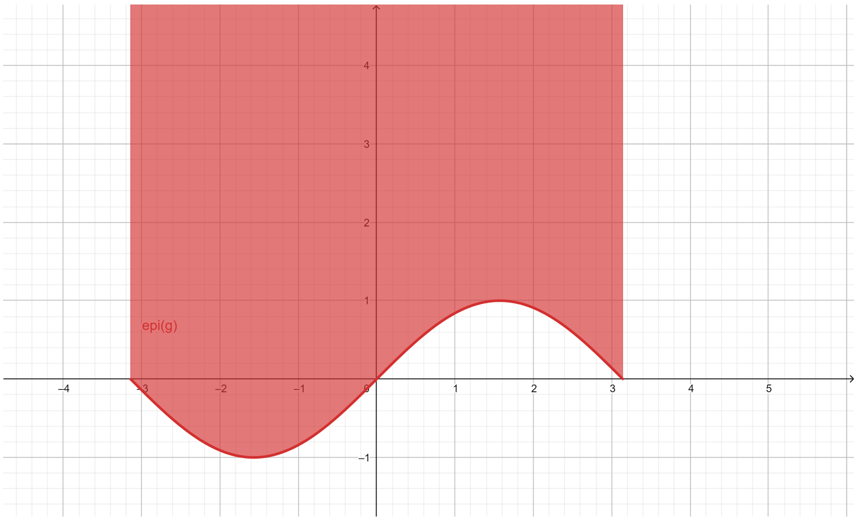

Figura 3.2: A figura acima ilustra o Exemplo 3.2 representando o epigráfico de uma função real de uma variável. Note que o epigráfico, neste caso, não é um subconjunto convexo do $\mathbb{R}^2$.

Definição 3.5 (Funções Convexas)**.** *Seja $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ uma função com domínio $X \subseteq \mathbb{R}^n$. Dizemos que f é convexa quando o conjunto*

$$epi(f) = \{(x, \mu) \in X \times \mathbb{R} \colon f(x) \leq \mu\} \subset \mathbb{R}^{n+1}.$$

for um conjunto convexo.

Observação 3.2. *Nas condições da Definição 3.5 se $epi(f)$ é um conjunto convexo então o domínio da função, $dom(f) = X$, também é convexo. Com efeito, considere a aplicação*

$$T : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n, \quad T(x, \mu) = x.$$

Afirmamos que T é uma transformação linear. Com efeito, dados $(x, \mu), (y, \lambda) \in \mathbb{R}^n \times \mathbb{R}$ e $\alpha \in \mathbb{R}$, temos que

$$T((x, \mu) + \alpha(y, \lambda)) = T((x + \alpha y, \mu + \alpha\lambda)) = x + \alpha y = T(x, \mu) + \alpha T(y, \lambda).$$

Observe que $T(epi(f)) = X$. Logo se $epi(f)$ é convexo, X é convexo de acordo com a Propriedade 2.4, pois é a imagem de um conjunto convexo por uma transformação linear.

Definição 3.6 (Função Côncava)**.** *A função $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ é uma função côncava se $g : \mathbb{R}^n \to \overline{\mathbb{R}}$, tal que $g(x) = -f(x)$ para todo x no domínio de f for convexa.*

Propriedade 3.2. *Considere uma função convexa $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ com domínio $X \subseteq \mathbb{R}^n$. Então dados quaisquer $x, y \in X$ e $\theta \in [0, 1]$ tem-se*

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

A recíproca também é verdadeira.

Demonstração. De acordo com a Definição 3.5, f é convexa quando $epi(f)$ for um conjunto convexo. Logo dados $x, y \in X$, tem-se

$$(x, f(x)), (y, f(y)) \in epi(f).$$

Dado $\theta \in [0, 1]$ tem-se

$$(1-\theta)(x, f(x)) + \theta(y, f(y)) \in epi(f) \Rightarrow ((1-\theta)x + \theta y, (1-\theta)f(x) + \theta f(y)) \in epi(f).$$

Portanto

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

Suponha que a recíproca seja verdadeira. Dados $(x, \mu), (y, \lambda) \in epi(f)$ tem-se

$$f(x) \leq \mu \text{ e } f(y) \leq \lambda.$$

Dado $\theta \in [0, 1]$ considere o seguinte elemento

$$(1-\theta)(x, \mu) + \theta(y, \lambda) = ((1-\theta)x + \theta y, (1-\theta)\mu + \theta\lambda).$$

Da hipótese,

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

Logo

$$(1-\theta)(x, \mu) + \theta(y, \lambda) \in epi(f),$$

concluindo a demonstração. □

Definição 3.7. *A função $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ é estritamente convexa, quando o domínio X for convexo e*

$$f((1-\theta)x + \theta y) < (1-\theta)f(x) + \theta f(y).$$

para todos $x, y \in X$ e $\theta \in [0, 1]$.

Observação 3.3. *Considere o conjunto*

$$H = \{h \colon h : \mathbb{R}^n \to \overline{\mathbb{R}}\}$$

que é o espaço de todas as funções reais. Dados $f, g \in H$ e $\alpha \in \mathbb{R}$ Definimos a soma e o produto por um escalar conforme a seguir

i. $(f + g)(x) = f(x) + g(x)$, *para todo* $x \in dom(f) \cap dom(g)$.

ii. $(\alpha f)(x) = \alpha f(x)$, *para todo* $x \in dom(f)$.

A partir dessas definições verifica-se que H é um espaço vetorial sobre o corpo dos reais. Afirmamos que o conjunto

$$\Omega = \{h : \mathbb{R}^n \to \overline{\mathbb{R}} \colon h \text{ função convexa}\}.$$

é um subespaço vetorial de H. Com efeito, dados $f, g \in \Omega$ e $\alpha \in \mathbb{R}$, suponha que $X = dom(f) \cap dom(g) \neq \emptyset$. Sabemos da Propriedade 2.8 que X é convexo, logo dados $x, y \in X$ e $\theta \in [0,1]$

$$(f+\alpha g)((1-\theta)x+\theta y) = f((1-\theta)x+\theta y) + \alpha g((1-\theta)x+\theta y) \leq$$
$$(1-\theta)(f(x)+\alpha g(x)) + \theta(f(y)+\alpha g(y)) = (1-\theta)(f+\alpha g)(x) + \theta(f+\alpha g)(y).$$

Da Proposição 3.2 a função $f+\alpha g$ é convexa e, portanto, $f+\alpha g \in \Omega$.

3.3 Desigualdades de Jensen, Hölder e Minkowski

Propriedade 3.3 (Desigualdade de Jensen)**.** *Considere uma função definida na reta real estendida $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Então f é convexa se, e somente se,*

$$f(\theta_1 x_1 + ... + \theta_k x_k) \leq \theta_1 f(x_1) + ... + \theta_k f(x_k),$$

onde $\theta_i \geq 0$ para $i = 1, ..., k$ e $\sum_{i=1}^{k} \theta_i = 1$.

Demonstração. Por indução em $k \in \mathbb{N}$. Para o caso $k = 2$ o resultado é obtido a partir da Propriedade 3.2. Suponha que seja válido para algum $k \in \mathbb{N}$. Sejam $\theta_1, ..., \theta_{k+1}$ reais não-negativos com soma unitária. Se $\theta_1 + ... + \theta_k = 0$, então $\theta_{k+1} = 1$, a demonstração segue de imediato. Caso contrário, defina $\theta'_i = \theta_i/(\theta_1 + ... + \theta_k)$, para $i = 1, ..., k$. Daí, para $x_j \in dom(f)$, $j = 1, ..., k+1$ tem-se

$$f(\theta_1 x_1 + \theta_2 x_2 + ... + \theta_{k+1} x_{k+1}) = f((\theta_1 + ... + \theta_k)(\theta'_1 x_1 + ... + \theta'_k x_k) + \theta_{k+1} x_{k+1}).$$

Porém, pelo caso $k = 2$, tem-se

$$f((\theta_1+...+\theta_k)(\theta'_1 x_1+...+\theta'_k x_k)+\theta_{k+1} x_{k+1}) \leq (\theta_1+...+\theta_k) f\left(\sum_{j=1}^{k} \theta'_j x_j\right) + \theta_{k+1} f(x_{k+1}),$$

de modo que utilizando a hipótese de indução com $\theta'_1 + ... + \theta'_k = 1$, tem-se

$$f(\theta'_1 x_1 + ... + \theta'_k x_k) \leq \theta'_1 f(x_1) + ... + \theta'_k f(x_k)$$

mas, como $(\theta_1 + ... + \theta_k)\theta'_i = \theta_i$ para $i = 1, ..., k$, a demonstração se conclui-se. □

Exemplo 3.3 (Desigualdade entre Médias)**.** *A utilização de propriedades das funções convexas podem resultar em demonstrações de desigualdade elementares que são apresentadas no ensino médio ou superior. É o caso da desigualdade generalizada entre a média aritmética e a média geométrica. A função*

$$-log : (0, \infty) \to \mathbb{R},$$

definida nos números reais positivos é convexa, pois conforme caracterização a ser apresentada na seção seguinte temos $(-log(x))'' = 1/x^2 \geq 0$. *Dados* $\lambda_1, \lambda_2, ...\lambda_m$ *reais positivos com soma unitária e* $x_1, x_2, ..., x_m$ *reais positivos quaisquer, tem-se*

$$-log(\lambda_1 x_1 + ... + \lambda_m x_m) \leq -\lambda_1 log(x_1) - ... - \lambda_m log(x_m).$$

A partir da desigualdade acima, concluímos a desigualdade generalizada entre médias, ou seja:

$$\lambda_1 x_1 + ... + \lambda_m x_m \geq x_1^{\lambda_1} x_2^{\lambda_2} ... x_m^{\lambda_m}.$$

No caso particular em que $\lambda_1 = ... = \lambda_m = 1/m$, *tem-se*

$$\frac{x_1 + x_2 + ... + x_m}{m} \geq (x_1 x_2 ... x_m)^{1/m}.$$

Propriedade 3.4 (Desigualdade de Hölder)**.** *Considere* (a_k) *e* (b_k)*, sequências finitas de números reais ou complexos e dados* $p, q > 1$*, reais, tais que* $1/p + 1/q = 1$. *Logo,*

$$|\sum_{j=1}^n a_j b_j| \leq (\sum_{j=1}^n |a_j|^p)^{1/p} (\sum_{j=1}^n |b_j|^q)^{1/q}.$$

Demonstração. Primeiramente, observe que se $\sum_{j=1}^n |a_j|^p = 0$, então $a_j = 0$, para todo j; concluindo a demonstração. De forma análoga, a afirmação vale para sequência (b_k). Portanto, vamos considerar os casos em que

$$t = (\sum_{j=1}^n |a_j|^p)^{1/p} > 0,$$
$$w = (\sum_{j=1}^n |b_j|^q)^{1/q} > 0.$$

Veja que a inequação é homogênea, ou seja, dado $s > 0$, substituir sa_j em vez de a_j, não altera o resultado da inequação. Portanto, considere as substituições

$$a_i \to a_i/t, \text{ para todo } 1 \leq i \leq n.$$
$$b_i \to b_i/w, \text{ para todo } 1 \leq i \leq n.$$

Logo,

$$|\sum_{j=1}^n (a_j/t)(b_j/w)| \leq \sum_{j=1}^n |(a_j/t)(b_j/w)| = \sum_{j=1}^n (|a_j|/t)(|b_j|/w),$$

utilizando a desigualdade triangular. Veja que,

$$\sum_{j=1}^n (|a_j|/t)(|b_j|/w) = \sum_{j=1}^n (|a_j|^p/t^p)^{1/p} (|b_j|^q/w^q)^{1/q}.$$

Mas,

$$\sum_{j=1}^n (|a_j|^p/t^p)^{1/p} (|b_j|^q/w^q)^{1/q} \leq (1/p) \sum_{j=1}^n (|a_j|^p/t^p) + (1/q) \sum_{j=1}^n (|b_j|^q/w^q),$$

utilizando a desigualdade generalizada entre as médias com pesos $1/p$ e $1/q$ do Exemplo 3.3. Agora, observe que

$$\sum_{j=1}^{n}(|a_j|^p/t^p) = 1.$$
$$\sum_{j=1}^{n}(|b_j|^q/w^q) = 1.$$

Portanto,

$$|\sum_{j=1}^{n}(a_j/t)(b_j/w)| \leq 1/p + 1/q = 1.$$

Daí,

$$|\sum_{j=1}^{n}(a_j/t)(b_j/w)| \leq 1 \Rightarrow$$
$$\Rightarrow |\sum_{j=1}^{n} a_j b_j| \leq |t||w| = tw = (\sum_{j=1}^{n}|a_j|^p)^{1/p}(\sum_{j=1}^{n}|b_j|^q)^{1/q}.$$

□

Definição 3.8. *Seja p um número real com $1 \leq p < \infty$. Definimos um espaço das sequências de números reais (p-somáveis) como o conjunto:*

$$\ell^p = \{(a_k)_{k\geq 1} \colon \sum |a_j|^p < \infty\}.$$

Observação 3.4. *Dados $(a_k), (b_k) \in \ell^p$ e $\alpha \in \mathbb{R}$, considere as operações*

i. $(a_k) + (b_k) = (c_k)$, *quando* $c_j = a_j + b_j$, *para todo* $j \in \mathbb{N}$.

ii. $\alpha(a_k) = (c_k)$, *quando* $c_j = \alpha a_j$, *para todo* $j \in \mathbb{N}$.

Com essas definições, constatamos que o espaço ℓ^p é um espaço vetorial. Para $p > 1$ definimos a função $\|.\|_p : \ell^p :\to \mathbb{R}_+$

$$\|(a_k)\|_p = (\sum |a_j|^p)^{1/p}.$$

A função está bem definida, conforme definição do espaço ℓ^p. Um resultado interessante é que essa função representa, de fato, uma norma em ℓ^p. Para demonstrar isso, utilizaremos a Desigualdade de Minkowski.

Propriedade 3.5 (Desigualdade de Minkowski)**.** *Considere $(a_k), (b_k) \in \ell^p$, $p > 1$. Com a introdução da função $\|.\|_p$ na Observação 3.4, vale*

$$\|(a_k) + (b_k)\|_p \leq \|(a_k)\|_p + \|(b_k)\|_p.$$

Demonstração. Primeiramente, observe que

$$\|(a_k) + (b_k)\|_p = (\sum |a_j + b_j|^p)^{1/p} \geq 0.$$

Se este termo é nulo, então

$$0 \leq (\sum |a_j|^p)^{1/p} + (\sum |b_j|^p)^{1/p} = \|(a_k)\|_p + \|(b_k)\|_p,$$

concluindo a demonstração. Portanto, considere o caso em que o lado esquerdo da desigualdade seja positivo. Defina

$$\lambda = \|(b)_k\|_p / (\|(a_k)\|_p + \|(b_k)\|_p).$$

Observe que $0 \leq \lambda \leq 1$. Veja que se $\lambda = 0$, então $\|(b)_k\|_p = 0 \Rightarrow (b_k) = 0$, concluindo a demonstração. O mesmo ocorre para caso $\lambda = 1$.
Considere que $0 < \lambda < 1$. Defina as sequências (x_k) e (y_k) tal que:

$$x_j = a_j/(1-\lambda) \text{ e } y_j = b_j/\lambda, \text{ para todo } j \in \mathbb{N}.$$

Veja que para $z \in \mathbb{R}$ e $p > 1$, real, tem-se que a função $|z|^p$ é convexa. Daí, pela Desigualdade de Jensen, para cada $j \in \mathbb{N}$ tem-se

$$|a_j + b_j|^p = |(1-\lambda)x_j + \lambda y_j|^p \leq (1-\lambda)|x_j|^p + \lambda|y_j|^p = |a_j|^p/(1-\lambda)^{p-1} + |b_j|^p/\lambda^{p-1}.$$

Somando para todos os termos da sequência teríamos:

$$\|(a_k) + (b_k)\|_p^p \leq \|(a_k)\|_p^p / (1-\lambda)^{p-1} + \|(b_k)\|_p^p / \lambda^{p-1}.$$

Substituindo λ, tem-se

$$\|(a_k)\|_p^p / (1-\lambda)^{p-1} + \|(b_k)\|_p^p / \lambda^{p-1} = \|(a_k) + (b_k)\|_p^{p-1} (\|(a_k)\|_p + \|(b_k)\|_p),$$

concluindo a demonstração. □

3.4 Caracterização das Funções Convexas

Os lemas e teoremas a seguir auxiliarão na caracterização de funções convexas.

Lema 3.1. *Seja $f : [a,b] \to \mathbb{R}$ uma função definida num intervalo real $I = [a,b]$ com $a < b$. Se f é convexa, então, para qualquer $c \in (a,b)$, tem-se*

$$\tfrac{f(c)-f(a)}{c-a} \leq \tfrac{f(b)-f(a)}{b-a} \leq \tfrac{f(b)-f(c)}{b-c}.$$

Demonstração. Seja $\lambda \in [0,1]$, com $\lambda = (c-a)/(b-a)$. Observe que:

$$c = (1-\lambda)a + \lambda b.$$

Daí, como f é convexa:

$$f(c) = f((1-\lambda)a + \lambda b) \leq (1-\lambda)f(a) + \lambda f(b).$$

Logo,

$$f(c) - f(a) \leq \lambda(f(b) - f(a)).$$

Substituindo λ, segue a primeira desigualdade. A segunda desigualdade é obtida de maneira análoga considerando $\lambda = (b-c)/(b-a)$. □

Teorema 3.1. *Seja $f : I \to \mathbb{R}$ uma função derivável, onde $I \subseteq \mathbb{R}$ é um intervalo. Então as seguintes afirmações são equivalentes:*

i. A função f é convexa.

ii. A derivada $f' : I \to \mathbb{R}$ é não-decrescente.

iii. Para quaisquer $a, x \in I$ tem-se $f(x) \geq f(a) + f'(a)(x - a)$.

Observação 3.5. *O intervalo $I \subseteq \mathbb{R}$ do Teorema 3.1 é um conjunto convexo, pois é um intervalo na reta real.*

Demonstração. Primeiramente, considere que a função f seja convexa. Sejam $a, b \in I$ quaisquer com $a < b$. Do Lema 3.1, para todo $t \in (a, b)$, tem-se

$$(f(t) - f(a))/(t - a) \leq (f(b) - f(t))/(b - t). \tag{3.1}$$

Tomando o limite para $t \to a^+$, a Equação 3.1 fica:

$$f'(a) \leq (f(b) - f(a))/(b - a).$$

Tomando o limite para $t \to b^-$, a Equação 3.1 fica:

$$f'(b) \geq (f(b) - f(a))/(b - a).$$

Portanto, $f'(a) \leq f'(b)$, concluindo que $i. \Rightarrow ii.$
Agora, se $ii.$ é verdadeiro. Dados $a, x \in I$. Se $a = x$, então $iii.$ segue de imediato. Se $a < x$ pelo Teorema do Valor Médio existe um $c \in (a, x)$, tal que:

$$f'(c) = (f(x) - f(a))/(x - a).$$

Por hipótese, f' é não-decrescente, logo

$$f'(a) \leq f'(c) = (f(x) - f(a))/(x - a),$$

provando $iii.$ Se $x > a$ pelo Teorema do Valor Médio existe um $c \in (x, a)$ tal que:

$$f'(c) = (f(a) - f(x))/(a - x).$$

Por hipótese, f' é não-decrescente, logo

$$f'(a) \geq f'(c) = (f(a) - f(x))/(a - x).$$

A desigualdade em $iii.$ é demonstrada. Por fim, se $iii.$ é verdadeiro. Dados $a, b \in I$ com $a < b$ e $\lambda \in [0, 1]$. O intervalo $[a, b] \subset I$ é convexo, então

$$(1 - \lambda)a + \lambda b = c \in [a, b] \subset I.$$

Logo, isolando λ temos

$$\lambda = (c-a)/(b-a) \text{ e } 1-\lambda = (b-c)/(b-a).$$

Da hipótese temos

$$\begin{aligned} f(a) &\geq f(c) + f'(c)(a-c) \\ f(b) &\geq f(c) + f'(c)(b-c). \end{aligned}$$

Multiplicando a inequação de cima por $(b-c)$ e a de baixo por $(c-a)$, ambos valores não-negativos e somando temos

$$(b-a)f(c) \leq f(a)(b-c) + f(b)(c-a).$$

Substituindo os valores de c e identificando os valores de λ conclui-se

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

Logo f é convexa.

$\square$

Observação 3.6. *Do Teorema 3.1, se a função f for de classe C^2, então f é convexa se, e somente se, $f''(x) \geq 0$ para todo $x \in dom(f)$, pois pelo ítem (b) a derivada primeira é não-decrescente. Veja que é fundamental, para a caracterização de convexidade, que o $dom(f)$ seja convexo. Com efeito, seja $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, tal que $f(x) = x^{-2}$. Observe que $f''(x) > 0$, porém f não é convexa, pois seu domínio não o é.*

Observação 3.7. *O Teorema 3.1 caracteriza funções convexas deriváveis em dimensão 1. Para a caracterização de funções convexas de dimensão finita, superior a 1, precisaremos do seguinte teorema.*

Teorema 3.2. *A função $f : \mathbb{R}^n \to \mathbb{R}$ é convexa se, e somente se, sua restrição a qualquer segmento de reta no seu domínio for convexa. Noutras palavras, dados $a \in dom(f)$ e $v \in \mathbb{R}^n$, tal que $a + tv \in dom(f)$, para $t \in I$, onde $I \subseteq \mathbb{R}$ é um intervalo, então f é convexa, se e somente se, $g : I \to \mathbb{R}$, definida por $g(t) = f(a+tv)$ for convexa, desde que $dom(f)$ seja convexo.*

Demonstração. Suponha que f seja convexa. Então dados $\lambda \in [0,1]$ e $t, s \in I$, tem-se

$$g((1-\lambda)t + \lambda s) = f(a + ((1-\lambda)t + \lambda s)v) = f((1-\lambda)(a+tv) + \lambda(a+sv)).$$

Veja que $a+tv, a+sv \in dom(f)$, então da convexidade de f

$$f((1-\lambda)(a+tv)+\lambda(a+sv)) \leq (1-\lambda)f(a+tv)+\lambda f(a+sv) = (1-\lambda)g(t)+\lambda g(s).$$

Logo g é convexa. Por outro lado, suponha que qualquer restrição a um segmento de reta no $dom(f)$ seja convexo. Então dados $a, b \in dom(f)$ e $\lambda \in [0,1]$, tem-se

$$f((1-\lambda)a + \lambda b) = f(a + \lambda(b-a)).$$

Defina $g : [0,1] \to \mathbb{R}$ com $g(t) = f(a + t(b-a))$. Da convexidade de g tem-se

$$g(\lambda) = g((1-\lambda)0 + \lambda 1) \leq (1-\lambda)g(0) + \lambda g(1) = (1-\lambda)f(a) + \lambda f(b).$$

Observe que da definição $g(\lambda) = f((1-\lambda)a + \lambda b)$, concluindo a demonstração. □

Observação 3.8. *Dos Teoremas 3.1 e 3.2, concluímos que para quaisquer $u, v \in dom(f)$ tem-se que $f \in C^1$ é convexa se, e somente se,*

$$f(v) \geq f(u) + \langle \nabla f(u), v - u \rangle,$$

demonstrando a condição de convexidade para o caso de funções diferenciáveis no $\mathbb{R}^n$. De maneira análoga, para funções de classe C^2, f é convexa se, e somente se, a forma quadrática hessiana de f for não-negativa.

