

UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
MESTRADO PROFISSIONALIZANTE EM MATEMÁTICA

*FUNÇÕES CONVEXAS E O PROBLEMA DE STEINER PARA TRÊS
PONTOS*

Dário Souza Rocha

MANAUS

2020

UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
PROGRAMA DE MESTRADO PROFISSIONALIZANTE EM MATEMÁTICA

Dário Souza Rocha

*FUNÇÕES CONVEXAS E O PROBLEMA DE STEINER PARA TRÊS
PONTOS*

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Disney Douglas de Lima Oliveira

MANAUS
2015

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

R672f Rocha, Dário Souza
Funções Convexas e o Problema de Steiner para Três Pontos /
Dário Souza Rocha. 2020
69 f.: il. color; 31 cm.

Orientadora: Disney Douglas Lima de Oliveira
Dissertação (Mestrado Profissional em Matemática em Rede
Nacional) - Universidade Federal do Amazonas.

1. Subdiferencial. 2. Função Convexa. 3. Subgradiente. 4.
Problema de Steiner. I. Oliveira, Disney Douglas Lima de II.
Universidade Federal do Amazonas III. Título

DÁRIO SOUZA ROCHA

FUNÇÕES CONVEXAS E O PROBLEMA DE STEINER PARA TRÊS
PONTOS

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Matemática.

Aprovado em 06 de Março de 2020.

BANCA EXAMINADORA



Prof. Dr. Disney Douglas de Lima Oliveira

Presidente



Prof. Dr. Nilomar Vieira de Oliveira

Membro



Prof. Dra. Jeanne Moreira de Sousa

Membro

AGRADECIMENTOS

Agradeço inicialmente a DEUS por ter me concedido a saúde, sabedoria e perseverança necessários para que eu conduzisse minhas escolhas para ter êxito neste projeto.

Agradeço à minha mãe Darcy cujo amor e anos de labuta diária para que eu tivesse uma boa educação e instrução me proporcionaram meios para que eu chegasse até aqui. Agraço aos meus filhos Darcy maria, Dário Filho, Heitor, Davi e Dário César pela compreensão por tantas vezes que me fiz ausente para estudar.

Agradeço aos meus amigos pelo incentivo e pelas cobranças para que eu concluísse esta etapa da minha vida.

Agradeço aos companheiros que estiveram comigo durante o curso pela amizade, pelas brincadeiras e pelas palavras de incentivos.

Agradeço aos professores, tutores e coordenadores do Profmat/UFAM que nos ajudaram ao longo do curso, ao meu orientador, Professor Disney Douglas de Oliveira Lima, pelo apoio e por tomar a orientação nestes últimos meses. Ao Professor Alfredo Wagner Martins pelo desprendimento e atenção com que sempre me tratou e que foi meu orientador até o seu falecimento, os meus mais sinceros agradecimentos. E fim a todos vocês, meus sinceros agradecimentos.

RESUMO

Este trabalho objetiva apresentar aplicações das teorias matemáticas (definições, teoremas e propriedades) como ferramenta para enriquecer o ensino-aprendizagem no ambiente escolar. Como sugestão temos aqui a apresentação e desenvolvimento da teoria sobre funções convexas e seus principais resultados (otimização convexa). Procuramos enfatizar tópicos como o corpo dos complexos, onde apresentamos a forma polar através da fórmula de Euler, tópicos de álgebra linear e análise real, conjuntos convexos e um breve histórico de seu estudo, funções convexas com ênfase à ideia de subdiferencial, como principal pré-requisito para o estudo de otimização. Resolvemos como aplicação desta teoria o problema de Steiner para três pontos do plano, que tem uso principalmente em tecnologias de comunicação dentre outros.

Palavras-chave: Ensino-aprendizagem, Conjunto Convexo, Função Convexa, Subdiferencial, Otimização.

ABSTRACT

This paper aims to present applications of mathematical theories (definitions, theorems and properties) as a tool to enrich teaching-learning in the school environment. As a suggestion we have here the presentation and development of the theory about convex functions and their main results (convex optimization). We seek to emphasize topics such as the body of complexes, where we present the polar form through Euler's formula; topics of linear algebra and real analysis; convex sets and a brief history of their study; convex functions with emphasis on the sub-differential idea, as the main requirement for the optimization study. As an application of this theory, we solved Steiner's problem for three points of the plane, which is mainly used in communication technologies among others.