Definição 3.9. *Seja $S \subset \mathbb{R}^n$ um conjunto aberto, convexo e não-vazio. Seja $f : S \to \mathbb{R}$ uma função diferenciável. Dizemos que o ponto $a \in S$ é crítico quando*

$$\nabla f(a) = 0.$$

Teorema 3.3. *Todo ponto crítico de uma função convexa (resp. côncava) $f : S \to \mathbb{R}$ de classe C^1 é um ponto de mínimo global (resp. máximo global).*

Demonstração. Ver [11], página 79, Corolário 7. □

3.5 Exemplos de Funções Convexas e Propriedades

Nesta seção, apresentaremos alguns exemplos de funções convexas reais cuja a sua convexidade pode ser verificada pela aplicação direta dos teoremas da seção anterior.

i. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^{\alpha x}$, onde $\alpha \in \mathbb{R}$, constante.

ii. $p : \mathbb{R} \to \overline{\mathbb{R}}$, $p(x) = \begin{cases} \sum_{j=0}^{m} a_j x^j, \text{ se } x \geq 0, a_j \geq 0, \forall j \geq 2 \\ \infty, \text{ caso contrário} \end{cases}$

iii. $f : \mathbb{R} \to \overline{\mathbb{R}}$, $f(x) = \begin{cases} -\alpha \log x, \text{ se } x > 0, \alpha \geq 0 \\ \infty, \text{ caso contrário} \end{cases}$

iv. $f : \mathbb{R}^n \to \mathbb{R}$, com $f(x) = \alpha\langle Qx, x\rangle$, onde Q é uma matriz quadrada $n \times n$, simétrica e positiva semi-definida, ou seja, para todo $y \in \mathbb{R}^n$, $y^TQy \geq 0$ e $\alpha \geq 0$.

Definição 3.10. *Uma função convexa $f : \mathbb{R} \to \overline{\mathbb{R}}$ é considerada própria se satisfaz as seguintes condições:*

i. $epi(f) \neq \emptyset$.

ii. $dom(f) \neq \emptyset$.

iii. $f(x) > -\infty$, para todo $x \in dom(f)$.

A função que não satisfaz alguma das propriedades acima é dita imprópria. Quando não explicitamente mencionado de forma contrária, nesse trabalho vamos considerar apenas as funções convexas próprias.

Exemplo 3.4. *A função $f : \mathbb{R} \to [-\infty, +\infty]$ definida por*

$$f(x) = \begin{cases} +\infty & se\ |x| < 1, \\ 0 & se\ |x| = 1, \\ -\infty & se\ |x| > 1 \end{cases}$$

é imprópria.

A seguir, apresentamos um importante teorema com respeito às funções convexas.

Teorema 3.4. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função convexa e $\alpha \in [-\infty, \infty]$ qualquer. O conjunto de nível $\{v\colon f(v) \leq \alpha\}$ é convexo.*

Demonstração. Dado um α qualquer, se o conjunto é vazio ou possui apenas um elemento, a conclusão é imediata. Caso contrário, sejam $u, v \in \mathbb{R}^n$ elementos do conjunto de nível e $0 \leq \lambda \leq 1$. Então, pela convexidade de f, tem-se

$$f(\lambda u + (1-\lambda)v) \leq \lambda f(u) + (1-\lambda)f(v) \leq \lambda\alpha + (1-\lambda)\alpha = \alpha.$$

Portanto, $\lambda u + (1-\lambda)v$ é também um elemento do conjunto de nível, concluindo a demonstração. □

Observação 3.9. *Observe que poderíamos ter utilizado a desigualdade estrita na definição do conjunto de nível, concluindo o mesmo resultado. Observe ainda que o conjunto de nível $\{v\colon f(v) \leq \alpha\}$ é a intersecção em $\mathbb{R}^n$ do conjunto $epi(f)$ no hiperplano horizontal $\{(v, \mu)\colon \mu = \alpha\} \subset \mathbb{R}^{n+1}$, tal que o conjunto de nível, pode ser visualizado geometricamente, como uma secção horizontal de $epi(f)$.*

Propriedade 3.6 (Funções Compostas)**.** *Considere uma função convexa* $f : \mathbb{R}^n \to [-\infty, +\infty]$ *e uma função real convexa* $\phi : \mathbb{R} \to [-\infty, +\infty]$ *não-decrescente com* $\phi(+\infty) = +\infty$. *Então* $h = \phi \circ f : \mathbb{R}^n \to [-\infty, +\infty]$ *é convexa.*

Demonstração. Dados $x, y \in \mathbb{R}^n$ e $\lambda \in [0, 1]$. Então, da convexidade de f, tem-se

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

Como ϕ é não-decrescente e convexa, aplicando a função em ambos lados, tem-se

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

Portanto,

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

□

Propriedade 3.7. *Sejam* $f : \mathbb{R}^n \to [-\infty, +\infty]$ *e* $g : \mathbb{R}^n \to [-\infty, +\infty]$ *funções convexas. Dados* $\alpha, \beta \geq 0$ *a função* $h : \mathbb{R}^n \to [-\infty, +\infty]$ *definida por:*

$$h(x) = \alpha f(x) + \beta g(x).$$

é convexa em $dom(h) = dom(f) \cap dom(g) \neq \emptyset$ *convexo.*

Demonstração. Dados $x, y \in dom(f) \cap dom(g)$ e $\lambda \in [0, 1]$. Então, da convexidade de f, g, tem-se

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

Como $\alpha, \beta \geq 0$ tem-se

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

Portanto,

$$\begin{gathered} h((1-\lambda)x + \lambda y) = \alpha f(1-\lambda)x + \lambda y) + \beta g(1-\lambda)x + \lambda y) \leq \\ (1-\lambda)(\alpha f(x) + \beta g(x)) + \lambda(\alpha f(y) + \beta g(y)) = (1-\lambda)h(x) + \lambda h(y). \end{gathered}$$

□

Observação 3.10. *Podemos construir uma função convexa no* $\mathbb{R}^n$ *tomando um subconjunto convexo* $F \subset \mathbb{R}^{n+1}$ *e considerando uma função* f, *tal que*

$$f(x) = \inf\{\mu \colon (x, \mu) \in F\}.$$

Com efeito, dados $(x, \alpha), (y, \beta) \in F$, *com* $f(x) = \alpha$ *e* $f(y) = \beta$. *Como* F *é um conjunto convexo, dado* $\lambda \in [0, 1]$ *tem-se*

$$(1-\lambda)(x,\alpha)+\lambda(y,\beta)\in F \Rightarrow ((1-\lambda)x+\lambda y,(1-\lambda)\alpha+\lambda\beta)\in F.$$

Logo, por definição,

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

concluindo a demonstração.

Teorema 3.5. *Sejam $f_1, f_2, ..., f_k : \mathbb{R}^n \to \mathbb{R}$ funções convexas. Defina $f : \mathbb{R}^n \to \mathbb{R}$ como sendo:*

$$f(x) = \inf\{f_1(x_1)+...+f_k(x_k) \colon x_1+...+x_k = x\}.$$

Então f é uma função convexa.

Demonstração. Defina $F_i = epi(f_i)$. O conjunto $F = F_1 + ... + F_k$ é convexo, pois é a soma de conjuntos convexos (vide Seção 2.4). Por definição de F, temos que se $(x,\mu)\in F$, então $x = x_1+...+x_k$ e $\mu = \mu_1+...+\mu_k$, onde $(x_i,\mu_i)\in F_i$. Logo

$$f_i(x_i)\leq\mu_i.$$

Daí,

$$f(x)\leq f_1(x_1)+...+f_k(x_k)\leq\mu_1+...+\mu_k=\mu.$$

Veja que f pode ser construído como $f(x) = \inf\{\mu \colon (x,\mu)\in F\}$. Conforme Observação 3.10, como F é convexo então f é convexa. □

Dizemos que f definida no teorema anterior é a *convolução infimal* de $f_1, f_2, ..., f_k$. Denotamos por:

$$f = f_1 \star f_2 \star ... \star f_k.$$

Quando há apenas duas funções envolvidas, escrevemos

$$f\star g(x) = \inf_{y\in\mathbb{R}^n}\{f(x-y)+g(y)\}.$$

Veja que, pela demonstração anterior, a propriedade da convolução preserva convexidade.

Definição 3.11. *Seja $\lambda > 0$ e $f : \mathbb{R}^n \to \mathbb{R}$ uma função qualquer. Então a função $(f\lambda) : \mathbb{R}^n \to \mathbb{R}$ é definida por*

$$(f\lambda)(x) = \inf\{\mu \colon (x,\mu)\in\lambda epi(f)\}.$$

A aplicação de transformações lineares a funções convexas, também preservam convexidade, conforme o teorema a seguir:

Teorema 3.6. *Seja A uma transformação linear do $\mathbb{R}^n$ no $\mathbb{R}^m$. Para toda função convexa $g : \mathbb{R}^m \to \mathbb{R}$, a função $gA : \mathbb{R}^n \to \mathbb{R}$, definida por $gA(x) = g(A(x))$ é convexa. Adicionalmente, para toda função convexa $h : \mathbb{R}^n \to \mathbb{R}$, a função $Ah : \mathbb{R}^m \to \mathbb{R}$ definida por $(Ah)(y) = \inf\{h(x) \colon Ax = y\}$ é convexa.*

Demonstração. Ver [16]. □

3.6 Funções Semicontínuas Inferiormente

Definição 3.12. *Uma função $f : \mathbb{R}^n \to (-\infty, +\infty]$ é semicontínua, inferiormente, em $u \in \mathbb{R}^n$, se $f(u)$ existe e dado $\lambda \in \mathbb{R}$ qualquer, tal que $\lambda < f(u)$, existe um $\delta > 0$, tal que:*

$$\lambda < f(v), \text{ para todo } v \in B(u; \delta).$$

A função é dita *semicontínua inferiormente (sci)* se ela for *semicontínua inferiormente* em todo $u \in \mathbb{R}^n$.

Proposição 3.1. *Uma função $f : \mathbb{R}^n \to (-\infty, +\infty]$, é sci em $v \in \mathbb{R}^n$ se, e somente se:*

$$f(v) \leq \lim_{i \to \infty} f(v_i)$$

para toda sequência $\{v_k\}_{k \in \mathbb{N}}$ de $\mathbb{R}^n$ que converge para v e que o limite de $f(v_k)$ existe.

Demonstração. Considere a sequência $\{v_k\}_{k\in\mathbb{N}}$ de $\mathbb{R}^n$ que converge para v e seja

$$\lim_{k\in\mathbb{N}} f(v_k) = L.$$

Suponha que $L < f(v)$. Escolha $L' \in \mathbb{R}$, tal que

$$L < L' < f(v).$$

Como f é *sci*, existe um $\delta > 0$, tal que $f(w) > L'$ para todo $w \in B(v; \delta)$. Como $v_k \to v$, dado δ existe um $k_0 \in \mathbb{N}$, tal que $v_k \in B(v; \delta)$ para todo $k > k_0$. Portanto $f(v_k) > L'$, o que é um absurdo, pois $f(v_k)$ deve-se aproximar de L para k suficientemente grande, não podendo está a uma distância $L' - L$ do seu limite. Por outro lado, considere a sequência $\{v_k\}_{k\in\mathbb{N}}$ de $\mathbb{R}^n$ que converge para v e que o $\lim_{k\in\mathbb{N}} f(v_k) = L$. Seja $\lambda < f(v) \leq L$. Pela convergência da sequência $\{f(v_k)\}$, dado $\epsilon = (L - \lambda)$, existirá $k_0 \in \mathbb{N}$, tal que para todo $k > k_0$, $|f(v_k) - L| < \epsilon = (L - \lambda) \Rightarrow f(v_k) > \lambda$, portanto f é *sci.* □

Para funções *semicontínua inferiormente* podemos anunciar as seguintes propriedades:

Propriedade 3.8. *Sejam* $f, g : \mathbb{R}^n \to (-\infty, +\infty]$ *funções sci e* $\alpha > 0$. *Então a função* $h : \mathbb{R}^n \to (-\infty, +\infty]$ *definida por*

$$h(x) = \alpha f(x) + g(x).$$

é sci.

Demonstração. Dado $u \in \mathbb{R}^n$ qualquer, $h(u)$ está bem definida, pois as funções f, g são por hipótese *sci* em todo $\mathbb{R}^n$. Considere $\lambda \in \mathbb{R}$ tal que $\lambda < h(u)$. Escolha $\lambda_1, \lambda_2 \in \mathbb{R}$ tal que

$$\lambda_1 = (\lambda - g(u))/\alpha.$$
$$\lambda_2 = (\lambda - \alpha f(u)).$$

Observe que por construção, $\lambda_1 < f(u)$ e $\lambda_2 < g(u)$. Logo existem δ_1, δ_2 tal que

$$f(v) < \lambda_1 \text{ para todo } v \in B(u; \delta_1).$$
$$g(v) < \lambda_2 \text{ para todo } v \in B(u; \delta_2).$$

Seja $\delta = \min\{\delta_1, \delta_2\}$. Então, para todo $v \in B(u; \delta)$, temos

$$h(v) = \alpha f(v) + g(v) < \alpha\lambda_1 + \lambda_2 = 2\lambda - h(u) < \lambda.$$

□

Propriedade 3.9. *Seja* $(f_\alpha)_{\alpha \in I}$ *uma família de funções sci definidas em todo o* $\mathbb{R}^n$. *Logo a função* $\sup_{\alpha \in I} f_\alpha$, *definida como:*

$$\left(\sup_{\alpha \in I} f_\alpha\right)(x) = \sup_{\alpha \in I} f_\alpha(x)$$

é sci.

Demonstração. Primeiramente observe que para cada $\alpha \in I$ a função

$$f_\alpha : \mathbb{R}^n \to (-\infty, +\infty]$$

é *sci*, logo está bem definida. Portanto a função

$$h = (\sup_{\alpha \in I} f_\alpha) : \mathbb{R}^n \to (-\infty, +\infty]$$

está bem definida. Dado $u \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, tal que $\lambda < h(u)$. Dado um $\epsilon > 0$, tal que $\lambda < h(u) - \epsilon$, existe pelo menos um índice $\gamma \in I$, tal que

$$f_\gamma(u) > h(u) - \epsilon > \lambda.$$

Daí, existe um $\delta > 0$ tal que

$$\lambda < f_\gamma(v) \text{ para todo } v \in B(u; \delta).$$

Mas, da definição de h temos que $f_\gamma(v) \leq h(v)$, concluindo a demonstração. □

Observação 3.11. *Poderíamos trabalhar com o conceito de semicontinuidade inferior fraca (scif), em que as propriedades acima são, também, satisfeitas. O conceito de scif é semelhante ao anterior, exceto que a convergência da sequência é fraca, o que permite trabalharmos em espaços de Banach reflexivos, em especial, espaços de Hilbert. No caso de estarmos em espaços euclidianos de dimensão finita, a convergência fraca equivale a convergência.*

A partir das definições acima, podemos enunciar o seguinte teorema:

Teorema 3.7. *Seja* $f : \mathbb{R}^n \to (-\infty, +\infty]$. *Então, as seguintes condições são equivalentes:*

i. A função f *é sci em* $\mathbb{R}^n$.

ii. O conjunto $\{u \in \mathbb{R}^n \colon f(u) \leq \beta\}$ *é fechado para todo* $\beta \in \mathbb{R}$.

iii. O $epi(f)$ *é um conjunto fechado em* $\mathbb{R}^{n+1}$.

Demonstração. Primeiramente, se $i.$ é verdadeiro. Dado $\beta \in \mathbb{R}$ qualquer. Se não existisse $u \in \mathbb{R}^n$, tal que $\beta < f(u)$, então o conjunto

$$\{v \in \mathbb{R}^n \colon f(u) \leq \beta\} = \mathbb{R}^n.$$

Noutro caso, dado $v \in \mathbb{R}^n$, tal que $\beta < f(v)$, existe um $\delta > 0$, tal que

$$\beta < f(u), \text{ para todo } u \in B(v; \delta).$$

Então o conjunto $\{u \in \mathbb{R}^n \colon \beta < f(u)\}$ é a união de bolas abertas em $\mathbb{R}^n$, portanto é aberto. Logo seu complementar é fechado, concluindo que $ii.$ é verdadeiro. Considere que $ii.$ seja verdadeiro. Considere o conjunto complementar de $epi(f)$ definido por

$$epi(f)^c = \{(x, \mu) \colon f(x) > \mu\}.$$

Seja $T : \mathbb{R}^{n+1} \to \mathbb{R}^n$ uma transformação linear definida por

$$T(x, a) = x \text{ para } x \in \mathbb{R}^n \text{ e } a \in \mathbb{R}.$$

Afirmamos que T é contínua. Com efeito seja $\{z_k\}_{k \in \mathbb{N}}$ uma sequência do $\mathbb{R}^{n+1}$ que converge para $z = (x, a)$. Escrevemos $z_j = (x_j, a_j)$. Para todo $\epsilon > 0$ existe um $N \in \mathbb{N}$, tal que para todo $m > N$, tem-se

$$\|z_m - z\| < \epsilon.$$

onde $\|.\|$ é a norma Euclidiana. Logo

$$\|x_m - x\| \leq \|z_m - z\| < \epsilon,$$

mostrando que $T(z_k) \to T(z)$. Por hipótese, o conjunto

$$\{u \in \mathbb{R}^n \colon f(u) > \beta\}$$

é aberto. Logo a imagem inversa de um aberto pela aplicação contínua é aberta. Mas veja que a reunião de todos as imagens inversas abertas é exatamente o $epi(f)^c$, concluindo que de fato $epi(f)$ é fechado. Por fim, se *iii.* for verdadeiro. O conjunto

$$epi(f) = \{(x, \mu) \colon f(x) \leq \mu\}$$

é, por hipótese, fechado. Então o conjunto

$$epi(f)^c = \{(x, \mu) \colon f(x) > \mu\}$$

é aberto. Dado $u \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, tal que $\lambda < f(u)$, temos que $(u, \lambda) \in epi(f)^c$, então existe um $\delta > 0$, tal que

$$B((u, \lambda); \delta) \subset epi(f)^c.$$

Afirmamos que

$$\lambda < f(v) \text{ para todo } v \in B(u; \delta).$$

Com efeito, temos que

$$\|(v, \lambda) - (u, \lambda)\| = \|v - u\| < \delta.$$

Portanto $(v, \lambda) \in B((u, \lambda); \delta) \subset epi(f)^c$, concluindo que f é *sci*. □

3.7 Continuidade e Diferenciabilidade de Funções Convexas

O teorema que apresentamos na sequência é um importante resultado sobre continuidade de funções convexas. Mas antes de enunciar e provar o teorema, considere as seguintes definições e lemas:

Definição 3.13. *Seja $f : \mathbb{R}^n \to [-\infty, +\infty]$. Suponha que a dimensão de $dom(f)$ seja n. Defina $g : \mathbb{R}^n \to [-\infty, +\infty]$ uma função tal que:*

$$g(x) = \begin{cases} f(x) & \text{se } x \in int(dom(f)). \\ \mu & \text{se } x \in bd(dom(f)) \text{ e } (x, \mu) \in bd(epi(f)). \end{cases} \tag{3.2}$$

Observe que $epi(g) = cl(epi(f))$. Dizemos que $g = cl(f)$ e chamamos g da envoltória *sci* de f. Ainda, dizemos que f é uma função fechada quando $f = cl(f)$ o que é equivalente ao conjunto $epi(f)$ ser fechado.

Observação 3.12. *De acordo com o Teorema 3.7 f é fechada se, e somente se, f é sci.*

Observação 3.13. *Quando o domínio da função f é um conjunto fechado então o $epi(f)$ é fechado.*

Lema 3.2. *Seja $f : \mathbb{R}^n \to [-\infty, +\infty]$ uma função convexa, própria. Então a função $cl(f)$, conforme Definição 3.13, é uma função convexa, própria e fechada, tal que os valores de $cl(f)$ coincide com f, exceto em alguns pontos da fronteira relativa do $dom(f)$.*

Demonstração. Ver [16], Seção 7, Teorema 7.4. □

Teorema 3.8 (Continuidade)**.** *A função convexa $f : \mathbb{R}^n \to [-\infty, +\infty]$ é contínua em qualquer conjunto aberto e convexo $C \subset dom(f)$.*

Demonstração. Podemos definir, se necessário, uma outra função que seja definida em C e identicamente igual a f em C. Portanto, basta demonstrar para o caso de $C = ri(dom(f))$. Adicionalmente, podemos considerar o caso em que a dimensão de $dom(f) = n$, caso contrário basta tomar uma transformação linear que leva o conjunto em dimensão inferior em $\mathbb{R}^n$ que preserva as propriedades de convexidade. Com essas considerações, a demonstração basta em $C = int(dom(f))$. Pelo Lema 3.2, $cl(f)(x) = f(x)$ para $x \in C$. Portanto, f é *sci* em C. Daí, precisamos demonstrar que f é *semicontínua superior* em C o que é equivalente ao conjunto

$$F = \{(x, a) \in C \times \mathbb{R} \colon f(x) \geq a\}$$

ser fechado. Mas veja que o conjunto $int(epi(f))$ definido por

$$int(epi(f)) = \{(x, \mu) \in C \times \mathbb{R} \colon f(x) < \mu\}$$

é aberto, pois C é aberto. Então, para qualquer $a \in \mathbb{R}$, o conjunto

$$F^c = \{(x, a) \in C \times \mathbb{R} \colon f(x) < a\}$$

é a intersecção entre $int(epi(f))$ e o semiespaço aberto

$$A = \{(x, \mu) \in \mathbb{R}^{n+1} \colon \mu < a\}$$

sendo, portanto, um conjunto aberto em $\mathbb{R}^{n+1}$. Logo seu complementar é fechado. □

Corolário 3.9. *Seja $f : \mathbb{R}^n \to [-\infty, \infty]$ uma função convexa com $int(dom(f)) \neq \emptyset$. Dado $x \in int(dom(f))$ existe um $\epsilon > 0$, tal que*

$$f \text{ é contínua em } B(x; \epsilon) \subset int(dom(f)).$$

Corolário 3.10. *Se a função convexa $f : \mathbb{R}^n \to [-\infty, \infty]$ for finita em todo $\mathbb{R}^n$, então ela é contínua no $\mathbb{R}^n$.*

Por fim, apresentaremos uma propriedade interessante com respeito a existência de derivada num sentido para funções convexas.

Propriedade 3.10 (Existência da derivada direcional)**.** *Seja* $f : \mathbb{R}^n \to \mathbb{R}$ *uma função convexa com domínio com interior não vazio. Seja* $x \in int(dom(f))$ *e* $y \in \mathbb{R}^n$*, tal que exista um* $\lambda > 0$*, com* $x + ty \in dom(f)$*, para todo* $|t| < \lambda$*. Então o seguinte limite*

$$f'(x, y) = \lim_{t \to 0^+} (f(x + ty) - f(x))/t.$$

existe.

Demonstração. Defina as funções $g_y : \mathbb{R}^+ \to \mathbb{R}$ e $h_y : \mathbb{R}^+ \to \mathbb{R}$, tal que

$$\begin{aligned} g_y(t) &= (f(x + ty) - f(x))/t \text{ e} \\ h_y(t) &= (-f(x - ty) + f(x))/t, \end{aligned}$$

para $|t| < \lambda$. Afirmamos que g_y é monótona não-decrescente e h_y é monótona não-crescente. Com efeito, dados $\lambda > s > t > 0$, tem-se que $x + ty$ pode ser escrito como uma combinação convexa de x e $x + sy$, contido no domínio de f, portanto,

$$f(x + ty) = f((1 - t/s)x + t/s(x + sy)) \leq (1 - t/s)f(x) + t/sf(x + sy) \Rightarrow$$
$$(f(x + ty) - f(x))/t \leq (f(x + sy) - f(x))/s \Rightarrow g_y(t) \leq g_y(s).$$

De maneira análoga, tem-se $h_y(s) \leq h_y(t)$. Além disso, da convexidade de f temos que $2f(x) \leq f(x + 2ty) + f(x - 2ty)$, sempre que $0 < 2t < \lambda$. Daí, $h_y(2t) \leq g_y(2t)$. Logo g_y possui uma cota inferior e como h_y é não-crescente, existirá o

$$\inf_{t>0}(f(x + ty) - f(x))/t.$$

□

3.8 A Conjugada de Funções Convexas

Há duas formas de caracterizar uma curva no $\mathbb{R}^n$. A primeira é como o lugar geométrico de pontos e a segunda é como envelope de tangentes. Veja na figura a seguir, um exemplo do gráfico de uma cônica no $\mathbb{R}^2$.

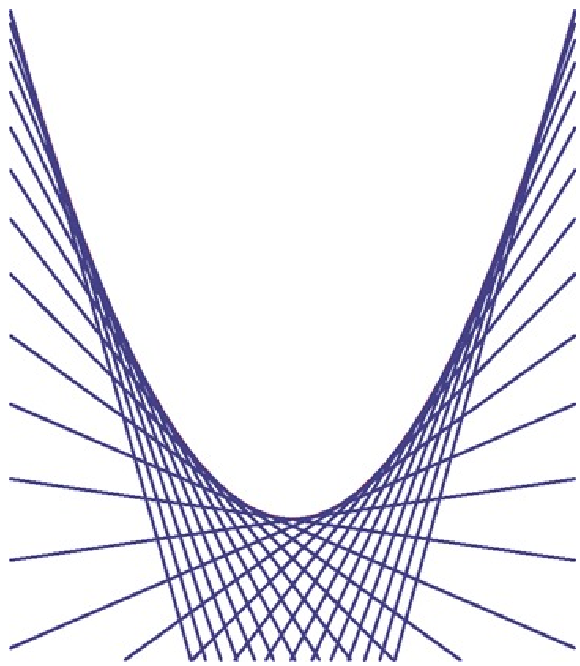

Figura 3.3: A figura mostra uma parábola no $\mathbb{R}^2$ representada pelo envelope de tangentes.

De fato a observação acima foi demonstrada no Teorema 2.17. A ideia do conjugado vem diretamente da afirmação que o epigráfico de uma função convexa própria fechada é um conjunto convexo e fechado, o que pode ser entendido como a intersecção de semiespaços fechados do $\mathbb{R}^{n+1}$ que contém este conjunto.

Teorema 3.11. *Uma função convexa e fechada $f : \mathbb{R}^n \to (-\infty, \infty]$ é o supremo, ponto a ponto, de uma coleção de funções afins $h : \mathbb{R}^n \to \mathbb{R}$, tal que*

$$h(x) \leq f(x) \text{ para todo } x \in dom(f).$$

Demonstração. Ver [16], Seção 12, Teorema 12.1. □

Observação 3.14. *De fato, foi demonstrado no Teorema 2.17, que um conjunto convexo e fechado, é a intersecção dos semiespaços fechados que contém o mesmo. Portanto a dualidade entre funções convexas fechadas e funções afins tangentes é equivalente a dualidade entre o gráfico da função e hiperplanos tangentes que é equivalente a dualidade entre epigráficos convexos fechados e os semiespaços que contém o mesmo.*

Exemplo 3.5. *Para exemplificar o Teorema 3.11, considere a seguinte função convexa e fechada:*

$$f : [0,1] \to [0,1], \text{ tal que } f(x) = x^2.$$

Logo o gráfico de f em $\mathbb{R}^2$ pode ser visto como o lugar geométrico dos pontos (x, x^2), com $x \in [0,1]$. Analogamente, o gráfico de f pode ser observado como o supremo ponto a ponto das funções afim $h_c(x) = 2x + c$, tangentes ao gráfico de f em cada ponto do domínio, veja a figura a seguir:

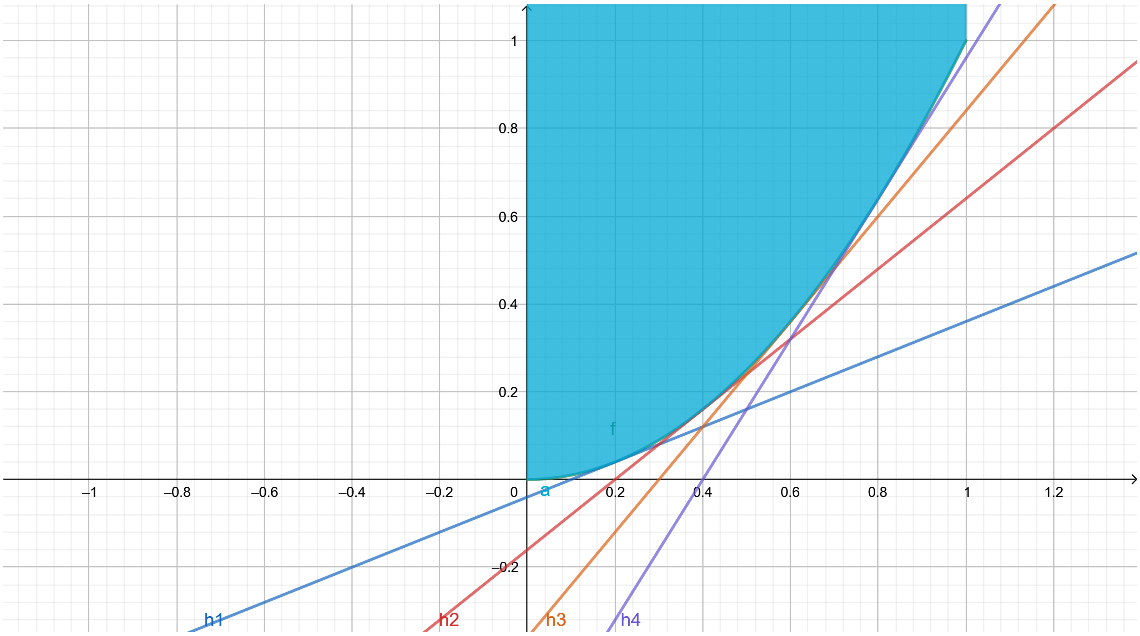

Figura 3.4: A figura mostra o gráfico da função f e das funções tangentes ponto a ponto h cuja a interseção dos semiplanos superiores resultará $epi(f)$.

Definição 3.14. *Seja $f : \mathbb{R}^n \to [-\infty, \infty]$ uma função convexa e fechada. Defina $f^* : \mathbb{R}^n \to [-\infty, \infty]$ por*

$$f^*(x^*) = \sup_{x \in \mathbb{R}^n} \{\langle x, x^* \rangle - f(x)\}.$$

Dizemos que f^* é a função conjugada de f.

Observação 3.15. *Observamos que a função f^* está bem definida. Com efeito, $epi(f)$ é um conjunto convexo e fechado, portanto, do Teorema 2.14, para cada $(x, \mu) \in bd(epi(f))$ existe um semiespaço*

$$HS = \{y \in \mathbb{R}^{n+1} \colon \langle y, w \rangle \geq b\},$$

onde $w \in \mathbb{R}^{n+1}$ e $b \in \mathbb{R}$, tal que

$$\langle (x, \mu), w \rangle = b \ \text{e}$$
$$\langle (z, \alpha), w \rangle \geq b.$$

para todo $(z, \alpha) \in epi(f)$. *Em particular*

$$\langle (z, f(z)), w \rangle \geq b,$$

para todo $z \in dom(f)$. *Suponha sem perda de generalidade que a* $(n+1)$*-ésima coordenada de* w *seja não nula. A inequação pode ser rescrita como*

$$f(z) \geq \langle z, w' \rangle - b', \text{ para todo } z \in dom(f),$$

onde $w' \in \mathbb{R}^n$ *e* $b' \in \mathbb{R}$. *Logo* b' *é uma cota superior para a função* $g : \mathbb{R}^n \to \mathbb{R}$*:*

$$g(z) = \langle z, w' \rangle - f(z).$$

Portanto o $\sup_{z \in \mathbb{R}^n} g(z)$ *existe e denotaremos pela função* $f^*(w')$.

Exemplo 3.6. *Seja* $f : \mathbb{R} \to \mathbb{R}$ *definida por* $f(x) = |x|$. *Observe que* f *é convexa e fechada. A conjugada* $f^* : \mathbb{R} \to \mathbb{R}$ *é definida por*

$$f^*(y) = \begin{cases} 0, & se\ |y| \leq 1. \\ \infty, & se\ |y| > 1. \end{cases}.$$

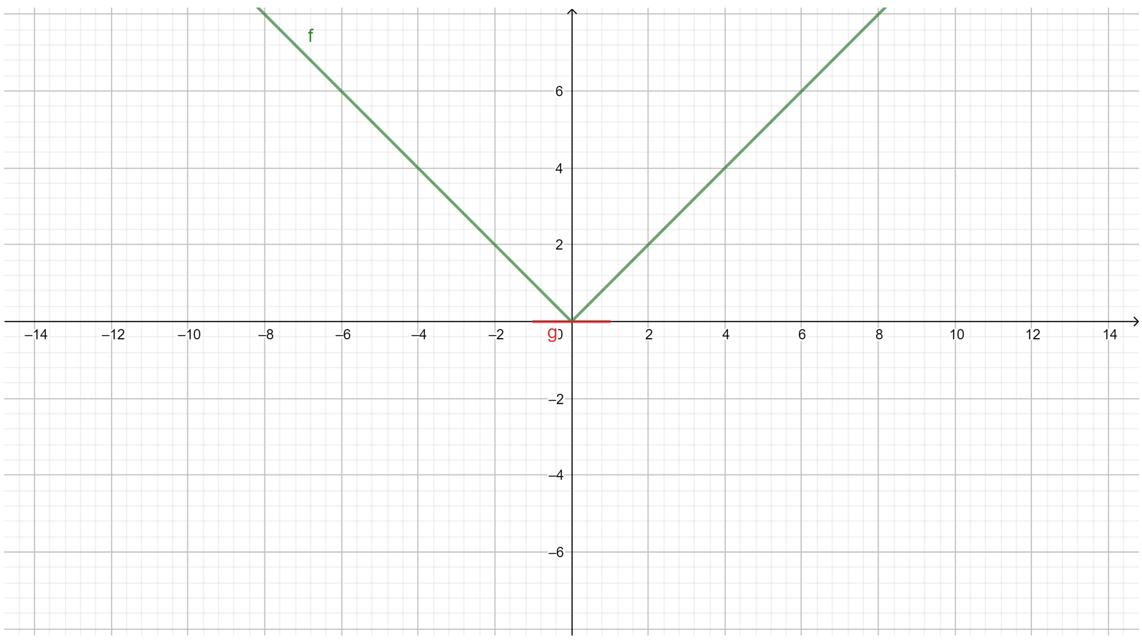

Figura 3.5: A figura mostra o gráfico da função f em verde e da sua conjugada em vermelho conforme definida no Exemplo 3.6.

Propriedade 3.11. *Dado* $\lambda \geq 0$ *e* f *uma função convexa própria, então* $(\lambda f)^* = f^* \lambda$ *e* $(f\lambda)^* = \lambda f^*$. *Observe que essa propriedade segue diretamente da definição do supremo.*

Definição 3.15. *Sejam* $f_1, ..., f_m$ *funções convexas e próprias do* $\mathbb{R}^n$*. Defina o operador* $\Box$*, como sendo*

$$(f_1 \Box ... \Box f_m)^*(x^*) = \sup_x \{\langle x, x^* \rangle - \inf_{x_1+....+x_m=x}\{f_1(x_1) + ... + f_m(x_m)\}\}.$$

Propriedade 3.12. *Sejam* $f_1, ..., f_m$ *funções convexas e próprias do* $\mathbb{R}^n$*. Então,*

$$(f_1 \Box ... \Box f_m)^* = f_1^* + ... + f_m^*.$$

Demonstração. Utilizando as propriedades de supremo e ínfimo temos:

$$\sup_x \{\langle x, x^* \rangle - \inf_{x_1+....+x_m=x}\{f_1(x_1) + ... + f_m(x_m)\}\} =$$
$$\sup_x \sup_{x_1+....+x_m=x}\{\langle x, x^* \rangle - f_1(x_1) - ... - f_m(x_m)\} =$$
$$\sup_{x_1,...,x_m}\{\langle x_1, x^* \rangle + ... + \langle x_m, x^* \rangle - f_1(x_1) - ... - f_m(x_m)\} = f_1^*(x_1^*) + ... + f_m^*(x_m^*).$$

$\Box$

Capítulo 4

Transformada de Legendre

Neste capítulo vamos apresentar de forma conceitual e prática a transformada de Legendre, introduzido pelo matemático francês Adrien-Marie Legendre (1752-1833), em seu trabalho denominado *L'intégration de quelques équations aux différences Partielles.* As motivações para o uso da transformada de Legendre são extensas e englobam seu aspecto teórico e prático. Na matemática aplicada, apresentamos exemplos na mecânica clássica e economia. Por fim, apresentaremos algumas propriedades e uma tabela com as funções convexas e suas respectivas transformadas. Os conceitos, as definições e os teoremas desse capítulo foram consultados nas referências [9, 15, 16].

4.1 Definição

Sejam $S \subseteq \mathbb{R}^n$ um conjunto aberto, convexo e $f : S \to \mathbb{R}$ uma função convexa de classe C^1 com gradiente $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ uma transformação linear invertível (injetiva). Suponha que o conjunto

$$S^* = \{s \in \mathbb{R}^n \colon s = \nabla f(x), \text{ para } x \in \mathbb{R}^n\}$$

seja não-vazio. Definimos a transformada de Legendre de f como sendo a função $F : S^* \to \mathbb{R}$ definida implicitamente pela igualdade

$$F(\nabla f(x)) = \langle \nabla f(x), x\rangle - f(x).$$

Observação 4.1. *Veja que as Definições em 4.1 e em 3.14 não são distintas. De fato, no caso em que a função é de classe C^1 o supremo da função definida em 3.14 é atingido em $x^* = \nabla f(x)$, pois a função $\langle x^*, x\rangle - f(x)$ é de classe C^1 e côncava em x (é um ponto de máximo global, vide Teorema 3.3). No capítulo seguinte trabalharemos o conceito quando a função não é diferenciável, onde apresentaremos a Transformada de Fenchel. A apresentação da Transformada de Legendre em capítulo* *anterior à apresentação da Transformada de Fenchel deve-se à razões históricas. Adicionalmente, a condição de invertibilidade do operador gradiente é essencial para a Definição 4.1.*

4.2 A Obtenção da Transformada de Legendre

No exemplo seguinte, mostraremos como obter a transformada de legendre a partir de uma função convexa real.

Exemplo 4.1 (Função Quadrática)**.** *Seja $f : \mathbb{R} \to \mathbb{R}$, tal que $f(x) = ax^2 + bx + c$, onde $a > 0$. Claramente, f é convexa, pois $f''(x) = 2a > 0$ e o domínio de f é a reta real que é convexa. Então, a transformada de legendre de f é $F : \mathbb{R} \to \mathbb{R}$ definida por $F(s) = (a(s-b)^2 - 4a^2c)/4a^2$. Com efeito, para obtenção da transformada de Legendre temos que fixado s, encontrar o supremo da função $sx - f(x)$, que é uma função quadrática em x, onde o valor máximo é obtido de quando $s - f'(x) = 0$, ou seja, quando $x = (s-b)/2a$. A substituição do valor de x, na definição da transformada, resulta na função apresentada.Observe que a inversa da derivada de f é a derivada de f^*. Na verdade, essa é uma importante propriedade da transformada de Legendre, conforme será demonstrada adiante neste capítulo.*

Exemplo 4.2. *Seja $f : \mathbb{R}^2 \to \mathbb{R}$ definida por*

$$f(x, y) = x^2 + y^2.$$

A transformada de Legendre de f é a função $F : \mathbb{R}^2 \to \mathbb{R}$, definida por:

$$F(x', y') = (x'^2 + y'^2)/4.$$

Com efeito, dado (x, y) no domínio de f então a variável $s = (x', y')$ do domínio de F será $s = (x', y') = \nabla f(x, y) = (2x, 2y)$. Portanto, substituindo $(x, y) = (x'/2, y'/2)$ na definição da transformada, temos:

$$F(x', y') = \langle (x', y'), (x'/2, y'/2) \rangle - (x'/2)^2 - (y'/2)^2 = (x'^2 + y'^2)/4.$$

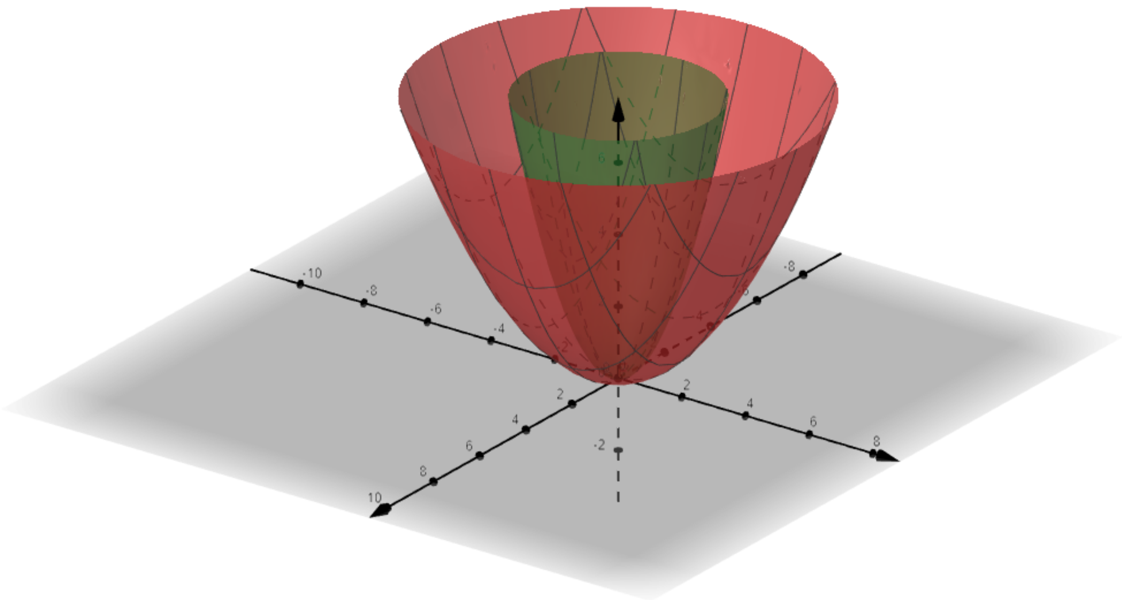

Figura 4.1: A figura mostra o gráfico da função f em verde e da sua transformada de Legendre F em vermelho conforme definido no Exemplo 4.2.

4.3 Propriedades da Transformada de Legendre

Propriedade 4.1. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função convexa de classe C^1, tal que o operador gradiente seja invertível. Defina $F : \mathbb{R}^n \to \mathbb{R}$ a transformada de Legendre de f. Então as seguintes propriedades são válidas:*

i. $(\nabla f)^{-1} = \nabla F : \mathbb{R}^n \to \mathbb{R}^n$.

ii. $f(x) + F(s) \geq \langle x, s\rangle$, *para todos* $x, s \in \mathbb{R}^n$.

Demonstração. Por definição da transformada de Legendre de f:

$$F(s) = \langle s, x\rangle - f(x)$$

é uma função de classe C^1, com $s = \nabla f(x) \in \mathbb{R}^n$. Como ∇f é um operador linear invertível, então podemos escrever $x = x(s) = (\nabla f)^{-1}(s)$. Portanto, aplicando a regra da cadeia, temos

$$\nabla F(s) = x(s) + s^T x'(s) - f'(x(s))^T x'(s) = x(s) = (\nabla f)^{-1}(s), \text{ para todo } s \in \mathbb{R}^n.$$

Aqui usamos as notações f' e ∇f para representar o operador linear *diferencial* de f, usamos a notação do elemento do $\mathbb{R}^n$ como vetor coluna $n \times 1$ e usamos como domínio $\mathbb{R}^n$ para facilitar notação. A demonstração é análoga para os casos em que $dom(f) \subset \mathbb{R}^n$, concluindo a demonstração de i.

Adicionalmente, por definição, o valor máximo (supremo) da função côncava de classe C^1, $\langle s, x\rangle - f(x)$ em x é atingido, conforme Teorema 3.3, quando $x = \nabla F(s)$, então $F(s) \geq \langle s, x\rangle - f(x)$, para todo $x \in \mathbb{R}^n$. De maneira análoga, $f(x) \geq \langle s, x\rangle - F(s)$, para todo $s \in \mathbb{R}^n$, concluindo a demonstração de *ii*. $\square$

4.4 Tabela com Transformadas de Legendre

Nesta seção, listamos uma tabela contendo alguns exemplos de funções convexas e suas transformadas de Legendre, quando esta pode ser determinada de forma explícita. O domínio das funções e das suas transformadas, quando não explícito, será o domínio convexo onde a função está bem definida.

Tabela de Transformada de Legendre

$f(x)$	$F(s)$
$e^{\alpha x}$, $\alpha > 0$	$(s/\alpha)(\log(s/\alpha) - 1)$
$-\alpha \log(x)$, $\alpha > 0$	$-\alpha - \alpha \log(-\alpha s)$
$\log(1 + e^x)$	$s \log(s) + (1 - s)\log(1 - s)$
$ax^2 + bx + c$, $a > 0$	$(a(s - b)^2 - 4a^2c)/4a^2$
$cx^{-\alpha}$, $c, \alpha > 0$	$s(-\alpha c/s)^{1/(\alpha+1)} - c(-\alpha c/s)^{-\alpha/(\alpha+1)}$
$\|x\|^p/p$, $p > 1$	$\|s\|^q/q$, $p^{-1} + q^{-1} = 1$
$c \sin(x)$, $c > 0$	$s \arccos(s/c) + \sqrt{c^2 - s^2}$
$c \cos(x)$, $c > 0$	$s \arcsin(-s/c) + \sqrt{c^2 - s^2}$
$c \tan(x)$, $c > 0$	$s \arccos(\sqrt{c/s}) - \sqrt{cs - c^2}$
$c \cosh(x)$, $c > 0$	$s \sinh^{-1}(s/c) - \sqrt{c^2 + s^2}$
$\alpha\langle Qx, x\rangle$, $\alpha > 0$, Q *psdi*	$\langle Q^{-1}s, s\rangle/4\alpha$

A seguir, complementamos a tabela acima com os respectivos domínio das funções, respeitando a ordem da tabela anterior.

Tabela dos domínios das funções

$f(x)$	$dom(f(x))$	$dom(F(s))$
$e^{\alpha x}$, $\alpha > 0$	$\mathbb{R}$	$(0, \infty)$
$-\alpha \log(x)$, $x > 0, \alpha > 0$	$(0, \infty)$	$(-\infty, 0)$
$\log(1 + e^x)$	$\mathbb{R}$	$(0, 1)$
$ax^2 + bx + c$, $a > 0$	$\mathbb{R}$	$\mathbb{R}$
$cx^{-\alpha}$, $c, \alpha > 0$	$(0, \infty)$	$(-\infty, 0)$
$\|x\|^p/p$, $p > 1$	$\mathbb{R}$	$\mathbb{R}$
$c \sin(x)$, $c > 0$	$[\pi, 3\pi/2]$	$[-c, c]$
$c \cos(x)$, $c > 0$	$[\pi, 3\pi/2]$	$[-c, c]$
$c \tan(x)$, $c > 0$	$[0, \pi/2)$	$[c, \infty)$
$c \cosh(x)$, $c > 0$	$(0, \infty)$	$[c, \infty]$
$\alpha\langle Qx, x\rangle$, $\alpha > 0$, Q *psdi*	$\mathbb{R}^n$	$\mathbb{R}^n$

4.5 Aplicações da Transformada de Legendre

Nos exemplos a seguir, mostraremos a aplicação da transformada de Legendre em outros campos do conhecimento como a mecânica clássica e a economia.