Keywords: Teaching-learning, Convex Set, Convex Function, Subdifferential, Optimization

LISTA DE SÍMBOLOS

| | |
|------------------|--|
| \mathbb{N} | Conjunto dos números naturais. |
| \mathbb{Z} | Conjunto dos números inteiros. |
| \mathbb{Z}^+ | Conjunto dos números inteiros não negativos. |
| \mathbb{Z}_*^+ | Conjunto dos números inteiros não nulos e não negativos. |
| \mathbb{Q} | Conjunto dos números racionais. |
| \mathbb{I} | Conjunto dos números irracionais. |
| \mathbb{R} | Conjunto dos números reais. |
| \mathbb{C} | Conjunto dos números Complexos. |
| \exists | Existe. |
| \forall | Para qualquer que seja. |
| $=$ | Igual. |
| \neq | Diferente. |
| \equiv | Congruente. |
| \cong | Aproximado. |
| \sim | Semelhante. |
| $>$ | Maior. |
| $<$ | Menor. |
| \cap | Interseção. |
| \cup | União. |
| \in | Pertence. |
| \notin | Não pertence. |
| // | Paralelo. |
| \perp | Perpendicular. |
| \overline{AB} | Segmento AB. |
| AB | Medida do segmento AB. |
| \widehat{ABC} | Medida do ângulo ABC. |
| \widehat{B} | Ângulo B. |
| <i>o.p.v</i> | Opostos pelo vértice. |
| \triangle | Triângulo. |
| S_{\triangle} | Área do triângulo. |
| S_C | Área do Círculo. |
| S_Q | Área do Quadrado. |
| \square | Indica o fim de uma demonstração. |

Lista de Figuras

| | | |
|------|--|----|
| 1.1 | Representação de $z = (a, b)$ no plano de Argand-Gauss. | 4 |
| 1.2 | Representação de $z = a + bi$ e $\bar{z} = a - bi$ no plano de Argand-Gauss. | 12 |
| 1.3 | Módulo do complexo z . | 14 |
| 1.4 | A adição $z + z_1$ e o simétrico $-z$ do complexo z . | 16 |
| 1.5 | Desigualdade triangular e a distância do ponto z ao ponto z_1 . | 17 |
| 1.6 | Forma trigonométrica do complexo z . | 18 |
| 1.7 | As raízes $\omega_0 = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$ e $\omega_1 = -\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}$ no plano de Argand-Gauss. | 21 |
| 1.8 | Representação das raízes $\sqrt[3]{z}$ e $\sqrt[5]{z}$ no plano de Argand-Gauss. | 22 |
| 1.9 | Representação de a) norma euclidiana, b) norma do máximo e c) norma da soma. | 27 |
| | | |
| 2.1 | Os conjuntos D_1 e D_2 são convexos e o D_3 é não convexo | 29 |
| 2.2 | Cone convexo e não convexo respectivamente. | 31 |
| 2.3 | O cone Dual K^* é convexo para qualquer cone K . | 32 |
| 2.4 | Exemplos de fecho cônico de um conjunto. | 32 |
| 2.5 | Ilustração da definição 2.5, onde d é uma direção tangente a D em \bar{x} . No desenho, consideramos $\ d\ = 1$, portanto a distancia entre os pontos \bar{x} e $\bar{x} + td$ é t . Quando $t \rightarrow 0_+$, $dist(\bar{x} + td, D)$ é de ordem menor que t , o comprimento do passo na direção d a partir de \bar{x} . | 33 |
| 2.6 | Gráfico de uma Função convexa | 35 |
| 2.7 | A função f é convexa \Leftrightarrow O epígrafo de f é convexo. | 37 |
| 2.8 | Conjunto de nível $L_{f, \mathbb{R}}(c) = [x_1, x_2] \cup [x_3, x_4]$. | 38 |
| 2.9 | A projeção de $(0, 0)$ sobre D não é única, pois $\ (x, y)\ _S = 1$, é qualquer ponto de D . | 42 |
| 2.10 | Ilustração do epígrafo da função $f(x) = \max\{f_1(x), f_2(x)\}$ que é convexo. Assim o supremo de funções convexas é uma função convexa. | 44 |
| 2.11 | Gráficos das funções g convexa, h convexa não decrescente e pela proposição 2.7 f é não convexa. | 46 |
| 2.12 | y_1 e y_2 são subgradientes de f em x , ou seja, $f(z) \geq f(x) + \langle y_i, z - x \rangle, i = 1, 2$. | 51 |
| 2.13 | Triângulo formado pelos pontos z_1, z_2 e z_3 , destacando o ângulo interno β de z_1 . | 56 |
| 2.14 | Triângulo formado pelos pontos z_1, z_2 e z_3 , destacando os ângulo α e β formados pelos pontos z_1, z_2, z_3 e z_0 . | 58 |