Exemplo 4.3 (Mecânica Clássica)**.** *Em física, na mecânica clássica, trabalhamos com equações de posição e velocidade de partículas em relação a um sistema de coordenadas cartesianas. Os elementos a serem trabalhados são vetores e a depender da escolha do sistema de coordenadas, o problema pode se tornar extremamente difícil de trabalhar. Por exemplo, num sistema de N partículas, onde para cada partícula tem-se 3 variáveis para a posição e 3 variáveis para velocidade, o total de variáveis do sistema é 6N. Uma forma de tentar contornar as dificuldades e complicações apresentadas é trabalhar com energia, em vez de forças, ou seja, transformar grandezas vetoriais em grandezas escalares. Primeiramente, veja que num sistema conservativo n-dimensional tem-se que:*

$$-\nabla V = \overrightarrow{F}$$

onde $V : \mathbb{R}^n \to \mathbb{R}$ *é uma função potencial definida num sistema n-dimensional e* $\overrightarrow{F} = \boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^n$ *é uma aplicação n-dimensional. Neste exemplo utilizaremos as notações* $\boldsymbol{a}$ *(a em negrito) para representar um vetor* $\overrightarrow{a}$ *no espaço n-dimensional e* $\cdot$ *para representar o produto interno entre vetores, ou seja* $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{a} \cdot \boldsymbol{b}$*. Para um movimento infinitesimal da posição* $\boldsymbol{r} = (r_1, r_2, ..., r_n)$*,* $\delta\boldsymbol{r} = (\delta r_1, \delta r_2, ..., \delta r_n)$*, de uma partícula de massa constante m, tem-se*

$$\boldsymbol{F} \cdot \delta\boldsymbol{r} = m\frac{d^2\boldsymbol{r}}{dt^2} \cdot \delta\boldsymbol{r},$$

onde: $\cdot$ *é uma representação do produto interno entre dois vetores do* $\mathbb{R}^n$*. Utilizando a regra do produto, a equação do lado direito pode ser reescrita como:*

$$m\left[\frac{d}{dt}\left(\frac{d\boldsymbol{r}}{dt} \cdot \delta\boldsymbol{r}\right) - \frac{d\boldsymbol{r}}{dt} \cdot \frac{d\delta\boldsymbol{r}}{dt}\right] = m\sum_k \left[\frac{d}{dt}\left(\frac{d\boldsymbol{r}}{dt} \cdot \frac{\partial \boldsymbol{r}}{\partial r_k}.\delta r_k\right) - \frac{d\boldsymbol{r}}{dt} \cdot \frac{\partial \dot{\boldsymbol{r}}}{\partial r_k}.\delta r_k\right]$$

onde utilizamos a notação $\dot{\boldsymbol{a}}$ *para representar a derivada de* $\boldsymbol{a}$ *em relação à variável tempo t. É fácil verificar pela regra da cadeia que a seguinte relação é válida:*

$$\frac{\partial \boldsymbol{r}}{\partial r_k} = \frac{\partial \dot{\boldsymbol{r}}}{\partial \dot{r}_k}$$

e considerando que o vetor velocidade $\boldsymbol{v} = \dot{\boldsymbol{r}}$ *ou* $\dot{r_k} = v_k$ *e definindo a energia cinética como sendo* $T = 0.5\,(m\boldsymbol{v} \cdot \boldsymbol{v})$*, tem-se que a equação do lado direito fica:*

$$m\sum_k \left[\frac{d}{dt}\left(\boldsymbol{v} \cdot \frac{\partial \boldsymbol{v}}{\partial v_k}\right) - \boldsymbol{v} \cdot \frac{\partial \boldsymbol{v}}{\partial r_k}\right]\delta r_k = \sum_k \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{r}_k}\right) - \frac{\partial T}{\partial r_k}\right]\delta r_k.$$

Combinando as equações, tem-se

$$\sum_k \left[-\frac{\partial V}{\partial r_k}\right]\delta r_k = \sum_k \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{r}_k}\right) - \frac{\partial T}{\partial r_k}\right]\delta r_k.$$

Daí,

$$\Rightarrow \sum_k \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{r}_k}\right) - \frac{\partial (T-V)}{\partial r_k}\right]\delta r_k = 0.$$

Em um sistema conservativo, tem-se que a função potencial depende apenas da posição, portanto: $\partial V/\partial \dot{r}_k = 0$*. A equação pode ser reescrita, definido a função Lagrangiana* $L : \mathbb{R}^{2n+1} \to \mathbb{R}$*, definida por* $L = L(t, \boldsymbol{r}, \dot{\boldsymbol{r}}) = T(t, \boldsymbol{r}, \dot{\boldsymbol{r}}) - V(t, \boldsymbol{r})$*, como:*

$$\sum_k \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}_k}\right) - \frac{\partial L}{\partial r_k}\right]\delta r_k = 0$$

onde $\delta \boldsymbol{r}$ *é tomado de forma arbitrária. Podemos tomar para cada* k*, a variação* $\delta \boldsymbol{r} = (0, ..., \delta r_k, ..., 0)$*, implicando na equação:*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}_k}\right) - \frac{\partial L}{\partial r_k} = 0$$

para $k = 1, 2, ..., n$*. Em 1833, o matemático William Rowan Hamilton reformulou as equações da mecânica clássica, introduzindo o conceito da função Hamiltoniana, a partir da Lagrangiana, através da transformada de Legendre, ou seja:* $L : \mathbb{R}^{2n+1} \to \mathbb{R}$ *a lagrangiana de um sistema, em que as variáveis são posição* $\boldsymbol{r}$*, velocidade* $\dot{\boldsymbol{r}} = \boldsymbol{v}$ *e o tempo* t*. A Hamiltoniana* $H : \mathbb{R}^{2n+1} \to \mathbb{R}$ *é definida por:*

$$H(t, \boldsymbol{r}, \boldsymbol{p}) = \boldsymbol{p} \cdot \boldsymbol{v} - L(t, \boldsymbol{r}, \boldsymbol{v})$$

ou seja, a função $H(t, \boldsymbol{r}, .) : \mathbb{R}^n \to \mathbb{R}$ *é a transformada de Legendre da função* $L(t, \boldsymbol{r}, .) : \mathbb{R}^n \to \mathbb{R}$*, onde a variável* $\boldsymbol{p} = (p_1, ..., p_n)$ *é definida como:*

$$p_k = \frac{\partial L}{\partial v_k} = \frac{\partial T}{\partial v_k} = m v_k$$

sendo o momento conjugado da variável velocidade. Observe que essa é exatamente a definição da Transformada de Legendre introduzida em 4.1. A partir dessa definição podemos derivar as equações diferenciais de Hamilton. Com efeito:

$$dL = \sum_k \left[\frac{\partial L}{\partial r_k} dr_k + \frac{\partial L}{\partial \dot{r}_k} d\dot{r}_k\right] + \frac{\partial L}{\partial t} dt = \sum_k \left[\frac{\partial L}{\partial r_k} dr_k + d(p_k \dot{r}_k) - \dot{r}_k dp_k\right] + \frac{\partial L}{\partial t} dt$$

rearranjando a equação, tem-se

$$d(\sum_k p_k \dot{r}_k - L) = \sum_k \left[-\frac{\partial L}{\partial r_k} dr_k + \dot{r}_k dp_k\right] - \frac{\partial L}{\partial t} dt$$

porém, o lado esquerdo da equação é exatamente o diferencial dH, que pode ser rescrito como:

$$dH = \sum_k \left[\frac{\partial H}{\partial r_k} dr_k + \frac{\partial H}{\partial p_k} dp_k\right] + \frac{\partial H}{\partial t} dt$$

donde concluímos que:

$$\frac{\partial H}{\partial r_k} = -\frac{\partial L}{\partial r_k} \text{ ; } \frac{\partial H}{\partial p_k} = \dot{r}_k \text{ e } \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$$

Utilizando a equação de Legendre, temos que:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}_k}\right) = \frac{\partial L}{\partial r_k}.$$

e sabendo que $\frac{\partial L}{\partial \dot{r}_k} = p_k$, conclui-se que:

$$\frac{\partial H}{\partial r_k} = -\dot{p}_k \text{ ; } \frac{\partial H}{\partial p_k} = \dot{r}_k \text{ e } \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$$

Exemplo 4.4 (Microeconomia)**.** *Defina $l : \mathbb{R}^+ \to \mathbb{R}$ a função lucro de uma firma para um determinado produto pode ser modelado por $l(p) = \sup_{p>0}(pq - c(q))$, onde p é a variável de preço, q é variável quantidade e $c : \mathbb{R} \to \mathbb{R}$ é a função custo. Assumindo que c é uma função convexa de classe C^1, entramos nas hipóteses da definição da transformada de Legendre, onde a função l é a transformada de Legendre da função custo c.*

Exemplo 4.5 (Termodinâmica)**.** *A energia interna U é uma função fundamental de estado da Termodinâmica. U é uma função das variáveis de um sistema, por exemplo a entropia S, o volume V e o componente molar da massa N_i. Essas são denominadas variáveis extensivas. Por outro lado, as seguintes variáveis são denominadas intensivas: Pressão P, Temperatura T e Potencial Químico μ_i. As variáveis intensivas são mais fáceis de medir na prática, em laboratórios. Considere o caso de um sistema simplificado, como a fundição de uma liga metálica, contendo n elementos, não sujeita a campos externos que afetem sua energia eletromagnética e tensorial. A energia interna do sistema pode ser modelada pela função $U : \mathbb{R}^3 \to \mathbb{R}$, como $U(S, V, N_i) = TS - PV + \mu_i N_i$. Considere a função $u : \mathbb{R} \to \mathbb{R}$, fixados V e N_i, $u(S) = U(S, V, N_i)$. Suponha que u seja convexo, de classe C^1 e que sua derivada seja invertível. Então, a transformada de Legendre de u, definida por $h : \mathbb{R} \to \mathbb{R}$, $h(T) = TS - u(S)$, onde $T = u'(S)$, em que a variável passa a ser a temperatura em vez de entropia. A função correspondente, denominada de Potencial de Helmholtz é a função $H : \mathbb{R}^3 \to \mathbb{R}$, em que $H(T, V, N_i) = TS(T) - U(S(T), V, N_i)$.*

Capítulo 5

Transformada de Fenchel

A transformada de Fenchel, também conhecida na literatura como transformada de Legendre-Fenchel, generaliza os conceitos apresentados no capítulo anterior. Os conceitos, as definições e os teoremas desse capítulo foram consultados nas referências [15, 16].

5.1 Caracterização de Funções Convexas e Semicontínuas Inferiormente

Antes de iniciar o desenvolvimento da Transformada de Fenchel, vamos apresentar o seguinte Lema, que será de fundamental apoio nas proposições seguintes deste capítulo.

Lema 5.1. *Seja* $f : \mathbb{R}^n \to (-\infty, +\infty]$. *As seguintes proposições são equivalentes:*

i. f *é convexa e sci.*

ii. f *é o supremo de todas as funções afins que são menores que ou iguais* f*.*

Demonstração. Se $i.$ é verdadeiro, do Teorema 3.7, temos que f é convexa e fechada. Daí, o Teorema 3.11 garante que f é o supremo ponto a ponto de uma coleção de funções afins que são menores que ou iguais a f, concluindo que $ii.$ é verdadeiro. Se $ii.$ é verdadeiro, dizer que f é o supremo de todas as funções afins que são menores que ou iguais f é equivalente a dizer que o $epi(f)$ é a interseção de todos os semiespaços fechados definidos pelos hiperplanos que tangenciam f em $x \in dom(f)$. Portanto, da Seção 2.4, $epi(f)$ será um conjunto convexo e fechado, portanto f será uma função convexa e fechada. Mais uma vez o Teorema 3.7 garante que f é convexa e *sci*, concluindo que $i.$ é verdadeiro. $\square$

5.2 O Conjunto das Funções Convexas e Semicontínuas Inferiormente

Denotamos por $\Gamma_0(\mathbb{R}^n)$ o conjunto de todas as funções convexas e *sci* $f : \mathbb{R}^n \to (-\infty, +\infty]$, cujo o domínio seja não vazio. A Transformada de Fenchel f^* de uma função $f \in \Gamma_0(\mathbb{R}^n)$ é a função $f^* : \mathbb{R}^n \to (-\infty, +\infty]$, definida por:

$$f^*(v) = \sup_{u \in dom(f)} \{\langle v, u \rangle - f(u)\}, \tag{5.1}$$

onde

$$dom(f^*) = \{v \in \mathbb{R}^n : \sup_{u \in dom(f)} \{\langle v, u \rangle - f(u)\} < \infty\}. \tag{5.2}$$

Observação 5.1. *A Transformada de Fenchel é a conjugada da função conforme Definição 3.14, pois de acordo com o Teorema 3.7, a função convexa é fechada se, e somente se, é semicontínua inferiormente.*

Exemplo 5.1. *A função $f : \mathbb{R} \to \mathbb{R}$ definida por:*

$$f(x) = \begin{cases} -x, & se\ x \leq 0. \\ x^2, & se\ x > 0. \end{cases}$$

Observe que essa função não é de classe C^1 pois não é diferenciável em 0. Entretanto, a função f é convexa e sci. Portanto podemos obter transformada de Fenchel que é dada pela função $f^* : \mathbb{R} \to \mathbb{R}$ como

$$f^*(y) = \begin{cases} y^2/4, & se\ y \geq 0. \\ 0, & se\ -1 \leq y < 0. \\ \infty, & se\ y < -1. \end{cases}$$

Com efeito, devemos achar o supremo da função $yx - f(x)$ quando x varia na reta real. Daí, temos dois casos. Caso $x \geq 0$ implica em achar o supremo de $xy - x^2$ quando $x \in \mathbb{R}$. Observe que neste caso, o supremo é obtido no ponto crítico $y - 2x = 0$, ou seja, $x = y/2$. Daí, temos que $y \geq 0$ e o supremo é a função $yy/2 - (y/2)^2 = y^2/4$. Caso $x < 0$ implica em achar o supremo $xy + x = (1 + y)x$. Esse supremo será 0 quando $y \geq -1$ e ∞ quando $y < -1$. Portanto, combinando os dois casos, obtemos a função conjugada.

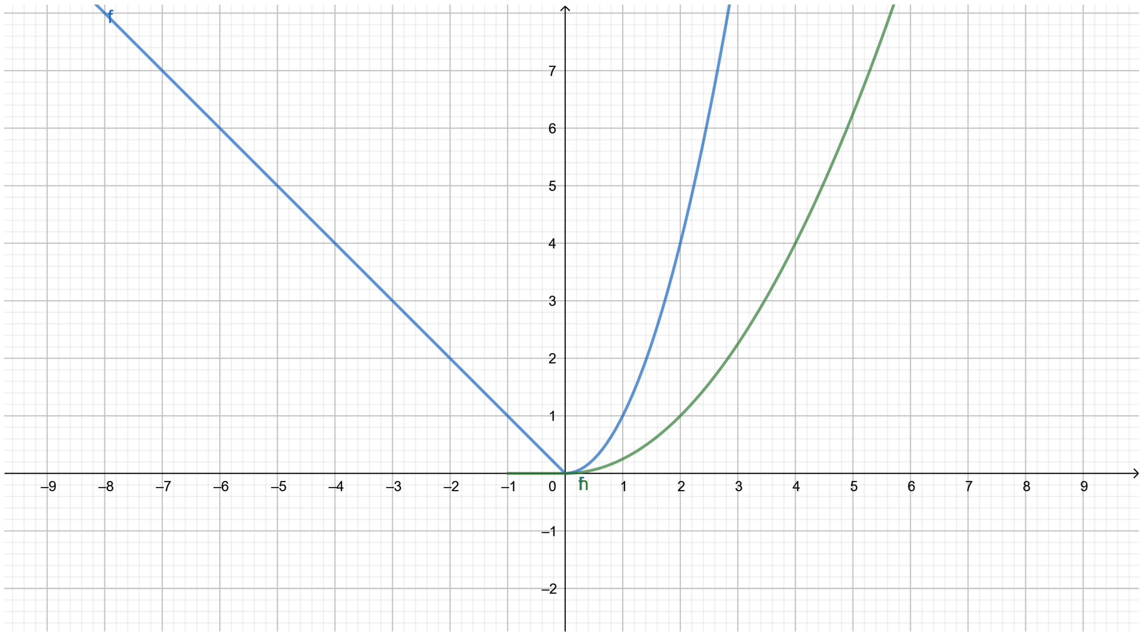

Figura 5.1: A figura mostra o gráfico da função f em azul e da sua transformada de Legendre F em verde conforme definido no Exemplo 5.1.

Exemplo 5.2. *Dado $1 < p < \infty$. Seja $f : \mathbb{R} \to \mathbb{R}$ uma função definida por*

$$f(x) = |x|^p/p.$$

Observe que essa função f é convexa e sci. Portanto podemos obter transformada de Fenchel que é dada pela função $f^* : \mathbb{R} \to \mathbb{R}$ como

$$f^*(y) = |y|^q/q,$$

onde $1/q + 1/p = 1$. Com efeito, devemos achar o supremo da função $yx - f(x)$ quando x varia na reta real. Daí, temos dois casos. Caso $x \geq 0$ implica em achar o supremo de $xy - x^p/p$. Mas veja que o supremo dessa função é atingido no ponto crítico $y - x^{p-1} = 0$. Daí, $y \geq 0$ e substituindo $x = y^{1/(p-1)}$ temos:

$$\sup_{x \geq 0}\{xy - x^p/p\} = yy^{1/(p-1)} - y^{p/(p-1)}/p = y^q - y^q/p = y^q/q,$$

pois $q = p/(p-1)$. Quando $x < 0$, temos que achar o supremo

$$\sup_{x<0}\{xy - (-x)^p/p\} = \sup_{z>0}\{-zy - (z)^p/p\}.$$

O supremo da função do lado direito da equação é atingido no ponto crítico $-y - z^{p-1} = 0$, válido quando $y < 0$. Realizando as substituições, o caso é análogo ao anterior, concluindo a obtenção da conjugada.

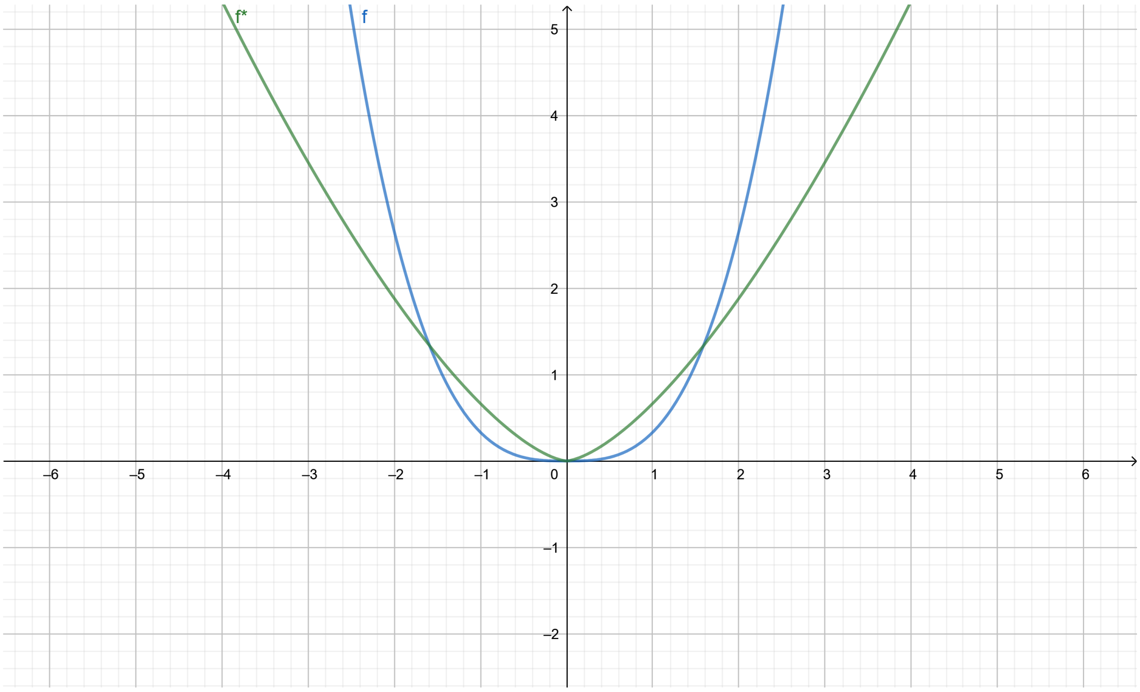

Figura 5.2: A figura mostra o gráfico da função f em azul e da sua transformada de Legendre F em verde conforme definido no Exemplo 5.1.

Observação 5.2. *A função afim e contínua* $\langle v, .\rangle - \alpha$ *é majorada por* f *se, e somente se,*

$$\alpha \geq \langle v, u\rangle - f(u),$$

para todo $u \in \mathbb{R}^n$*, ou seja, se, e somente se,* $\alpha \geq f^*(v)$*. Mas veja que, conforme Lema 3.1., a função* f^* *é convexa e sci. Desde que* $\alpha \geq f^*(v)$*, o* $\mathrm{dom} f^* \neq \emptyset$ *e, portanto,* $f^* \in \Gamma_0(\mathbb{R}^n)$*. Uma consequência imediata da transformada de Fenchel é a seguinte desigualdade:*

$$f(u) + f^*(v) \geq \langle u, v\rangle,$$

para todos $u, v \in \mathbb{R}^n$*. Uma outra consequência imediata é que dados* $f_1, f_2 \in \Gamma_0(\mathbb{R}^n)$ *tais que* $f_1 \leq f_2$*, então* $f_2^* \geq f_1^*$*.*

5.3 O Teorema de Fenchel-Moreau

Teorema 5.1 (Fenchel-Moreau)**.** *Se* $f \in \Gamma_0(\mathbb{R}^n)$*, então* $f^{**} = f$*.*

Demonstração. Primeiramente, observe que, por definição

$$f^{**}(u) = \sup_v \{\langle u, v\rangle - f^*(v)\} \leq f(u), \tag{5.3}$$

pois $\langle u,v\rangle - f^*(v) \leq f(u)$. Para cada $u \in \mathbb{R}^n$, $f(u) = \sup_{(v,\alpha)\in\mathbb{R}^n\times\mathbb{R}}\{\langle v,u\rangle - \alpha\}$, sempre que $f \geq \langle v,.\rangle - \alpha$. Mas,

$$\sup_{(v,\alpha)\in\mathbb{R}^n\times\mathbb{R}}\{\langle v,u\rangle - \alpha\} = \tag{5.4}$$

$$\sup_{v\in\mathbb{R}^n \text{ e } \alpha\geq f^*(v)}\{\langle v,u\rangle - \alpha\} \leq \tag{5.5}$$

$$\sup_{v\in dom(f^*)}\{\langle v,u\rangle - f^*(v)\} = f^{**}(u). \tag{5.6}$$

Das últimas 3 equações, tem-se

$$f(u) \leq f^{**}(u). \tag{5.7}$$

Portanto, de 5.3 e 5.7, $f(u) = f^{**}(u)$. □

5.4 Subdiferencial de uma Função

Definimos o subdiferencial da função $f \in \Gamma_0(\mathbb{R}^n)$ no ponto $u \in \mathbb{R}^n$ como sendo o conjunto:

$$\partial f(u) = \{v \in \mathbb{R}^n \colon f(w) \geq f(u) + \langle v, w-u\rangle \text{ para todo } w \in \mathbb{R}^n\}.$$

Dizemos que f é subdiferenciável em u se $\partial f(u) \neq \emptyset$.

Exemplo 5.3. *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função linear, ou seja:*

$$f(x) = ax + b,\ a, b \in \mathbb{R}.$$

Observe que f é convexa e sci e o domínio é toda reta real. Dado $c \in \mathbb{R}$ o conjunto:

$$\partial f(c) = \{x \in \mathbb{R} \colon f(y) \geq f(c) + x(y-c),\ \textit{para todo}\ y \in \mathbb{R}\} = (-\infty, a].$$

Exemplo 5.4. *Seja $f : \mathbb{R} \to \mathbb{R}$ a função modular $f(x) = |x|$. Observe que f é convexa e sci e o domínio é todo $\mathbb{R}$. Então o subdiferencial em 0 é o conjunto*

$$\partial f(0) = \{x \in \mathbb{R} \colon f(y) \geq f(0) + xy,\ \textit{para todo}\ y \in \mathbb{R}\} = [-1, 1].$$

Exemplo 5.5. *Seja $f : \mathbb{R} \to \mathbb{R}$ a função modular $f(x) = x^2$. Observe que f é convexa e sci e o domínio é todo $\mathbb{R}$. Então o subdiferencial em 1 é o conjunto*

$$\partial f(1) = \{x \in \mathbb{R} \colon f(y) \geq f(1) + x(y-1),\ \textit{para todo}\ y \in \mathbb{R}\}.$$

Observe que deveríamos ter $y^2 - xy + x - 1 \geq 0$, para qualquer y, então o $\Delta \leq 0 \Rightarrow x^2 - 4x + 4 = (x-2)^2 \leq 0$, em que a única solução é $x = 2$. Portanto $\partial f(1) = \{2\}$.