Sumário

| | |
|---|-----------|
| Introdução | 1 |
| 1 Preliminares | 2 |
| 1.1 Tópicos Álgebra Linear e Topologia | 2 |
| 1.1.1 Corpo | 2 |
| 1.2 Os Complexos | 3 |
| 1.2.1 Forma algébrica dos números complexos | 8 |
| 1.2.2 Módulo de um número complexo | 13 |
| 1.2.3 Forma polar de um número complexo | 17 |
| 1.3 Espaço Vetorial | 22 |
| 1.4 Produto interno | 23 |
| 1.5 Norma | 25 |
| 1.6 Bolas | 26 |
| 2 Conjuntos Convexos | 28 |
| 2.1 História | 28 |
| 2.2 Conjuntos Convexos | 29 |
| 2.3 Operador Projeção | 40 |
| 2.4 Funções Convexas | 43 |
| 2.4.1 Propriedades fundamentais das funções convexas | 43 |
| 2.4.2 Funções convexas diferenciáveis | 45 |
| 2.4.3 Funções convexas não diferenciáveis | 50 |
| 2.5 O Problema de Steiner | 55 |
| 2.5.1 A solução é um dos pontos dados | 56 |
| 2.5.2 A solução é um ponto diferente dos pontos dados | 57 |
| Considerações Finais | 59 |

Introdução

A otimização é um tema de grande importância no mundo de hoje, principalmente para resolver problemas de ciências e engenharia. Levar esta problemática para escola é fazer o aluno estar inteirado de sua importância e para motivá-lo para aprender e a pensar a matemática.

Apresentamos neste trabalho um problema de Otimização que Jacob Steiner (1796 - 1863) propôs no início do século XIX, conforme vemos em [1], que consiste no seguinte:

"Três vilarejos A , B e C deverão ser ligados por um sistema viário de comprimento total mínimo."

Matematicamente queremos um ponto P do plano cuja soma das distâncias a três pontos dados A , B e C seja mínima. Esse problema já foi resolvido geometricamente.

Neste trabalho propomos a utilização da teoria de otimização convexa para resolvê-lo. Abordamos a teoria de números complexos, definindo e mostrando que é uma estrutura algébrica chamada corpo. Apresentamos suas principais propriedades operatórias e para a parte que trata da forma polar introduzimos a versão obtida pela Fórmula de Euler:

$$e^{i\theta} = \cos \theta + i \operatorname{sen} \theta,$$

onde $\theta \in \mathbb{R}$ e i é a unidade imaginária.

Apresentamos também resultados de álgebra linear e de análise real que serão utilizados no desenvolvimento do trabalho, com destaque para a definição das normas euclidiana, norma da soma e norma do máximo, pois para resolver o problema de Steiner utilizamos a função norma euclidiana.

No último capítulo tratamos do desenvolvimento da teoria de otimização convexa, onde expomos os principais resultados de conjuntos convexos, funções convexas diferenciáveis e não diferenciáveis. Introduzimos a ideia de subdiferencial que é um ponto importante na teoria de minimização convexa. Como exemplo de aplicações dos principais tópicos que convergiram para existência de uma solução do problema de Steiner, utilizamos a função norma euclidiana, que se mostrou adequada para existência de uma solução global para o problema e para determinar em que condições existe a solução, utilizamos a forma polar dos complexos.

Objetivo principal do trabalho é mostrar uma aplicação da teoria de otimização convexa, não nos preocupamos em provar os principais resultados utilizados visto que todas as provas dos resultados encontram-se em [2].

Capítulo 1

Preliminares

Abordaremos neste capítulo os temas preliminares necessários para o estudo das funções convexas, bem como para a solução do Problema de Steiner. Apresentaremos algumas noções importantes de álgebra linear e topologia no \mathbb{R}^n . Incluímos também o conjunto dos números complexos e suas principais propriedades. Veremos também a Fórmula de Euler que associa a exponencial complexa às funções trigonométricas.