Observação 5.3. *Segue algumas observações com respeito à subdiferenciabilidade:*

i. *f é subdiferenciável em u, se existe uma função afim h, tal que $h(u) = f(u)$ e $f(v) \geq h(v)$, para todo $v \in \mathbb{R}^n$. Com efeito, seja $h : \mathbb{R}^n \to \mathbb{R}$ uma função afim definida como:*

$$h(x) = \langle a, x \rangle + b,$$

onde $a \in \mathbb{R}^n$ e $b = f(u) - \langle a, u \rangle$. Se $f(v) \geq h(v)$ para todo $v \in \mathbb{R}^n$, então:

$$f(v) \geq \langle a, v \rangle + f(u) - \langle a, u \rangle = f(u) + \langle a, v - u \rangle.$$

Assim, $a \in \partial f(u)$, logo f é subdiferenciável em u.

ii. *Se $u \in \mathbb{R}^n$ é tal que $\inf_{x \in \mathbb{R}^n} f(x) = f(u)$, então $0 \in \partial f(u)$. A recíproca também é verdadeira. Com efeito, veja que se $\inf_{x \in \mathbb{R}^n} f(x) = f(u)$, então $f(v) \geq f(u)$, para todo $v \in \mathbb{R}^n$. Logo*

$$f(v) \geq f(u) + \langle 0, v - u \rangle,$$

para todo $v \in \mathbb{R}^n$. Portanto, $0 \in \partial f(u)$. Por outro lado, se $0 \in \partial f(u)$, então $f(v) \geq f(u) + \langle 0, v - u \rangle$, para todo $v \in \mathbb{R}^n$. Logo $f(u) = \inf_{x \in \mathbb{R}^n} f(x)$.

iii. *Se $v_i \in \partial f(u_i)$ para $i = 1, 2$, então:*

$$f(u_2) \geq f(u_1) + \langle v_1, u_2 - u_1 \rangle.$$

$$f(u_1) \geq f(u_2) + \langle v_2, u_1 - u_2 \rangle.$$

Daí,

$$\langle v_2 - v_1, u_1 - u_2 \rangle \leq 0 \Rightarrow \langle v_2 - v_1, u_2 - u_1 \rangle \geq 0.$$

iv. *$\partial f(u)$ é fechado e convexo. De fato, sejam $a, b \in \partial f(u)$ e $\lambda \in [0, 1]$ então:*

$$\begin{gathered} f(v) \geq f(u) + \langle a, v - u \rangle \text{ e} \\ f(v) \geq f(u) + \langle b, v - u \rangle, \end{gathered}$$

para todo $v \in \mathbb{R}^n$. Multiplicando a primeira inequação por $1 - \lambda$ e a segunda por λ e somando ambas, obtemos:

$$f(v) \geq f(u) + \langle (1 - \lambda)a, v - u \rangle + \langle \lambda b, v - u \rangle = f(u) + \langle (1 - \lambda)a + \lambda b, v - u \rangle.$$

Portanto, $(1 - \lambda)a + \lambda b \in \partial f(u)$. Considere uma sequência $(x_k)_{k \in \mathbb{N}}$ do $\partial f(u)$ que converge para $x \in \mathbb{R}^n$. Então para todo $v \in \mathbb{R}^n$,

$$f(v) \geq f(u) + \langle x_m, v-u \rangle,$$

para todo $m \in \mathbb{N}$. Afirmamos que:

$$\lim_{m\to\infty} \langle x_m, v-u \rangle = \langle x, v-u \rangle.$$

Com efeito, dado $\epsilon > 0$, existe um $m_0 \in \mathbb{N}$, tal que para todo $m \geq m_0$

$$\|x_m - x\| < \epsilon / \|v-u\|.$$

para $v \neq u$. Pela desigualdade de Schwarz,

$$|\langle x_m, v-u \rangle - \langle x, v-u \rangle| = |\langle x_m - x, v-u \rangle| \leq \|x_m - x\| \, \|v-u\| < \epsilon.$$

Portanto, concluímos que qualquer que seja $v \in \mathbb{R}^n$,

$$f(v) \geq f(u) + \langle x, v-u \rangle.$$

Logo $x \in \partial f(u)$, concluindo que $\partial f(u)$ é fechado.

Teorema 5.2. *Se $f \in \Gamma_0(\mathbb{R}^n)$, então as seguintes afirmações são equivalentes:*

i. $v \in \partial f(u)$.

ii. $f(u) + f^*(v) = \langle v, u \rangle$.

iii. $u \in \partial f^*(v)$.

Demonstração. Temos que se $v \in \partial f(u)$, então

$$\langle v, u \rangle - f(u) \geq \langle v, w \rangle - f(w),$$

para todo $w \in \mathbb{R}^n$. Então o supremo da função do lado direito é obtido em $w = u$, logo $\langle v, u \rangle - f(u) = f^*(v)$. Se *ii.* é verdadeiro, então escrevemos

$$\langle v, u \rangle - f^*(v) = f(u).$$

Pelo Teorema 5.1, temos que

$$f(u) = f^{**}(u) = \sup_{w\in\mathbb{R}^n} \{\langle w, u \rangle - f^*(w)\}.$$

Logo, para todo $w \in \mathbb{R}^n$,

$$\langle v, u \rangle - f^*(v) \geq \langle w, u \rangle - f^*(w),$$

concluindo que $u \in \partial f^*(v)$. Por fim, se *iii.* é verdadeiro, então

$$\langle v, u \rangle - f^*(v) \geq \langle w, u \rangle - f^*(w),$$

para todo $w \in \mathbb{R}^n$. Logo o supremo da função do lado direito é obtido quando $w = v$, mas esse supremo é $f^{**}(u) = f(u)$ pelo Teorema 5.1. Então,

$$\langle v, u\rangle - f^*(v) = f(u) \Rightarrow \langle v, u\rangle - f^{**}(v) = f^*(u) = \sup_{w\in\mathbb{R}^n}\{\langle w, u\rangle - f(w)\}.$$

Daí,

$$\langle v, u\rangle - f(v) \geq \langle w, u\rangle - f(w).$$

Portanto $v \in \partial f(u)$. □

Proposição 5.1. *Considere uma função $f \in \Gamma_0(\mathbb{R}^n)$. Então o conjunto*

$$F = \{(u, v) \in \mathbb{R}^n \times \mathbb{R}^n : v \in \partial f(u)\}$$

é fechado.

Demonstração. Seja $(z_k)_{k\in\mathbb{N}}$ uma sequência de F que converge para $z \in \mathbb{R}^{2n}$. Represente $z_m = (u_m, v_m)$ e $z = (u, v)$. Para todo $w \in \mathbb{R}^n$ temos

$$f(w) \geq f(u_m) + \langle v_m, w - u_m\rangle,$$

para todo $m \in \mathbb{N}$. Como f é *sci*, $f(u) \leq \lim_{m\in\mathbb{N}} f(u_m)$. Logo,

$$f(w) \geq f(u) - \langle v, w - u\rangle.$$

Daí, $v \in \partial f(u)$ e $z = (u, v) \in F$. Portanto, F é fechado. □

Observação 5.4. *Observe que o conjunto F da Proposição 5.1 representa o gráfico de $\partial f(u)$.*

5.5 Funções Diferenciáveis e Caracterização

O teorema a seguir ilustra a propriedade de diferenciabilidade para as transformadas de Fenchel.

Teorema 5.3. *Se a função convexa $f : \mathbb{R}^n \to \mathbb{R}$ é diferenciável em $u \in \mathbb{R}^n$, então*

$$\partial f(u) = \{\nabla f(u)\}.$$

Demonstração. Note que pela definição, $\nabla f(u) \in \partial f(u)$, pois para todo $w \in \mathbb{R}^n$, tem-se

$$f(w) - f(u) = f(u + w - u) - f(u) \geq \langle \nabla f(u), w - u\rangle$$

pela propriedade de funções convexas. Agora, considere $v \in \partial f(u)$, então, pela definição, tem-se

$$f(w) - \langle v, w\rangle \geq f(u) - \langle v, u\rangle$$

para todo $w \in \mathbb{R}^n$, ou seja a função $f(.) - \langle v, . \rangle$ tem um mínimo em u. Pela diferenciabilidade de f em u, conclui-se: $\nabla f(u) - v = 0$, demonstrando o teorema. $\square$

O teorema a seguir caracteriza a diferenciabilidade da conjugada.

Teorema 5.4. *Seja $f \in \Gamma_0(\mathbb{R}^n)$ uma função estritamente convexa que satisfaz:*

$$\lim_{\|x\| \to \infty} f(x)/\|x\| = +\infty.$$

Então $f^* \in C^1(\mathbb{R}^n, \mathbb{R})$.

Demonstração. Dado $v \in \mathbb{R}^n$ defina a função $g_v : \mathbb{R}^n \to \mathbb{R}$, como sendo:

$$g_v(w) = \langle v, w \rangle - f(w).$$

Note que a função g_v está bem definida pois $f \in \Gamma_0(\mathbb{R}^n)$ possui domínio não vazio. Adicionalmente g_v é estritamente côncava. Ainda, temos que:

$$\lim_{\|w\| \to \infty} g_v(w) = \lim_{\|w\| \to \infty} \|w\| \left((\langle v, w \rangle) / \|w\| - f(w)/\|w\| \right) = -\infty.$$

Como g_v é estritamente côncava, então ela possui exatamente um máximo global, digamos em $u \in \mathbb{R}^n$. Afirmamos que $\partial f^*(v) = \{u\}$. Com efeito, o máximo de g_v ocorre em

$$\sup_{w \in \mathbb{R}^n} \{\langle v, w \rangle - f(w)\} = \langle v, u \rangle - f(u).$$

Logo para todo $w \in \mathbb{R}^n$,

$$\langle v, u \rangle - f(u) \geq \langle v, w \rangle - f(w).$$

Daí, $v \in \partial f(u)$. Do Teorema 5.2, $u \in \partial f^*(v)$. Suponha que exista $z \neq u$ no $\mathbb{R}^n$, tal que $z \in \partial f^*(v)$. Do Teorema 5.2, $v \in \partial f(z)$. Então, escolhendo $w = u$, teríamos

$$\langle v, z \rangle - f(z) \geq \langle v, u \rangle - f(u).$$

Logo $g_v(z) \geq g_v(u)$, absurdo pois u é o único máximo global de g_v. Agora, vamos mostrar que $\partial f^* : \mathbb{R}^n \to \mathbb{R}^n$ é um mapeamento contínuo. Inicialmente, mostraremos que ∂f^* é limitado. Com efeito, dados $a \in dom(f)$ e seja $v \in \mathbb{R}^n$ tal que $\partial f^*(v) = \{u\}$. Do Teorema 5.2 $\partial f(u) = \{v\}$. Logo,

$$f(a) \geq f(u) + \langle v, a - u \rangle = \|u\| \left(f(u)/\|u\| + \langle v, a - u \rangle / \|u\| \right).$$

Veja que se $\|u\|$ não fosse limitada o lado direito da inequação tenderia para infinito o que é um absurdo. Seja $(z_k)_{k\in\mathbb{N}}$ uma sequência em $\mathbb{R}^n$ que converge para z, tal que $\partial f^*(z_m) = \{y_m\}$ para cada $m \in \mathbb{N}$. Observe que (y_k) é uma sequência de um conjunto limitado. Pelo Teorema de Weierstrass $y_k \to y$, para uma subsequência $k \in \mathbb{N}_1 \subset \mathbb{N}$. Portanto, $(z_m, y_m) \to (z, y)$ e da Proposição 5.1 $\partial f^*(z) = \{y\}$, demonstrando a continuidade.

Por fim, seja $\partial f^*(v) = \{u\}$. Dado $h \in \mathbb{R}^n$, não nulo, tal que $v + h \in dom(f^*)$ e $\partial f^*(v + h) = \{u_h\}$. Da definição de subdiferencial temos:

$$f^*(v) \geq f^*(v+h) + \langle u_h, -h\rangle.$$

Logo,

$$(f^*(v+h) - f^*(v) - \langle u, h\rangle)/\,\|h\| \leq (\langle u_h, h\rangle - \langle u, h\rangle)/\,\|h\| \leq \|u_h - u\|\,.$$

Da continuidade de ∂f^*, temos que $\lim_{h\to 0} u_h = u$. Daí, concluímos que f^* é diferenciável em v. Do Teorema 5.3, $u = \nabla f^*(v)$. Portanto, f^* é de classe C^1. □

5.6 Exemplos de Transformadas de Fenchel

Nesta seção, exploraremos alguns exemplos de funções que não são de classe C^1, mas que podemos obter a Transformada de Fenchel.

Exemplo 5.6. *A função módulo definida por $f : \mathbb{R} \to \mathbb{R}$ com $f(x) = |x|$. Para encontrar a transformada de Fenchel, devemos encontrar o supremo da função $xy - f(x)$ quando x varia na reta real. Vamos considerar dois casos. Caso $x \geq 0$, então devemos encontrar o supremo de $xy - x = x(y-1)$ quando x varia na reta. Esse supremo será $+\infty$ se $y > 1$ e 0 caso contrário. Se $x < 0$, então devemos encontrar o supremo de $xy + x = x(y+1)$, que será $+\infty$ se $y > -1$ e 0 caso contrário. Considerando ambos casos, temos que a transformada de Fenchel é $f^* : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ será $f^*(y) = +\infty$ se $y > -1$ e 0 caso contrário.*

Exemplo 5.7. *Considere função $f : \mathbb{R}^n \to \mathbb{R}$ definida na norma da soma como $f(x) = a\,\|x\|_1$, em que $a \in \mathbb{R}$. Para encontrar a transformada de Fenchel, devemos encontrar o supremo da função $\langle x, y\rangle - a\,\|x\|_1 = \sum_{i=1}^n x_i y_i - sgn(x_i) a x_i = \sum_{i=1}^n x_i(y_i - sgn(x_i)a)$, quando $x = (x_1, x_2, ..., x_n)$ varia em $\mathbb{R}^n$. Então, veja que o supremo da função será 0 se $|y_i| \leq a$ para todo $i \in \{1, 2, ..., n\}$ e $+\infty$ caso contrário. Logo, a transformada de Fenchel é $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ será $f^*(y) = +\infty$ se $|y_i| > a$ para algum $1 \leq i \leq n$ e 0 caso contrário.*

Por fim, apresentamos o seguinte exemplo de aplicação da transformada de Fenchel para estabelecer existência de solução de um sistema Hamiltoniano semilinear em determinadas condições, resultado obtido por [2].

Exemplo 5.8. *Considere o seguinte sistema Hamiltoniano semilinear:*

$$\begin{cases} -\Delta u + u = W_2(x)|v|^{p-1}v & \text{em } \mathbb{R}^N, \\ -\Delta v + v = W_1(x)|u|^{q-1}u & \text{em } \mathbb{R}^N, \\ u(x), v(x) \to 0 \text{ quando} & |x| \to \infty, \\ u, v > 0 & \text{em } \mathbb{R}^N. \end{cases}.$$

onde $p, q > 1$ são valores que possuem a seguinte restrição para $N \geq 3$:

$$1/(p+1) + 1/(q+1) > (N-2)/N.$$

Sobre certas condições das funções contínuas, limitadas e positivas W_1, W_2 os autores provam a existência de soluções para o caso periódico, utilizando na obtenção da solução a Transformada de Fenchel. Para maiores detalhes ver [2].

Capítulo 6

Elementos de Análise Convexa em Espaços de Dimensão Infinita

Podemos expandir alguns conceitos da análise convexa, apresentado nesse trabalho para espaços vetoriais de dimensão infinita. Para tal, precisamos introduzir alguns conceitos sobre topologia e espaços métricos. Os conceitos, as definições e os teoremas desse capítulo foram consultados nas referências [8, 11, 14].

6.1 Definições em Espaços Topológicos

Definição 6.1. *Um espaço topológico é um conjunto X e uma coleção τ de subconjuntos de X denominados conjuntos abertos, tais que:*

T1. $\emptyset \in \tau$ *e* $X \in \tau$.

T2. Se U_1 *e* $U_2 \in \tau$, *então* $U_1 \cap U_2 \in \tau$.

T3. A união de qualquer coleção de conjuntos abertos em τ é aberto.

Denotamos por (X, τ) o espaço topológico. Neste trabalho, escreveremos apenas X quando não houver perigo de confusão.

Exemplo 6.1. *Dizemos que a topologia padrão na reta real é o próprio conjunto dos reais, $X = \mathbb{R}$ e τ como sendo a coleção de todos os subconjuntos abertos da reta.*

Observação 6.1. *Dizemos que τ é uma topologia trivial de X se $\tau = \{\emptyset, X\}$. Dizemos que τ é uma topologia discreta, se $\tau = \{O\colon O \subset X\}$, ou seja, τ consiste de todos os subconjuntos de X.*

Definição 6.2. *Dado um espaço topológico X e um subconjunto $F \subset X$. Dizemos que F é fechado, se seu complementar $X \setminus F$ for aberto.*

Exemplo 6.2. *O conjunto $[0, 1] \subset \mathbb{R}$ é fechado pois seu complementar $(-\infty, 0) \cap (1, \infty)$ é um conjunto aberto da reta real.*

Definição 6.3. *Dado um espaço topológico X, dizemos que V é uma vizinhança aberta de $u \in X$, se $V \subset X$ é aberto e $u \in V$. Similarmente, V é uma vizinhança aberta de um subconjunto $S \subset X$, se $V \subset X$ é aberto e $S \subset V$. Uma vizinhança de um ponto ou conjunto é um conjunto que contém uma vizinhança aberta do ponto ou conjunto.*

Exemplo 6.3. *Dado $x \in \mathbb{R}$. Os conjuntos $V_1 = (x-2, x+3]$ e $V_2 = (x-1, x+1)$ são exemplos de vizinhanças de x. Apenas V_2 é aberto. Na topologia trivial, apenas conforme Observação 6.1, existe apenas uma vizinhança aberta para qualquer elemento de X, sendo o próprio X. Na topologia discreta, dado $x \in X$ qualquer subconjunto contendo x é uma vizinhança aberta, pois $\{x\}$ é um conjunto aberto.*

Definição 6.4. *Um espaço topológico é dito Hausdorff, se quaisquer dois elementos possuem vizinhanças disjuntas. O espaço topológico é regular se é Hausdorff e cada elemento e cada subconjunto fechado que não contém o elemento possuem vizinhanças disjuntas. O espaço é dito normal se é Hausdorff e dois subconjuntos fechados disjuntos possuem vizinhanças disjuntas.*

Observação 6.2. *A maior parte dos espaços que estudamos em geometria e análise são normais. A topologia discreta de qualquer topologia é normal. A topologia trivial não é nem Hausdorff, se o espaço possui mais de um elemento.*

Exemplo 6.4. *Considere $X = \{a, b, c, d\}$ e $\tau = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{a, b, c\}, X\}$. Veja que a única vizinhança aberta de $d \in X$ é o próprio X. As vizinhanças abertas de $a \in X$ são os conjuntos $\{a\}, \{a, b\}, \{a, b, c\}, X$. Logo os elementos d e a não possuem vizinhanças abertas disjuntas de X. Portanto a topologia (X, τ) não é Hausdorff.*

Exemplo 6.5. *Considere o espaço topológico $(\mathbb{R}, \tau)$ em que*

$$\tau = \{(-n, n) \colon n \in \mathbb{Z}, n \geq 1\}.$$

Afirmamos que $(\mathbb{R}, \tau)$ não é Hausdorff. Com efeito, dados $0, 1/2 \in \mathbb{R}$. Observe que

$$0, 1/2 \in (-1, 1) \subset (-2, 2).....$$

Logo dados U e V vizinhanças quaisquer de 0 e $1/2$, $U \cap V \neq \emptyset$.

Definição 6.5. *Um espaço topológico X é enumerável de primeira ordem se para cada $u \in X$ existe uma sequência de vizinhanças de u, $\{V_1, V_2, ...\} = \{V_n\}$, tal que para toda vizinhança V de u, existe um $k \in \mathbb{N}$, tal que $V_k \subset V$. Dizemos que B é uma base da topologia, se cada subconjunto aberto é uma união de elementos de B. Uma topologia é enumerável de segunda ordem, se contém uma base enumerável.*

Observação 6.3. *Grande parte dos espaços topológicos a serem estudados são enumeráveis de primeira e segunda ordem. Entretanto, há alguns espaços que não são enumeráveis de primeira ordem, como será mostrado nos exemplos a seguir.*

Exemplo 6.6. *A topologia padrão em $\mathbb{R}$ é enumerável de primeira ordem. Com efeito para qualquer vizinhança aberta de $x \in \mathbb{R}$, $(x - \epsilon, x + \epsilon)$ a sequência (V_k) de vizinhanças de x definidas por*

$$V_n = (x - 1/n, x + 1/n),\ n \in \mathbb{N},$$

satisfaz as condições da Definição 6.5.

Definição 6.6. *O fecho de um conjunto A da topologia X é definido como a interseção de todos os subconjuntos fechados da topologia X que contém A, denotamos por $cl(A)$. Similarmente, o interior de um conjunto A é a reunião de todos os subconjuntos abertos contidos em A, denotamos por $int(A)$. A fronteira de A é definida por:*

$$bd(A) = cl(A) \cap cl(X \setminus A).$$

Exemplo 6.7. *Considere o conjunto $\mathbb{R}$ e sua topologia padrão. Daí temos:*

$$\begin{aligned} cl((0,1]) &= [0,1], \\ int((0,1]) &= (0,1), \\ bd((0,1]) &= \{0,1\}. \end{aligned}$$

Definição 6.7. *Um subconjunto $A \in X$ é denso em X se $cl(A) = X$. O espaço X é separável se contém um subconjunto enumerável denso. Um elemento $u \in A$ é de acumulação se toda vizinhança de u em A contém um outro elemento diferente de u. O conjunto de pontos de acumulação é chamado de conjunto derivado de A, denotado por $der(A)$. Por outro lado, u é isolado se existe uma vizinhança de u que não contém nenhum ponto de A, exceto por u.*

Exemplo 6.8. *O conjunto dos números racionais $\mathbb{Q}$ é denso em $\mathbb{R}$ pois cada número $\mathbb{R}$ é limite de uma sequência de números $\mathbb{Q}$, ou seja $cl(\mathbb{Q}) = \mathbb{R}$. Como $\mathbb{Q} \subset \mathbb{R}$, temos que $\mathbb{R}$ é separável. Considere conjunto $A = (0,1] \cup \{2\}$. Dizemos que 2 é um ponto isolado de A. Ainda observe que $der(A) = [0,1]$.*

6.2 Convergência e Sequências de Cauchy

Definição 6.8. *Sejam um espaço topológico X e uma sequência $\{u_n\}$ de elementos de X. Dizemos que essa sequência converge para $u \in X$, se para toda vizinhança V de u em X, existe um $N \in \mathbb{N}$, tal que para todo $n \geq N$, tem-se $u_n \in V$.*

Exemplo 6.9. *Considere a sequência* (x_k) *dos* $\mathbb{R}$ *definida por*

$$x_n = 1/n, \text{ para } n \in \mathbb{N}.$$

Dizemos que a sequência (x_k) *converge para* $0 \in \mathbb{R}$. *Com efeito, para todo* $\epsilon > 0$ *existe um* $N \in \mathbb{N}$, *tal que* $N > 1/\epsilon$ *e para todo* $n \geq N$ *temos*

$$x_n \in (-\epsilon, \epsilon).$$

Observação 6.4. *Se* X *é um espaço topológico enumerável de primeira ordem, então um elemento pertence ao seu fecho se, e somente se, é o limite de uma sequência de elementos de* X.

Definição 6.9. *Considere* X *um conjunto. Uma métrica em* X *é uma função* $d : X \times X \to \mathbb{R}$, *tais que para todos* $m_1, m_2, m_3 \in X$, *tem-se*

M1. $d(m_1, m_2) = 0$ *se, e somente se,* $m_1 = m_2$.

M2. $d(m_1, m_2) = d(m_2, m_1)$.

M3. $d(m_1, m_3) \leq d(m_1, m_2) + d(m_2, m_3)$.

Observação 6.5. *O espaço métrico é denotado como o par* (X, d).

Exemplo 6.10. *A função distância definida por* $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, *tal que*

$$d(x, y) = \|x - y\|,$$

onde $\|.\|$ *é a norma euclidiana, é uma métrica.*

Definição 6.10. *Num espaço métrico definimos uma bola aberta de centro* $m \in X$ *de raio* $\epsilon > 0$, *como sendo:*

$$B\,(m; \epsilon) = \{m' \in X \colon d(m', m) < \epsilon\}.$$

A bola fechada é definida por:

$$B[m; \epsilon] = \{m' \in X \colon d(m', m) \leq \epsilon\}.$$

Exemplo 6.11. *No* $\mathbb{R}^2$ *a bola fechada unitária definida na norma euclidiana é o conjunto:*

$$B[0, 1] = \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 \leq 1\}.$$

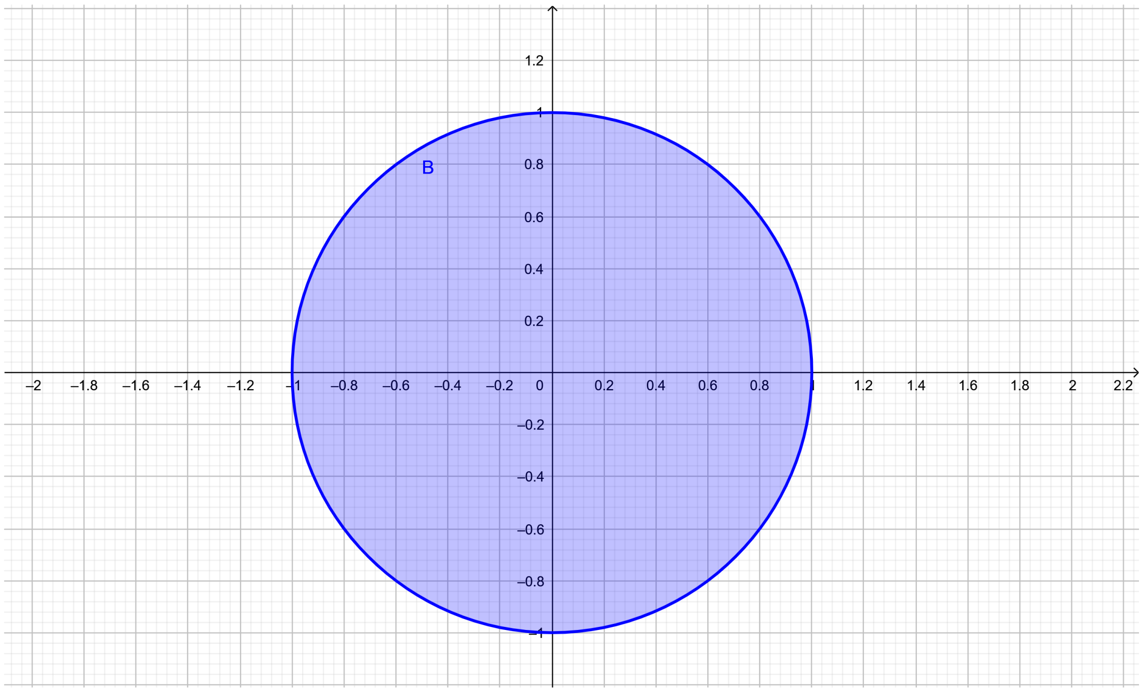

Figura 6.1: A figura representa a bola unitária no $\mathbb{R}^2$ quando utilizamos a métrica da norma euclidiana.

Definição 6.11. *A coleção de subconjuntos de X que são a união de bolas abertas define uma topologia métrica no espaço métrico (X, d). Dois espaços métricos são equivalentes se elas induzem a mesma topologia.*

Exemplo 6.12. *Os espaços métricos $(\mathbb{R}^2, \|.\|_1)$ e $(\mathbb{R}^2, \|.\|_2)$, com as normas expressas conforme Observação 1.1, são equivalentes.*

Observação 6.6. *Uma pseudométrica é definida pelas propriedades M2 e M3, porém dois elementos distintos podem ter distância nula.*

Exemplo 6.13. *Considere os espaço $\mathbb{R}^2$ e a função $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por:*

$$d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2|.$$

é uma pseudo-métrica. Com efeito, os elementos $(1, 0)$ e $(1, 1)$ são distintos porém

$$d((1, 0), (1, 1)) = |1 - 1| = 0.$$

Definição 6.12. *Seja X um espaço métrico com métrica d e seja $\{u_n\}$ uma sequência de X. Então a sequência $\{u_n\}$ é de Cauchy se para todo real $\epsilon > 0$, existe um inteiro N, tal que, para $n, m \geq N$ tem-se $d(u_n, u_m) < \epsilon$. O espaço é completo se toda sequência de Cauchy converge.*

Exemplo 6.14. *O $\mathbb{R}$ com sua topologia padrão é um espaço completo.*

6.3 Propriedades de Espaços Métricos

Definição 6.13. *Seja X um espaço métrico completo e $f : X \to X$ um mapeamento. Dizemos que $u \in X$ é um ponto fixo se $f(u) = u$.*

Exemplo 6.15. *A função $f : \mathbb{R} \to \mathbb{R}$ definida por*

$$f(x) = 4x(1 - x),$$

possui os pontos fixos 0 *e* $3/4$.

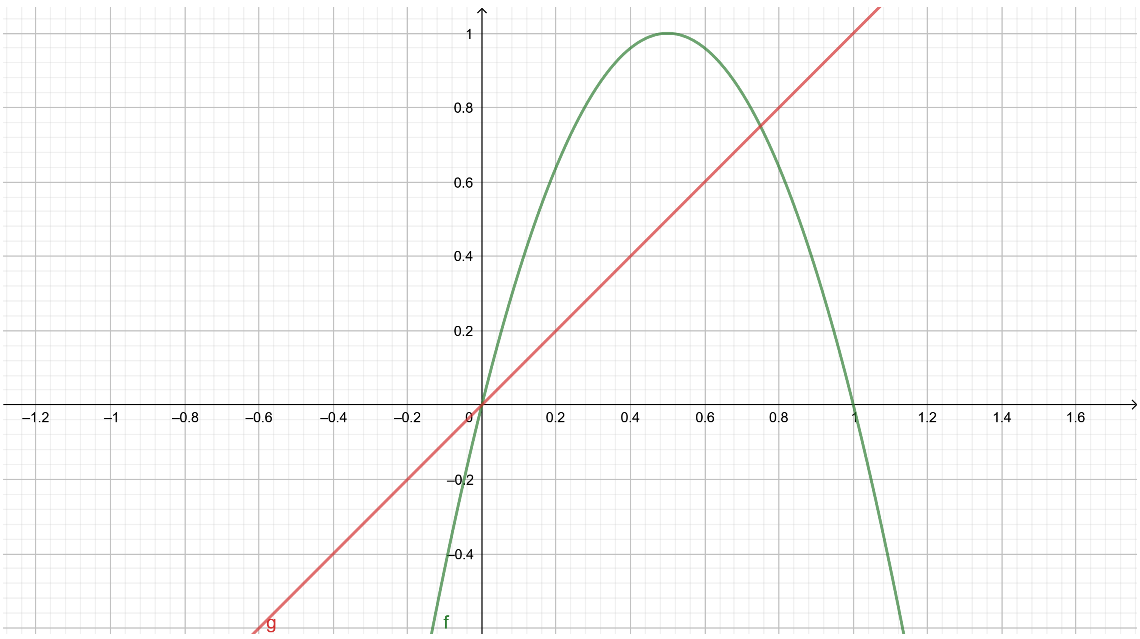

Figura 6.2: A figura representa os gráficos das funções $f(x)$, $g(x) = x$ e suas interseções que são os pontos fixos conforme Exemplo 6.15.

Teorema 6.1 (Teorema do Ponto Fixo)**.** *Seja X um espaço métrico completo e $f : X \to X$ um mapeamento. Suponha que exista uma constante $0 \leq k < 1$, tal que:*

$$d(f(m), f(n)) \leq kd(m, n)$$

para todos $m, n \in X$. Então f possui um único ponto fixo.

Demonstração. Seja $m_0 \in X$ um elemento qualquer. Defina a sequência $\{m_k\}_{k \in \mathbb{N}}$, pondo $m_{k+1} = f(m_k)$, para todo $k \geq 0$. Observe que:

$$d(m_{i+1}, m_i) = d(f(m_i), f(m_{i-1})) \leq kd(m_i, m_{i-1}).$$

Por indução verifica-se que

$$d(m_{i+1}, m_i) \leq k^i d(m_1, m_0).$$

Logo para quaisquer $i, j \in \mathbb{N}$ com $i < j$, tem-se

$$d(m_i, m_j) \leq (k^i + ... + k^{j-1}) d(m_1, m_0),$$

que tende a 0 quando $i, j \to +\infty$. Portanto, $\{m_k\}$ é uma sequência de Cauchy e, pela completeza de X, converge para um elemento m^*. Portanto, temos que:

$$d(m^*, f(m^*)) \leq d(m^*, m_i) + d(m_i, f(m_i)) + d(f(m_i), f(m^*)).$$

Mas,

$$d(m^*, m_i) + d(m_i, f(m_i)) + d(f(m_i), f(m^*)) \leq (1 + k) d(m^*, m_i) + k^i d(m_1, m_0).$$

Veja que o lado direito da equação tende a 0 quando $i \to \infty$, portanto $f(m^*) = m^*$. Para demonstrar a unicidade, suponha que exista $n \neq m^*$ em X, tal que $f(n) = n$. Então, temos

$$d(m^*, n) \leq d(m^*, f(m^*)) + d(f(m^*), f(n)) + d(f(n), n) = d(f(m^*), f(n)) = kd(m^*, n),$$

absurdo pois $0 \leq k < 1$. □

Exemplo 6.16. *Considere a métrica módulo $|.|$ em $\mathbb{R}$. Seja $f : \mathbb{R} \to \mathbb{R}$ uma função tal que:*

$$f(x) = x/2 + 1.$$

Logo f possui um único ponto fixo em $x = 2$. Com efeito, dados $x, y \in \mathbb{R}$ temos:

$$|f(x) - f(y)| = |x/2 - y/2| = 1/2|x - y|.$$

Daí, como $\mathbb{R}$ é completo, o Teorema 6.1 garante que f possui um único ponto fixo.

6.4 Continuidade

Definição 6.14. *Sejam X e Y espaços topológicos e um mapeamento $f : X \to Y$. Diz-se que f é contínuo em $u \in X$ se para toda vizinhança V de $f(u)$ existe uma vizinhança U de u, tal que $f(U) \subset V$.*

Exemplo 6.17. *Considere a função $f : \mathbb{R} \to \mathbb{R}$ definida por:*

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

Sabemos dos cursos fundamentais de análise que f é contínua. Veja que a imagem inversa $f^{-1}((0, 1)) = (-1, 0) \cup (0, 1)$ é aberto.

Proposição 6.1. *Sejam X e Y espaços topológicos e um mapeamento $f : X \to Y$. Se para todo conjunto aberto $S \subset Y$ a imagem inversa $f^{-1}(S) = \{u \in X : f(u) \in S\}$ é aberta em X, então f é contínua. Similarmente, se a imagem inversa de todo subconjunto fechado for fechada, então o mapeamento é contínuo.*

Demonstração. Ver [14], Seção 1.3, Proposição 1.3.2. □

Definição 6.15. *Se f é uma bijeção contínua e f^{-1} é contínuo, então f é chamado de homeomorfismo.*

Proposição 6.2. *Sejam X e Y espaços topológicos enumeráveis de primeira ordem e um mapeamento $f : X \to Y$. f é contínua se, e somente se, toda sequência $\{u_n\}$ converge para $u \in X$ a sequência $\{f(u_n)\}$ converge para $f(u)$, para todo $u \in X$.*

Demonstração. Ver [14], Seção 1.3, Corolário 1.3.3. □

Quando os espaços topológicos X e Y do mapeamento $f : X \to Y$ são espaços métricos, podemos definir continuidade, utilizando os conceitos aprendidos em análise real.

Definição 6.16. *Dados os espaços métricos (X, d_1) e (Y, d_2) e um mapeamento $f : X \to Y$. Dizemos que f é uniformemente contínua se para todo $\epsilon > 0$, existe um $\delta > 0$, tal que se $d_1(u, v) < \delta$, então $d_2(f(u), f(v)) < \epsilon$.*

Exemplo 6.18. *Dadas as funções $s, p : \mathbb{R}^2 \to \mathbb{R}$ definidas por:*

$$s(x, y) = x + y.$$
$$p(x, y) = xy.$$

Vamos estudar a continuidade das funções em $(a, b) \in \mathbb{R}^2$, considerando a norma do máximo em $\mathbb{R}^2$. Primeiramente para a função s, dado $\epsilon > 0$, seja $\delta = \epsilon/2$, tal que se $\|(x, y) - (a, b)\|_\infty < \delta$, então:

$$|s(x, y) - s(a, b)| = |x - a + y - b| \leq |x - a| + |y - b| < \delta + \delta = \epsilon.$$

Já para a função p, dado $\epsilon > 0$, seja

$$\delta = \min\{\sqrt{\epsilon/3}, \epsilon/3(|a| + 1), \epsilon/3(|b| + 1)\},$$

tal que se $\|(x, y) - (a, b)\|_\infty < \delta$, então:

$$|p(x, y) - p(a, b)| = |xy - ab| = |(x - a)(y - b) + (x - a)b + a(y - b)| \leq$$
$$|(x - a)(y - b)| + |(x - a)b| + |a(y - b)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

As funções s e p são contínuas em (a,b). Dizemos que s é uniformemente contínua, pois o δ depende apenas do ϵ. Entretanto p não é uniformemente contínua pois o δ depende do ponto escolhido.

Definição 6.17. *Sejam X um conjunto, Y um espaço métrico, $f_n : X \to Y$ uma sequência de mapeamentos para $n \in \mathbb{N}$ e $f : X \to Y$ um mapeamento. Dizemos que f_n converge uniformemente para f se para todo $\epsilon > 0$, existe um $N \in \mathbb{N}$, tal que para todo $n \geq N$, tem-se $d(f_n(u), f(u)) < \epsilon$, para todo $u \in X$.*

Exemplo 6.19. *Uma norma é um mapeamento uniformemente contínuo. De fato, seja X um espaço vetorial normado, $\|.\| : X \to \mathbb{R}$ uma norma e $v \in X$ qualquer, então dado $\epsilon > 0$, tome $\delta = \epsilon$. Tem-se*

$$\|u - v\| < \delta \Rightarrow |\, \|u\| - \|v\| \,| \leq \|u - v\| < \delta = \epsilon,$$

da desigualdade triangular.

6.5 Teorema de Bolzano-Weierstrass

Definição 6.18. *Seja S um conjunto. Uma relação de equivalência $\sim$ em S é uma relação binária tal que para todos u, v, w em S, tem-se*

i. $u \sim u$,

ii. $u \sim v$, *então* $v \sim u$,

iii. *Se* $u \sim v$ *e* $v \sim w$, *então* $u \sim w$,

onde $i., ii., iii.$ são as propriedades ditas reflexiva, simétrica e transitiva, respectivamente. A classe de equivalência contendo $u \in S$ é definida por:

$$[u] = \{v \in S \colon v \sim u\}$$

O conjunto das classes de equivalências é denotado por $S/\sim$ e a projeção canônica é definida pelo mapeamento $\pi : S \to S/\sim$ que leva u em $[u]$.

Observação 6.7. *Observe que S é a união disjunta de classes de equivalência e que, portanto, a coleção dos conjuntos $U \subset S/\sim$, tal que $\pi^{-1}(U)$ é aberta em S é uma topologia quociente, quando S é uma topologia.*

Exemplo 6.20 (Toro)**.** *Considere o espaço $\mathbb{R}^2$. A relação de $\sim$ definida por:*

$$(a,b) \sim (c,d) \text{ se } a - c \in \mathbb{Z} \text{ e } b - d \in \mathbb{Z}$$

é uma relação de equivalência. Então o conjunto das classes de equivalência denominado por $2 - toro$ é:

$$\mathbb{T}^2 = \mathbb{R}^2 / \sim .$$

Ao colocarmos que:

$$[(a, b)] + [(c, d)] = [(a, b) + (c, d)],$$

a classe de equivalência herda uma estrutura de grupo. O $\mathbb{T}^n$ é definido de forma similar.

Definição 6.19. *Seja X um conjunto. Uma cobertura por abertos de X é uma coleção*

$$\Lambda = \{A_\alpha \colon \alpha \in I\}$$

de conjuntos abertos A_α, tal que $X \subseteq \bigcup_{\alpha \in I} A_\alpha$.

Exemplo 6.21. *Considere o conjunto:*

$$B[0; 1] = \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 \leq 1\}.$$

Considere o conjunto das bolas abertas:

$$\Lambda = \{B((1, 0); 1), B((-1, 0); 1), B((0, -1); 1), B((0, 1); 1)\}.$$

Observe que Λ é uma cobertura aberta de $B[0, 1]$, conforme pode ser visualizada na figura a seguir.

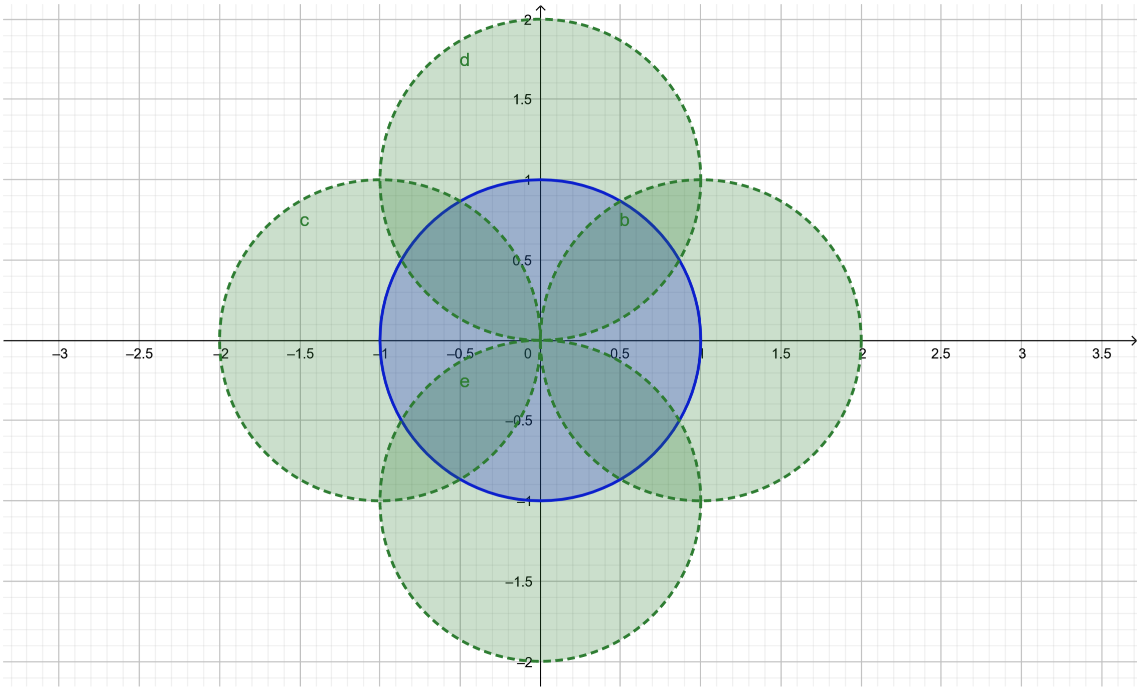

Figura 6.3: A figura representa o gráfico de uma bola fechada sendo coberta por 4 bolas abertas no $\mathbb{R}^2$.

Um dos conceitos mais importantes da topologia é o conceito de conjuntos compactos, cuja a definição é a seguinte:

Definição 6.20. *Seja X um espaço topológico. Se para toda cobertura de X por conjuntos abertos existir uma subcobertura finita, então X é compacto.*

Proposição 6.3. *As seguintes propriedades são válidas em topologias compactas:*

i. *Se X é compacto e $C \subset X$ é fechado, então C é compacto.*

ii. *A imagem de uma mapeamento contínuo de um conjunto compacto é compacta*

iii. *Se X é Hausdorff e $C \subset X$ é compacto, então C é fechado.*

iv. *Sejam $f : X \to Y$ um mapeamento contínuo, X compacto e Y Haussdorff. Então f é fechado. Se f é uma bijeção, X e Y são homeomorfos.*

Demonstração. Demonstração de *i.*: Toda cobertura aberta de X será uma cobertura aberta de C, como X é compacto, então admite uma subcobertura finita. Portanto C admite uma subcobertura finita, sendo assim compacto.
Demonstração de *ii.*: Considere $\{U_\alpha\}$ uma cobertura por abertos de $f(X)$. Então $\{f^{-1}(U_\alpha)\}$ é uma cobertura de abertos de X. Portanto existe $\{f^{-1}(U_{\alpha_k})\}$, com

$k = 1, 2, ..., n$, uma cobertura finita de X. Como f é contínua, $\{U_{\alpha_k}\}$ é uma cobertura de abertos finita de $f(X)$. Portanto a imagem é compacta.
Demonstração de *iii.*: Seja $v \in X \setminus C$ um elemento qualquer. Como X é Hausdorff, para todo $u \in C$, u e v possuem vizinhanças disjuntas. Então seja $\{U_\alpha\}$ a reunião de vizinhanças abertas de $u \in C$ com as correspondentes vizinhanças abertas disjuntas de $v \in X \setminus C$, $\{V_\alpha\}$. Como C é compacto, existe uma subcobertura $\{U_{\alpha_k}\}$ finita. Logo a intersecção finita das correspondentes vizinhanças de v, $\bigcap V_{\alpha_k}$ é aberta e disjunta de C. Demonstração de *iv.*: temos que $f(X)$ é compacto de *ii.* e fechado de *iii.*. Se ela é uma bijeção, tem-se que $f^{-1} : Y \to X$ é contínua pois a imagem inversa de todo subconjunto fechado de X é fechado em Y, logo f é um homeomorfismo. □

Teorema 6.2 (Teorema de Bolzano-Weierstrass)**.** *Seja X um espaço de Hausdorff enumerável de primeira ordem e compacto. Então toda sequência possui uma subsequência convergente.*

Demonstração. Suponha que X seja compacto e uma sequência (x_k) de X não possui nenhuma subsequência convergente. Assuma que os elementos da sequência são distintos. Da Proposição 6.3 o fecho

$$cl(\{x_k\}) = \{x_k\} \subset X$$

é compacto. Como X é enumerável de primeira ordem, então cada x_n da sequência contém uma vizinhança V_n que não contém nenhum outro elemento x_j, caso contrário x_n seria o limite desta sequência. Portanto a cobertura $\{V_k\}$ é uma cobertura do conjunto compacto $\{x_k\}$ que não contém nenhuma subcobertura finita, absurdo. □

Definição 6.21. *Um espaço métrico X é dito totalmente limitado se ele é um subconjunto de uma união finita de bolas abertas.*

Exemplo 6.22. *O subconjunto $Q \subset \mathbb{R}^2$ definido por:*

$$Q = \{(x, y) \in \mathbb{R}^2 \colon x \in [-1, 1]\} \cap \{(x, y) \in \mathbb{R}^2 \colon y \in [-1, 1]\}$$

é totalmente limitado, pois a bola aberta

$$B(0, 3) = \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 < 3\}$$

contém o conjunto Q.

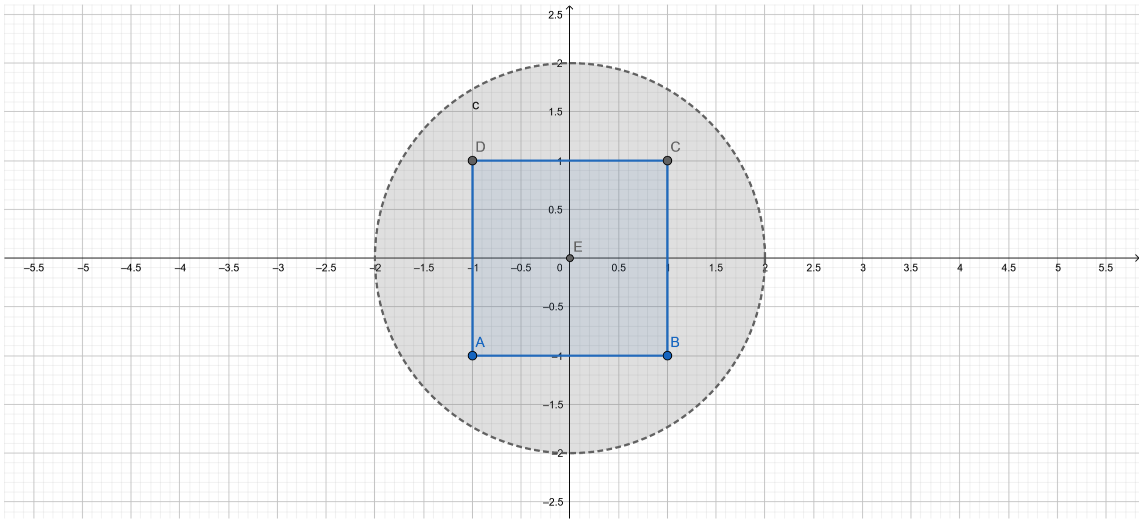

Figura 6.4: A figura representa o gráfico de um conjunto totalmente limitado Q contido numa bola aberta em $\mathbb{R}^2$.

Normas e produtos internos definidos no Capítulo 1, podem ser generalizados para espaços ou conjuntos mais genéricos, com as mesmas propriedades. Ou seja, uma norma satisfaz as propriedades de definição positiva, homogeneidade e desigualdade triangular.

Observação 6.8. *Observe que em um espaço com produto interno $(X, \langle ., . \rangle)$ definindo a norma $\|u\| = \langle u, u\rangle^{1/2}$, define um espaço normado $(X, \|.\|)$. Também, é fácil verificar que um espaço normado $(X, \|.\|)$ com definição $d(u_1, u_2) = \|u_1 - u_2\|$, define um espaço métrico (X, d). Dessa forma, temos que em ordem de generalidade, espaços com produto interno são menos gerais que espaços normados que são menos gerais que espaços métricos que são menos gerais que espaços topológicos.*

Deste ponto em diante, trabalharemos em espaços vetoriais de dimensão infinita, ou seja, definiremos algumas propriedades para conjuntos convexos nesses espaços. A definição de espaços vetoriais de dimensão infinita é similar para o caso finito. Da mesma forma, define-se subespaços vetoriais ou seja, os subconjuntos dos espaços vetoriais que são fechados com respeito à soma e à multiplicação por escalar.

6.6 Espaços Lineares, Normados, Convexos e Equivalência entre Normas

Definição 6.22. *Seja X um espaço vetorial sobre um corpo $\mathbb{R}$ e $S \subseteq X$ um subconjunto. A envoltória linear de S é a interseção de todos os subespaços vetoriais que contém S, chamamos $LS(S)$ ou $span(S)$.*

Observação 6.9. *Observe que o $LS(S)$ é o menor subespaço vetorial de X que contém S.*

Teorema 6.3. *Seja X um espaço vetorial sobre um corpo $\mathbb{R}$ e $S \subseteq X$ um subconjunto. Então*

$$LS(S) = \{\textstyle\sum_{j=1}^{N} \alpha_j u_j : u_j \in S,\ \alpha_j \in \mathbb{R}\}.$$

Demonstração. Considere o conjunto

$$Y = \{\textstyle\sum_{j=1}^{N} \alpha_j u_j : u_j \in S,\ \alpha_j \in \mathbb{R}\}.$$

Observe que $Y \subseteq X$ é um subespaço vetorial. Adicionalmente $S \subseteq Y$, pois qualquer elemento de $u \in S$, pode ser escrito como $1u \in Y$. Seja $Z \subseteq X$ um subespaço vetorial que contém S. Dado $u \in Y$ qualquer, $u = \sum_{j=1}^{N} \alpha_j u_j$, mas como $u_j \in S$, então $\sum_{j=1}^{N} \alpha_j u_j \in Z \Rightarrow u \in Z$. Logo, $Y \subseteq Z$, então Y é o menor subespaço vetorial que contém S. □

Definição 6.23. *Seja X um espaço vetorial sobre o corpo $\mathbb{R}$ e $Y \subseteq X$ um subespaço. Defina $\sim$ uma relação em X, tal que se $x \sim y$ se $x - y \in Y$.*

Observação 6.10. *Observe que se $\sim$ é uma relação de equivalência, o conjunto de equivalência em X de x é o $[x]$ e o subespaço quociente $X/Y = \{[x] : x \in X\}$ é um subespaço vetorial.*

Subconjuntos convexos de espaços vetoriais são definidos de forma similar ao caso $\mathbb{R}^n$, assim como a envoltória convexa de qualquer subconjunto.

Teorema 6.4. *Seja X um espaço vetorial normado sobre $\mathbb{R}$. Considere que a $dim(X) < \infty$, tal que $\{x_1, ...x_N\}$ é uma base de X. Dado $x \in X$, escrevemos*

$$x = \textstyle\sum_{j=1}^{N} \alpha_j x_j,$$

onde α_j são unicamente determinados para cada x e $j = 1, ..., N$. Então existe $c_0 > 0$ uma constante real, tal que

$$\textstyle\sum_{j=1}^{N} |\alpha_j| \leq c_0 \|x\|.$$

Demonstração. Suponha que não exista tal c_0, então para cada $k \geq 1$, existe $y_k \in X$, não nulo, tal que $\sum_{j=1}^{N} |\alpha_j(y_k)| > k \|y_k\|$, onde $\alpha_j(y_k)$ são os coeficientes unicamente determinados pela combinação linear da base de X. Defina $z_k = y_k/(\sum |\alpha_j(y_k)|)$, então $\sum |\alpha_j(z_k)| = 1$ e $k \|z_k\| < 1 \Rightarrow \|z_k\| < 1/k$. Mas veja que para cada $j = 1, ..., N$, $|\alpha_j(z_k)| \leq 1$, então essa sequência real em k é limitada, portanto, possui uma subsequência convergente. Então, por um processo indutivo, começando em $j =$

1, poderemos tomar subsequências que convergem, definindo assim os coeficientes limites $|\alpha_j(z_k)| \to |\alpha_j|$. Mas veja que a $\|z_k\| \to 0$, porém no limite $\sum |\alpha_j| = 1$, absurdo. □

Definição 6.24. *Duas normas num espaço vetorial X são equivalentes, se elas induzem a mesma topologia.*

Teorema 6.5. *Duas normas quaisquer $\|.\|_1$ e $\|.\|_2$ de um espaço vetorial X são ditas equivalentes se, e somente se, existe uma constante C, tal que para todo $u \in X$, tem-se*

$$\frac{1}{C}\|u\|_2 \leq \|u\|_1 \leq C\|u\|_2$$

Demonstração. A definição de induzir a mesma topologia, implica que dado uma bola definida em uma norma centrada na origem, existem duas bolas definidas na outra norma, também centrada na origem, que contém e está contida na bola de referência. A desigualdade apresentada, segue de forma simples. □

Teorema 6.6. *Seja X um espaço vetorial normado nos reais. Considere que a $dim(X) < \infty$. Então quaisquer duas normas de X são equivalentes.*

Demonstração. Seja $x \in X$ definido por

$$x = \textstyle\sum_{j=1}^N \alpha_j(x)x_j,$$

onde $\{x_1, ..., x_N\}$ é uma base de X. Defina a $\|.\|_0$, como sendo, $\|x\|_0 = \sum_{j=1}^N |\alpha_j(x)|$. Portanto, basta provar que qualquer norma $\|.\|$ em X é equivalente a $\|.\|_0$. Com efeito, a primeira inclusão segue da desigualdade triangular, pois:

$$\|x\| = \left\|\textstyle\sum_{j=1}^N \alpha_j(x)x_j\right\| \leq \textstyle\sum_{j=1}^N |\alpha_j(x)|\,\|x_j\| \leq C\textstyle\sum_{j=1}^N |\alpha_j(x)| = C\|x\|_0,$$

onde $C = \max_{1\leq j\leq N}\{\|x_j\|\}$. A outra inclusão segue do Teorema 6.4. □

6.7 Elementos de Teoria dos Conjuntos

Definição 6.25. *Seja X um conjunto. Uma relação R entre elementos x, y, z de X é dita uma ordem parcial, se satisfaz:*

i. xRx, *para todo* $x \in X$ *(reflexiva)*

ii. *Se* xRy *e* yRx, *então* $x = y$ *(antisimétrica)*

iii. *Se* xRy *e* yRz *então* xRz *(transitiva)*

A relação de ordem R é dita uma cadeia (ou totalmente ordenado), se para cada $x, y \in X$, tem-se que xRy ou yRx. Um conjunto é bem ordenado se é uma cadeia e todo subconjunto não vazio $Y \subset X$, possui um primeiro elemento, ou seja, existe um $p \in Y$, tal que pRy para todo $y \in Y$. Um elemento $b \in X$ é uma cota superior de uma cadeia $Y \subset X$, se yRb, para todo $y \in Y$. Um elemento máximo de um conjunto ordenado X é um elemento $m \in X$, tal que se $x \in X$ e mRx, então $x = m$. Podemos simbolizar R com $\leq$.

Ainda, na teoria de conjuntos, podemos enunciar o Lema de Zorn, ver lema a seguir, que é equivalente ao axioma da escolha. Este lema será utilizado na demonstração do teorema de Hahn-Banach, que por sua vez, será fundamental no entendimento das propriedades de conjuntos convexos em espaços vetoriais de dimensão infinita.

Lema 6.1 (Lema de Zorn)**.** *Seja X um conjunto não vazio com uma ordem parcial. Suponha que para todo subconjunto de X totalmente ordenado, exista uma cota superior. Então X possui ao menos um elemento máximo.*

6.8 A Existência de Base e o Teorema de Hanh-Banach

Definição 6.26 (Base de Hamel)**.** *Seja X um espaço vetorial. Dizemos que $\{x_\alpha \colon \alpha \in I\}$ é uma base de X se:*

i. O subconjunto $\{x_{\alpha_1}, .., x_{\alpha_n}\}$ é linearmente independente, para quaisquer $\alpha_j \in I$, distintos.

ii. $X = span(\{x_\alpha \colon \alpha \in I\})$.

Teorema 6.7 (Existência da Base de Hamel)**.** *Seja X um espaço vetorial não vazio. Então X possui uma Base de Hamel.*

Demonstração. Defina Ω como uma coleção de conjuntos

$$C_I = \{x_\alpha \colon x_\alpha \in X \text{ e } \alpha \in I\},$$

em que qualquer subconjunto finito de C_I é linearmente independente. Defina uma ordem parcial em Ω, denotada por $<$, satisfazendo:

$$\{x_\alpha \colon \alpha \in I\} < \{y_\alpha \colon \alpha \in I'\},$$

se para cada $\alpha \in I$, existe um $\alpha' \in I'$, tal que $x_\alpha = y_{\alpha'}$. Pode-se verificar que $<$ é uma relação de ordem. Considere $\Omega' \subset \Omega$ um subconjunto totalmente ordenado e afirmamos que essa cadeia possui uma cota superior. Seja $Y = \bigcup_{\theta \in J}\{x_\alpha \colon \alpha \in I_\theta\}$, que é uma reunião de elementos para todas as famílias de índices I em Ω'. De fato, por construção Y é uma cota superior em Ω'. Mas veja que Y é um elemento de Ω, pois é a reunião de conjuntos de Ω. Pelo Lema de Zorn, Ω possui um elemento máximo, que chamamos de B. Afirmamos que $X = span(B)$. Como todo elemento de B, por construção está em X, então $B \subseteq X$. Suponha que exista um elemento de X que não esteja em $span(B)$, digamos x. Tome um subconjunto de B finito, ou seja, $S = \{x_{\theta_1}, ..., x_{\theta_k}\}$, com elementos linearmente independentes. Considere a equação $\lambda x + \sum_j^k \lambda_j x_{\theta_j} = 0$. Se $\lambda = 0$, por hipótese, $\lambda_j = 0$ e todos os elementos seriam linearmente independentes, o que implicaria $x \in B$, absurdo. Logo $\lambda \neq 0$, o que implica que x é uma combinação linear de elementos de B, logo pertence ao $span(B)$, absurdo. Portanto, B é uma base para o espaço vetorial X. $\square$

A seguir apresentaremos o teorema de Hahn-Banach, que será de grande utilidade no desenvolvimento de análise convexa em espaços vetoriais de dimensão infinita.

Definição 6.27. *Seja X um espaço vetorial sobre $\mathbb{R}$. A função $p : X \to \mathbb{R}_+$ é dita positiva subaditiva se satisfaz as seguintes condições:*

i. $p(\alpha x) = \alpha p(x)$*, para* $\alpha \geq 0$ *e* $x \in X$*,*

ii. $p(x + y) \leq p(x) + p(y)$.

Exemplo 6.23. *Seja X um espaço vetorial normado sobre $\mathbb{R}$. A função norma $\|.\| : X \to \mathbb{R}$ é uma função positiva subaditiva.*

Teorema 6.8 (Teorema de Hahn-Banach)**.** *Seja X um espaço vetorial não vazio sobre os reais. Considere $Y \subseteq X$ um subespaço vetorial. Suponha que exista $t : Y \to \mathbb{R}$, função linear e $p : X \to \mathbb{R}$, função positiva subaditiva, tal que $t(y) \leq p(y)$, para todo $y \in Y$. Então existe uma função linear $T : X \to \mathbb{R}$, tal que:*

i. $T(y) = t(y)$*, para todo* $y \in Y$.

ii. $T(x) \leq p(x)$*, para todo* $x \in X$.

Demonstração. Suponha que Y seja um subespaço próprio de X, por outro lado o teorema seria trivial. Daí, existe um elemento $z \in X$ que não pertence a Y. Defina o conjunto

$$Y_z = \{y + \alpha z \colon y \in Y,\, \alpha \in \mathbb{R}\}.$$

Verifica-se que Y_z é fechado com relação à soma a à multiplicação por escalar sendo, portanto, um subespaço vetorial de X. Dados $y_1, y_2 \in Y$ e $\alpha_1, \alpha_2 \in \mathbb{R}$, tem-se

$$y_1 + \alpha_1 z = y_2 + \alpha_2 z \iff (\alpha_1 - \alpha_2)z = y_2 - y_1 \in Y \iff \alpha_1 = \alpha_2 \text{ e } y_1 = y_2,$$

pois $z \notin Y$. Logo os elementos de Y_z são unicamente determinados. Afirmamos que existe uma função linear $t_z : Y_z \to \mathbb{R}$ que satisfaz as propriedades $i.$ e $ii.$ Defina $t_z : Y_z \to \mathbb{R}$:

$$t_z(y + \alpha z) = t(y) + \alpha t_z(z).$$

Com essa definição, t_z é uma função linear e satisfaz a condição $i.$ Falta definir o valor de $t_z(z)$. Considere dois casos: Caso 1: $\alpha > 0$. Para que $ii.$ seja satisfeita devemos ter

$$t_z(y + \alpha z) \leq p(y + \alpha z) \Rightarrow t(y) + \alpha t_z(z) \leq p(y + \alpha z),$$

como $\alpha > 0$, então

$$t_z(z) \leq 1/\alpha(p(y + \alpha z) - t(y)) = p(y/\alpha + z) - t(y/\alpha).$$

Como $\alpha > 0$, então $t_z(z) \leq p(y + z) - t(y)$, para todo $y \in Y$. Portanto,

$$t_z(z) \leq \inf_y \{p(y + z) - t(y)\}.$$

Caso 2: $\alpha < 0$. Considerando $-\alpha > 0$, então aplicando os mesmos passos do Caso 1, conclui-se que

$$t_z(z) \geq \sup_y \{t(y) - p(y - z)\}.$$

Observe que dados $y_1, y_2 \in Y$ tem-se

$$t(y_1 + y_2) \leq p(y_1 + y_2) = p(y_1 + z + y_2 - z) \leq p(y_1 + z) + p(y_2 - z) \Rightarrow$$
$$p(y_1 + z) - t(y_1) \geq t(y_2) - p(y_2 - z).$$

Basta definir $t_z(z)$ qualquer valor entre as cotas descritas acima, que satisfaz $ii.$ Defina $\Omega = \{(Z, t_Z)\}$, como conjunto de todos os subespaços $Z \subseteq X$, tal que $Y \subseteq Z$. De fato, pela demonstração anterior, os elementos de Ω estão bem definidos e satisfazem $i.$ e $ii.$. Defina uma relação de ordem parcial em Ω, $<$, tal que $(Z_1, t_{Z_1}) < (Z_2, t_{Z_2})$, se $Z_1 \subseteq Z_2$ e $t_{Z_2}(z) = t_{Z_1}(z)$, para todo $z \in Z_1$. Seja $\Omega' \subseteq \Omega$, um conjunto totalmente ordenado. Seja $Z = \bigcup_{\alpha \in I} Z_\alpha$. Veja que $Z \in \Omega$ tendo em vista que é a reunião de elementos em Ω e, por essa definição, é uma cota superior de Ω'. Pelo Lema de Zorn, Ω possui um elemento máximo, digamos, (M, T). Temos que $M \subseteq X$, mas afirmamos que $M = X$. Suponha que $M \neq X$, então existe $w \in X$ que não pertence a M. Dessa forma, é possível construir um conjunto $W = \{m + \alpha w\}$ e uma função linear T_W que satisfaz as condições *i.* e *ii.*. Portanto, $M \subset W$, $W \subset X$ e teríamos em Ω, $(M,T) < (W,T_W)$, absurdo. $\square$

6.9 Pontos Interiores, Conjuntos Convexos, Hiperplanos e Semiespaços em Espaços Vetoriais de Dimensão Infinita

Definição 6.28. *Seja X um espaço vetorial sobre os reais e $Y \subseteq X$ um subconjunto. Dizemos que $y \in Y$ é um elemento interior a Y quando para todo $x \in X$ existe um $\epsilon = \epsilon(y) > 0$, tal que $y + tx \in Y$ para todo $t, |t| \leq \epsilon$.*

Definição 6.29. *Seja X um espaço vetorial sobre os reais e $K \subset X$ um conjunto convexo. Seja $w \in K$ um ponto interior de K. Definimos a função $p : X \to \mathbb{R}_+$, como sendo*

$$p(x) = \inf_{a>0}\{a \colon w + (1/a)x \in K\}.$$

Observação 6.11. *Observe que a função da Definição 6.29 está bem definida utilizando a definição de ponto interior.*

Proposição 6.4. *A função p da Definição 6.29 satisfaz as seguintes condições:*

i. $p(\alpha x) = \alpha p(x)$, *para todos* $\alpha > 0$ *e* $x \in X$.

ii. $p(x+y) \leq p(x) + p(y)$, *para todos* $x, y \in X$.

iii. *Dado* $x \in X$, *tal que* $w + x$ *é um elemento de* K *então* $p(x) \leq 1$.

iv. *Dado* $x \in X$, *então* $p(x) < 1$ *se e somente se* $w + x$ *é um ponto interior de* K.

Demonstração. Parte *i.*: Se $t \in \{a > 0 \colon w + (1/a)x \in K\}$ então $t' > t$, também pertence a esse conjunto. Com efeito, defina $\theta = t/t' \in [0,1]$. Pela convexidade de K, temos

$$(1-\theta)w + \theta(w + (1/t)x) \in K \Rightarrow w + (\theta/t)x \in K \Rightarrow w + (1/t')x \in K.$$

Considere:

$$b = p(x) = \inf\{a > 0 \colon w + (1/a)x \in K\}.$$

Dado $\alpha > 0$, tem-se:

$$b = \inf\{a > 0 \colon w + (1/\alpha a)\alpha x \in K\} = \inf\{c/\alpha > 0 \;;\; w + (1/c)\alpha x \in K\}.$$

Então se

$$d = \inf\{c > 0\colon w + (1/c)\alpha x \in K\} \Rightarrow b = d/\alpha.$$

Observe que $d = f(\alpha x) = \alpha b = \alpha f(x)$.
Parte *ii.*: Dados $x, y \in X$. Suponha que $p(x + y) > p(x) + p(y)$ e considere $\epsilon > 0$, tal que

$$p(x) + p(y) + \epsilon < p(x + y).$$

Tome $b, c \in \mathbb{R}$, tal que

$$p(x) \leq b \leq p(x) + \epsilon/2 \text{ e}$$
$$p(y) \leq c \leq p(y) + \epsilon/2.$$

Defina $\lambda = c/(b + c) \in [0, 1]$. Pela convexidade de K, tem-se:

$$\lambda(w + (1/b)x) + (1 - \lambda)(w + (1/c)y) \in K \Rightarrow w + (1/(b + c))(x + y) \in K.$$

Daí,

$$b + c \in \{a : w + (1/a)(x + y) \in K\}.$$

Logo, $p(x + y) \leq (b + c) \leq p(x) + p(y) + \epsilon$, absurdo.
Parte *iii.*: Seja $\delta > 0$, tal que $(w + x) + \delta x \in K$. Então

$$w + (1 + \delta)x \in K \Rightarrow w + (1/(1/(1 + \delta))x \in K.$$

Daí,

$$1/(1 + \delta) \in \{a > 0\colon w + (1/a)x \in K\}.$$

Logo $p(x) \leq 1/(1 + \delta) \leq 1$.
Parte *iv.*: Dado $y \in X$, precisamos encontrar um $\epsilon = \epsilon(y) > 0$, tal que

$$(w + x) + ty \in K,$$

para todo $|t| \leq \epsilon$. De $p(x) < 1$, existe um $b < 1$, tal que $w + (1/b)x \in K$. Veja que podemos escrever $c = 1/b - 1$, com $c > 0$. Se $w \in K$ é interior então existe um $\delta = \delta(y) > 0$, tal que $w + ty \in K$, para todo $|t| \leq \delta$. Como K é convexo, seja $\theta = 1/(1 + c) \in [0, 1]$. Logo

$$(1 - \theta)(w + ty) + \theta(w + (1 + c)x) \in K \Rightarrow (w + x) + (ct/(1 + c))y \in K.$$

Daí, escolha $\epsilon = (1 + c)\delta/c$. Se $|t| \leq \epsilon$ então $(w + x) + ty \in K$. Por outro lado, se $w + x \in int(K)$ existe um $\epsilon > 0$, tal que $w + x + tx \in K$, para todo $|t| \leq \epsilon$. Logo existe um t_0, tal que $1 + t_0 < 1$ e $1 + t_0 \in \{a > 0\colon w + (1/a)x \in K\}$ o que conclui a demonstração. $\square$

Definição 6.30. *Seja X um espaço vetorial sobre os reais e $t : X \to \mathbb{R}$ uma função linear não nula. O conjunto $H = \{x \in X : t(x) = c\}$ é denominado hiperplano, onde $c \in \mathbb{R}$. O conjunto $S = \{x \in X : t(x) < c\}$ é denominado semiespaço. Caso a desigualdade não seja estrita, dizemos que o semiespaço é fechado.*

Teorema 6.9. *Seja X um espaço vetorial sobre os reais e $K \subset X$ um subconjunto próprio de X, convexo e não vazio. Suponha que todos os elementos de K sejam interiores. Dado $y \in X$, tal que $y \notin K$, existe uma função linear $T : X \to \mathbb{R}$ e um $c \in \mathbb{R}$, tal que $T(x) < c$, para todo $x \in K$ e $T(y) = c$.*

Demonstração. Suponha sem perda de generalidade que $0 \in K$, ou seja, 0 é um ponto interior de K. Defina a função $p : X \to \mathbb{R}_+$, tal que

$$p(x) = \inf\{a > 0 : (1/a)x \in K\}.$$

Pela Proposição 6.4, a função p é positiva subaditiva. Seja $z \in K$ então z é um ponto interior e pela Proposição 6.4, $p(z) < 1$. Observe que

$$1 \notin \{a > 0 : (1/a)y \in K\}.$$

Logo, $p(y) \geq 1$. Defina

$$Y = \{\alpha y : \alpha \in \mathbb{R}\}.$$

Observe que Y é fechado com relação a soma e a multiplicação por escalar, então é um subespaço X. Defina $t : Y \to \mathbb{R}$ uma função linear, tal que

$$t(y) = p(y) \text{ e}$$
$$t(\alpha w) = \alpha t(w), \text{ para todo } w \in Y.$$

Veja que t está bem definida. Se $\alpha = 0$ então

$$t(0y) = 0t(y) = 0p(y) = 0.$$

Se $\alpha > 0$ então

$$t(\alpha y) = p(\alpha y) = \alpha t(y) = \alpha t(y).$$

Se $\alpha < 0$ então

$$\alpha t(w) = t(\alpha w) = p(\alpha w) \geq -p(-\alpha w) = \alpha p(w),$$

ou seja, $t(w) \leq p(w)$, para todo $w \in Y$. Daí, pelo Teorema de Hahn-Banach, existe uma função linear $T : X \to \mathbb{R}$, que é uma extensão, ou seja,

$$T(w) = t(w) \text{ para todo } w \in Y \text{ e}$$
$$T(x) \leq p(x) \text{ para todo } x \in X.$$

Se $z \in K$, então $T(z) \leq p(z) < 1 \leq p(y) = T(y) = c$. $\square$

Corolário 6.10. *Seja X um espaço vetorial sobre os reais e $K \subset X$ um subconjunto próprio de X, convexo e não vazio. Suponha que exista um elemento de K que seja ponto interior. Dado $y \in X$, tal que $y \notin K$, existe uma função linear $T : X \to \mathbb{R}$ e um $c \in \mathbb{R}$, tal que $T(x) \leq c$, para todo $x \in K$ e $T(y) = c$.*

Teorema 6.11. *Seja X um espaço vetorial sobre os reais e $K_1, K_2 \subset X$ subconjuntos próprios de X, convexos e não vazios. Suponha que $K_1 \cap K_2 = \emptyset$ e que K_1 possui um ponto interior. Então existe uma função linear $T : X \to \mathbb{R}$ e um $c \in \mathbb{R}$, tal que $T(x) \leq c \leq T(y)$, para todo $x \in K_1$ e $y \in K_2$.*

Demonstração. Defina $K = K_1 - K_2 = \{x - y \colon x \in K_1, \, y \in K_2\}$. É fácil verificar que K é um conjunto convexo. Adicionalmente, $0 \notin K$, pois $K_1 \cap K_2 = \emptyset$. Veja que se $z \in K_1$ é um ponto interior, então $z - y \in K$ é um ponto interior para todo $y \in K_2$. De fato, dado $x \in X$, existe um $\epsilon = \epsilon(x) > 0$, tal que $z + |t|x \in K_1$, para todo $|t| \leq \epsilon \Rightarrow z + |t|x - y = (z - y) + |t|x \in K$. Como $0 \notin K$, pelo corolário anterior existe uma função linear $T : X \to \mathbb{R}$, tal que $T(w) \leq T(0) = 0$, para todo $w \in K \Rightarrow T(x - y) \leq 0 \Rightarrow T(x) \leq T(y)$, para todo $x \in K_1, y \in K_2$. Defina $c = \sup_{x \in K_1}\{T(x)\}$, temos que $c \leq T(y)$, para todo $y \in K_2$, concluindo a demonstração. $\square$

Definição 6.31. *Seja X um espaço vetorial sobre o corpo $\mathbb{K}$ e $t : X \to \mathbb{K}$ uma função linear. Então t é contínua em $x \in X$ se para toda sequência $(x_n)_{n \in \mathbb{N}} \in X$, tal que $x_n \to x$, então $t(x_n) \to t(x)$. Dizemos que t é limitada se existe $c \in \mathbb{K}$, tal que para todo $x \in X$, tem-se $|t(x)| \leq c \|x\|$.*

Proposição 6.5. *Seja X um espaço vetorial sobre o corpo $\mathbb{K}$ e $t : X \to \mathbb{K}$ uma função linear é limitada se, e somente se, é contínua.*

Demonstração. Suponha que a função linear seja limitada. Dado $x \in X$, seja $(x_n)_{n \in \mathbb{N}}$ uma sequência de X convergente para x. Se a sequência converge, dado $\epsilon > 0$, existe $N \in \mathbb{N}$, tal que para todo $n \geq N$, tem-se $\|x_n - x\| < \epsilon / c$, onde c é a constante de limitação. Então $|t(x_n) - t(x)| = |t(x_n - x)| \leq c \|x_n - x\| < \epsilon$, portanto $t(x_n) \to t(x)$ e t é contínua. Agora, suponha que t não seja limitada, então dado $n \in \mathbb{N}$, podemos construir uma sequência $(x_n)_{n \in \mathbb{N}}$, tal que $t(x_n) \geq n \|x_n\|$. Defina a sequência $(y_n)_{n \in \mathbb{N}}$, tal que $y_n = x_n / (\sqrt{n} \|x_n\|)$. Então $\|y_n\| = 1/\sqrt{n}$ e $t(x_n) \geq n \|x_n\| \Rightarrow t(y_n) \geq n \|x\| / (\sqrt{n} \|x_n\|) = \sqrt{n}$. Mas veja que $y_n \to 0$, mas $t(y_n)$ não converge para $t(0) = 0$, portanto t não é contínua o que conclui a demonstração. $\square$

6.10 Espaços de Banach e Hilbert

Definição 6.32. *Seja $(X, \|.\|)$ um espaço vetorial normado. Se sua métrica correspondente é completa, dizemos que $(X, \|.\|)$ é um espaço de Banach. Se $(X, \langle ., . \rangle)$ é um espaço vetorial com produto interno, cuja a respectiva métrica é completa, dizemos que $(X, \langle ., . \rangle)$ é um espaço de Hilbert. Ambos os espaços podem ser definidos nos campos dos números reais ou complexos.*

Por exemplo, pode ser demonstrado que o espaço $\mathbb{R}^n$ é completo. O $\mathbb{R}^n$ com sua respectiva norma é um espaço de Banach. O $\mathbb{R}^n$ com seu respectivo produto interno é um espaço de Hilbert.

Proposição 6.6. *Seja X um espaço vetorial sobre os reais com produto interno. Então para todos $x, y \in X$, tem-se*

i. (Desigualdade de Scwarz) $\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle$.

ii. (Regra de Paralelograma) $2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2$, onde $\|.\| : X \to \mathbb{R}$ é uma função, com $\|x\|^2 = \langle x, x\rangle$.

Demonstração. i. Seja $t \in \mathbb{R}$, temos que $\langle x + ty, x + ty\rangle \geq 0 \Rightarrow \langle x, x\rangle + 2t\langle x, y\rangle + t^2\langle y, y\rangle \geq 0$, como $t \in \mathbb{R}$, a inequação do segundo grau será válida se, e somente se, $\Delta \leq 0 \Rightarrow 4\langle x, y\rangle^2 - 4\langle x, x\rangle\langle y, y\rangle \leq 0$, o que conclui a demonstração.
ii. Temos que $\|x + y\|^2 + \|x - y\|^2 = \langle x+y, x+y\rangle + \langle x-y, x-y\rangle = 2\langle x, x\rangle + 2\langle y, y\rangle$. □

Teorema 6.12. *Seja X um espaço vetorial sobre os reais com produto interno. A função definida por $\|.\| : X \to \mathbb{R}$, como $\|x\| = \sqrt{\langle x, x\rangle}$ é uma norma.*

Demonstração. Veja que a primeira propriedade da norma é satisfeita pois é equivalente a propriedade 4 de produtos internos. Para a segunda propriedade teríamos, $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha^2\langle x, x\rangle} = |\alpha|\sqrt{\langle x, x\rangle} = |\alpha|\,\|x\|$. Para a terceira propriedade, dados $x, y \in X$, tem-se $\|x + y\|^2 = \langle x + y, x + y\rangle = \langle x, x\rangle + \langle y, y\rangle + 2\langle x, y\rangle \leq \langle x, x\rangle + \langle y, y\rangle + 2\|x\|\,\|y\| = (\|x\| + \|y\|)^2$, utilizando a desigualdade de Schwarz, demonstrando a desigualdade triangular. □

6.11 Conjuntos e Funções Convexas em Espaços de Hilbert

A partir das definições, proposições e teoremas acima, podemos introduzir os conceitos de conjuntos e funções convexas em espaços normados de dimensão infinita.

Teorema 6.13. *Seja X um espaço de Hilbert e $K \subset X$ um subconjunto próprio, convexo e fechado. Dado $x \in X$, defina $d : X \to \mathbb{R}$, como $d_K(x) = \inf\{\|x - z\| : z \in K\}$. Então dado $x \in X$, existe um único $w \in K$, tal que $d_K(x) = \|x - w\|$.*

Demonstração. Dado $x \in X$, seja $d = d_K(x)$ e defina uma sequência $(y_n)_{n \in \mathbb{N}} \in K$ tal que $\|x - y_n\|^2 < d^2 + 1/n$. Afirmamos que (y_n) é uma sequência de Cauchy. Dados $m, n \in \mathbb{N} \Rightarrow (y_n + y_m)/2 \in K$. Pela regra do paralelogramo, tem-se

$$\|y_n - y_m\|^2 = 2\,\|x - y_n\|^2 + 2\,\|x - y_m\|^2 - 4\,\|x - (y_m + y_n)/2\|^2$$

Portanto, pelas desigualdades no lado direito da equação, tem-se

$$\|y_n - y_m\|^2 < 2d^2 + 2/n + 2d^2 + 2/m - 4d^2 = 2/n + 2/m$$

O que demonstra que a sequência é de Cauchy. Como o espaço é completo, então (y_n) converge para $w \in K$, pois K é fechado. Suponha que existam $w_1, w_2 \in K$ tal que $d_K(x) = \|x - w_1\| = \|x - w_2\| = d$. Similarmente, utilizando a regra do paralelogramo, sabendo que pela convexidade de K, $(w_1 + w_2)/2 \in K$, tem-se:

$$\|w_1 - w_2\|^2 = 2\,\|x - w_1\|^2 + 2\,\|x - w_2\|^2 - 4\,\|x - (w_m + w_2)/2\|^2 < 0$$

absurdo. □

Corolário 6.14. *Seja X um espaço de Hilbert e K um subconjunto convexo e fechado de X. Então existe um elemento $u_l \in K$, tal que, $\|u_l\| = \inf\{\|u\| : u \in K\}$.*

Podemos introduzir alguns dos conceitos de funções convexas e transformadas de Fenchel apresentadas nos capítulos anteriores para o caso $\mathbb{R}^n$, para espaços de Hilbert sobre os reais.

Definição 6.33. *Seja X um espaço de Hilbert sobre um corpo $\mathbb{K}$ com a sua respectiva norma. Dizemos que a sequência $(x_n) \in X$ converge para x se $\|x_n - x\| \to 0$, quando $n \to \infty$. Dizemos que a sequência possui uma convergência fraca se para todo $y \in X$, tem-se $\langle x_n, y\rangle \to \langle x, y\rangle$, quando $n \to \infty$.*

Definição 6.34. *Seja X um espaço de Hilbert sobre um corpo dos reais. Dizemos que $f : X \to (-\infty, \infty]$ é semicontínua, inferiormente, (sci) em $x \in X$, se dado $\lambda \in \mathbb{R}$, tal que $\lambda < f(x)$, existe uma vizinhança de x, denotada por $V(x)$, tal que:*

$$\lambda < f(y), \text{ para todo } y \in V(x).$$

Definição 6.35. *Seja X um espaço de Hilbert sobre os reais. Seja $f : X \to (-\infty, +\infty]$ uma função, então definimos o conjunto epi$(f) = \{(u, k) \in X \times \mathbb{R} \colon f(u) \leq k\}$. Dizemos que a função f é convexa se o conjunto epi(f) for convexo.*

Proposição 6.7. *Seja X um espaço de Hilbert sobre os reais. Então $f : X \to (-\infty, +\infty]$ é uma função sci se, e somente se, epi(f) é fechado.*

Demonstração. Se f é uma função *sci*, então dado $x \in X$ e $\lambda \in \mathbb{R}$, com $\lambda < f(x)$ existe uma vizinhança de x, $V(x) \subset X$, tal que $\lambda < f(y)$, para todo $y \in V(x)$. Então o conjunto $C_\lambda = \{(x, \lambda) \in X \times \mathbb{R} \colon f(x) > \lambda\}$ é aberto. Daí, a reunião de conjuntos $\bigcup C_\lambda$ é aberta, portanto, seu complementar é fechado, mas veja que o complementar é exatamente o *epi*(f). Por outro lado, se *epi*(f) é fechado, então seu complementar é um conjunto aberto, ou seja, o conjunto $\{(x, \lambda) \in X \times \mathbb{R}\ ;\ f(x) > \lambda\}$ é aberto, daí, fixado um $\lambda \in \mathbb{R}$, existe uma vizinhança de $x \in X$, tal que $f(y) > \lambda$ para todo y nessa vizinhança. □

Teorema 6.15. *Seja X um espaço de Hilbert sobre os reais. Seja $f : X \to (-\infty, +\infty]$ uma função. As seguintes afirmações são equivalentes:*

i. f é convexa e sci

ii. f é o supremo de todas as funções lineares contínuas que são menores ou iguais que f

Demonstração. Primeiramente, observe que se $K \subset X$ é um subconjunto próprio, convexo não vazio em que todos os elementos são interiores e $cl(K) \subset X$ então podemos afirmar que K é a interseção de todos os semiespaços definidos nos pontos da fronteira de K em X, que contém K. A existência dos semiespaços fora demonstrada anteriormente utilizando o teorema da Hahn-Banach. Então o item ii. é equivalente a afirmação que o *epi*(f) é a intersecção dos semiespaços definidos pela funções lineares contínuas, que é, por sua vez, convexo e fechado pois satisfaz a igualdade na fronteira do conjunto. Então o conjunto *epi*(f) é convexo e fechado, então f é convexa e fechada pelas definções e lema anterior. Para a implicação contrária ver [15]. □

Definição 6.36. *Seja X um espaço de Hilbert sobre os reais. Defina $\Gamma_0(X) = \{f : X \to (-\infty, \infty] \colon f$ convexa, sci e $D(f) \neq \emptyset\}$, onde $D(f) = \{x \in X \colon f(x) < \infty\}$. A conjugada de f é definida por $f^* : X \to (-\infty, \infty]$, tal que:*

$$f^*(y) = \sup_{x \in D(f)} \{\langle y, x\rangle - f(x)\}.$$

Teorema 6.16 (Desigualdade de Young)**.** *Seja X um espaço de Hilbert e $f \in \Gamma_0(X)$. Então:*

$$\langle y, x \rangle \leq f(x) + f^*(y),$$

para todos $x \in D(f)$ e $y \in D(f^*)$.

Demonstração. Ver [7], página 11. □

Proposição 6.8. *Seja X um espaço de Hilbert e $f \in \Gamma_0(X)$. Se $f \not\equiv +\infty$. Então $f^* \not\equiv +\infty$.*

Demonstração. Ver [7], Proposição 1.10, página 12. □

Teorema 6.17 (Fenchel-Moreau)**.** *Seja X um espaço de Hilbert e $f \in \Gamma_0(X)$ e $f \not\equiv +\infty$. Então $f^{**} = f$.*

Demonstração. Ver [7], Teorema 1.11, página 13. □

Exemplo 6.24. *Seja $K \in X$ um subconjunto não vazio. Defina:*

$$I_K(x) = \begin{cases} +1, & se\ x \in K, \\ +\infty, & se\ x \notin K. \end{cases}$$

A função I_K é denominada função indicadora de K. Observe que I_K é convexa se, e somente se, K é convexo e I_K é sci se, e somente se, K for fechado. Se K for um subespaço de X então temos $(I_K)^* = I_{K\perp}$

Capítulo 7

Aplicação de Funções Convexas no Ensino Médio

As funções convexas tem larga aplicação em problemas de matemática do Ensino Médio, especialmente, em problemas olímpicos, grande parte relacionada a demonstração de desigualdades de alto grau de dificuldade. Neste capítulo, apresentaremos algumas demonstrações de problemas do Ensino Médio.

7.1 Problemas Olímpicos

Exemplo 7.1. *Sejam A, B, C os ângulos de um triângulo. Prove que:*

$$\sin A + \sin B + \sin C \leq 3\frac{\sqrt{3}}{2}$$

Demonstração. Defina $f : (0, \pi) \to (0, 1]$ como sendo $f(x) = \sin x$. Observe que f é côncava, pois $f'' \leq 0$, para todo $x \in dom f$. Outra forma de observar que f é côncava, sem utilizar a regra da derivada segunda é desenhar o gráfico de f, no domínio definido. Então, pela Desigualdade de Jensen, tem-se

$$1/3f(A) + 1/3f(B) + 1/3f(C) \leq f((A + B + C)/3) = f(\pi/3) = \frac{\sqrt{3}}{2}$$

finalizando a demonstração.

□

Exemplo 7.2. *Sejam $a, b, c \in \mathbb{R}_+^*$. Prove que:*

$$a^a b^b c^c \geq \left(\frac{a+b+c}{3}\right)^{(a+b+c)}$$

Demonstração. Defina $f : \mathbb{R}_+^* \to \mathbb{R}$ como sendo $f(x) = x \ln x$. Observe que f é convexa, pois $f''(x) = 1/x > 0$, para todo $x \in dom f$. Então, pela Desigualdade de

Jensen, tem-se

$$1/3f(a) + 1/3f(b) + 1/3f(c) \geq f((a+b+c)/3).$$

Logo,

$$a \ln a + b \ln b + c \ln c \geq 3\left(\tfrac{a+b+c}{3}\right) \ln\left(\tfrac{a+b+c}{3}\right).$$

Com isso,

$$\ln a^a b^b c^c \geq \ln\left(\frac{a+b+c}{3}\right)^{(a+b+c)},$$

concluindo a demonstração. $\square$

Exemplo 7.3. *(IMO 1995) Sejam $a, b, c \in \mathbb{R}_+^*$, tal que $abc = 1$. Prove que:*

$$\frac{1}{a^3(b+c)} + \frac{1}{b^3(c+a)} + \frac{1}{c^3(a+b)} \geq \frac{3}{2}$$

Demonstração. Sejam

$$x = 1/a,\ y = 1/b \text{ e } z = 1/c,$$

então $xyz = 1/(abc) = 1$. Defina $f : \mathbb{R}_+^* \to \mathbb{R}$ como sendo $f(x) = 1/x$. Observe que f é convexa. Além disso, temos que

$$1/(a^3(b+c)) = x^3yz/(y+z) = x^2/(y+z).$$

Daí,

$$\frac{x^2}{(y+z)} + \frac{y^2}{(z+x)} + \frac{z^2}{(x+y)} = xf(\frac{y+z}{x}) + yf(\frac{z+x}{y}) + zf(\frac{x+y}{z}).$$

Pela desigualdade de Jensen tem-se:

$$xf(\frac{y+z}{x}) + yf(\frac{z+x}{y}) + zf(\frac{x+y}{z}) \geq (x+y+z)f(\frac{y+z+z+x+x+y}{x+y+z}).$$

O valor do lado direito é exatamente $(x+y+z)f(2)$. Utilizando a desigualdade entre médias aritméticas e geométricas, tem-se $(x+y+z) \geq 3(xyz)^{1/3} = 3$. Por definição $f(2) = 1/2$, concluindo o problema. $\square$

Exemplo 7.4. *(IMO 2001) Sejam $a, b, c \in \mathbb{R}_+^*$. Prove que:*

$$\frac{a}{\sqrt{a^2+8bc}} + \frac{b}{\sqrt{b^2+8ca}} + \frac{c}{\sqrt{c^2+8ab}} \geq 1$$

Demonstração. Defina $f : \mathbb{R}^3 \to \mathbb{R}$ como sendo $f(x, y, z) = \sum_{cycl} \frac{x}{\sqrt{x^2+8yz}}$. Observe que dado $t \in \mathbb{R}^*$, tem-se $f(tx, ty, tz) = f(x, y, z)$. Então a função é homogênea, portanto podemos assumir, sem perda de generalidade, que $a + b + c = 1$. Agora considere a função convexa $g : \mathbb{R}_+^* \to \mathbb{R}$ tal que $g(x) = \frac{1}{\sqrt{x}}$. Pela Desigualdade de Jensen tem-se

$$ag(a^2 + 8bc) + bg(b^2 + 8ca) + cg(c^2 + 8ab) \geq g(a^3 + b^3 + c^3 + 24abc)$$

Ou seja:

$$\frac{a}{\sqrt{a^2 + 8bc}} + \frac{b}{\sqrt{b^2 + 8ca}} + \frac{c}{\sqrt{c^2 + 8ab}} \geq \frac{1}{\sqrt{a^3 + b^3 + c^3 + 24abc}}$$

Mas veja que: $1 = (a+b+c)^3 = a^3+b^3+c^3+6abc+3(a^2b+a^2c+b^2a+b^2c+c^2a+c^2b) \geq a^3 + b^3 + c^3 + 6abc + 3 \times 6(a^6b^6c^6)^{1/6} = a^3 + b^3 + c^3 + 24abc$, utilizando a desigualdade das médias aritmética e geométrica, o que finaliza o problema. $\square$

Capítulo 8

Conclusão

Neste trabalho, apresentamos os conceitos de análise convexa, com ênfase nas transformadas de Legendre e Fenchel para os casos no $\mathbb{R}^n$, bem como a expansão desses conceitos para casos mais gerais, em espaços vetoriais de dimensão infinita. Nos casos mais gerais, observamos que alguns conceitos em caso de dimensões finitas não se aplica para caso de dimensões infinitas. Por exemplo, a bola unitária em espaços topológicos métricos de dimensão infinita não é compacta. Expandimos o nosso entendimento de convergência, de sequências e definimos espaços completos, em que toda sequência de Cauchy é convergente. Ainda, vimos que nos casos de espaços vetoriais e funções lineares, alguns teoremas que envolvem a análise convexa em dimensão finita podem ser expandidas para dimensão infinita, a exemplo do Teorema da Separação por Hiperplanos. Para a expansão desses conceitos apresentamos algumas definições e utilizamos conceitos da teoria de conjuntos como o Lema de Zorn para demonstrar o teorema de Hahn-Banach, resultado clássico em análise convexa em conjuntos de dimensão infinita.

Apresentamos uma aplicação da Transformada de Legendre utilizado na mecânica clássica e também diversas propriedades importantes da análise convexa, como por exemplo, a continuidade das funções, o comportamento das mesmas em conjuntos convexos, com respeito a apresentação de mínimos ou máximo de funções. Vimos também, algumas aplicações de conceitos de funções convexas na demonstração de desigualdade clássicas e resoluções de problemas de alto nível de olimpíadas do Ensino Médio.

Futuros trabalhos neste tema envolve o desenvolvimento de teorias e aplicações dos conceitos desenvolvidos na análise convexa para resolução de problemas de otimização ou até mesmo no cálculo de variações bem como no desenvolvimento de novas técnicas ou conceitos em matemática pura e aplicada.

Referências Bibliográficas

[1] ACKER, Felipe; DICKSTEIN, Flávio.*Uma Introdução à Análise Convexa.* Rio de Janeiro: IMPA, 1983.

[2] ALVES, Claudianor Oliveira; CARRIÃO, Paulo César; MIYAGAKI, Olímpio Hiroshi. *On the existence of positive solutions of a perturbed Hamiltonian system in* $\mathbb{R}^N$. J. Math. Anal. Appl. 276 (2002) 673-690, 2002.

[3] AMORIM, Ronan Gomes de.*Introdução à Análise Convexa: conjuntos e funções convexas.* 2013. 79 f.Dissertação (Mestrado) - Curso de Matemática, Universidade Federal de Goias, Goiânia, 2013.

[4] BARTLE, Robert Gardener; SHERBERT, Donald. *Introduction to real analysis.* New Jersey: John Wiley & Sons, 2011.

[5] BERTSEKAS, Dimitri. *6.253 Convex Analysis and Optimization.* Spring 2012. Massachusetts Institute of Technology: MIT OpenCourseWare, https://ocw.mit.edu. License: Creative Commons BY-NC-SA.

[6] BOYD, Stephen; VANDENBERGHE, Lieven.*Convex Optimization.* New York: Cambridge University Press, 2004.

[7] BREZIS, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* New York: Springer, New York, 2011.

[8] *Functional Analysis Course.* IMPA 2019. Canal Instituto de Matemática Pura e Aplicada. Consulta em novembro de 2020.

[9] GLICKSMAN, Martin Eden. *Principles of Solidification.* New York: Springer. 2011.

[10] LIMA, Elon Lages. *Análise real v.1.* Rio de Janeiro: IMPA, 2018.

[11] LIMA, Elon Lages. *Análise real v.2.* Rio de Janeiro: IMPA, 2016.

[12] LIMA, Elon Lages. *Álgebra Linear* .Rio de Janeiro: IMPA, 2018.

[13] LEWIS, Adrian.*Convex Analysis Notes*. Ithaca: Cornell University OIRE 6328, 2015.

[14] MARSDEN, Jerrold E., RATIU, Tudor. *Manifold, Tensor Analysis, and Applications*. Santa Cruz-CA: Springer-Verlag Publishing Company. American Mathematics Society, 2001.

[15] MAWHIN, Jean; WILLEM, Michel. *Critical Point Theory and Hamiltonian Systems*. New York: Springer-Verlag, 1989.

[16] ROCKAFELLAR, Tyrrell. *Convex Analysis*. New Jersey: Princeton University Press, 1970.