1.1 Tópicos Álgebra Linear e Topologia

1.1.1 Corpo

Definição 1.1. *Seja \mathbb{K} um conjunto não vazio, munido das operações de adição (+) e produto (\cdot). A terna descrita e simbolizada por $(\mathbb{K}, +, \cdot)$ é denominada **corpo** se dados x, y e $z \in \mathbb{K}$ tivermos:*

$$(A) \quad x + y \in \mathbb{K};$$

$$(P) \quad x \cdot y \in \mathbb{K};$$

$$(A_1) \quad (x + y) + z = x + (y + z);$$

$$(P_1) \quad (x \cdot y) \cdot z = x \cdot (y \cdot z);$$

$$(A_2) \quad x + y = y + x;$$

$$(P_2) \quad x \cdot y = y \cdot x;$$

$$(A_3) \quad \exists \theta \in \mathbb{K} \text{ tal que } x + \theta = x;$$

$$(P_3) \quad \exists \mu \in \mathbb{K}, \text{ tal que } x \cdot \mu = x;$$

$$(A_4) \quad \forall x, \exists -x \in \mathbb{K} \text{ tal que } x + (-x) = \theta. \quad (P_4) \quad \forall x \neq 0, \exists x^{-1} \in \mathbb{K}, \text{ tal que } x \cdot x^{-1} = \mu.$$

$$(D) \quad x \cdot (y + z) = x \cdot y + x \cdot z.$$

Em álgebra linear é comum citar os elementos do corpo por escalar. O escalar θ é chamado de neutro da adição, $-x$ de simétrico de x , μ de neutro do produto e x^{-1} de inverso multiplicativo de x . Este último recebe esta nomenclatura porque é comum chamarmos o produto de

multiplicação. Observe que a definição de corpo nos permite concluir que θ , $-x$, μ e x^{-1} são únicos, com efeito:

- (i) Suponha que exista outro elemento neutro da adição em \mathbb{K} , digamos θ_1 . Assim $x + \theta_1 = x$, $\forall x \in \mathbb{K}$. Temos então que

$$\theta_1 = \theta_1 + \theta = \theta$$

- (ii) Suponha que dado $x \in \mathbb{K}$ exista $y \in \mathbb{K}$, que também é simétrico de x , ou seja $x + y = \theta$. Temos então que

$$y = y + \theta = y + (x + (-x)) = (y + x) + (-x) = \theta + (-x) = -x$$

- (iii) Procedendo como no item (i), mas para o neutro do produto, existe $\mu_1 \in (\mathbb{K})$ tal que $x \cdot \mu_1 = x$, $\forall x \in \mathbb{K}$. Assim

$$\mu_1 = \mu_1 \cdot \mu = \mu$$

- (iv) Suponha que dado $x \in \mathbb{K}$, $x \neq 0$, exista $y \in \mathbb{K}$, que também é inverso multiplicativo de x , ou seja $x \cdot y = e$. Temos então que

$$y = y \cdot \mu = y \cdot (x \cdot x^{-1}) = (y \cdot x) \cdot x^{-1} = \mu \cdot x^{-1} = x^{-1}.$$

Sabemos que o conjunto dos números racionais \mathbb{Q} e o dos números reais \mathbb{R} com suas operações usuais de adição e produto são corpos. Já o conjunto dos naturais \mathbb{N} e o dos números inteiros \mathbb{Z} não são corpos. Definiremos e mostraremos na próxima seção que o conjunto dos números complexos \mathbb{C} é um corpo.

1.2 Os Complexos

Definição 1.2. Definimos o conjunto dos números complexos como sendo o conjunto de todos os pares ordenados $z = (a, b)$, com $a, b \in \mathbb{R}$, e representaremos por \mathbb{C} :

$$\mathbb{C} = \{(a, b); a, b \in \mathbb{R}\}$$

Definir os complexos desta forma nos permite associar a cada complexo um ponto em um plano como o cartesiano que aqui chamaremos de Argand-Gauss, conforme vemos na figura [1.1](#). O eixo horizontal é chamado de real e o eixo vertical é chamado de eixo imaginário e a interseção entre os eixos de origem O do plano.

Sabemos ainda que dois pares ordenados $z = (a, b)$ e $z_1 = (a_1, b_1)$ são iguais se suas respectivas coordenadas também o são, ou seja, $a = a_1$ e $b = b_1$.

Vamos definir as operações e verificar propriedades que tornam o conjunto dos números complexos um corpo.

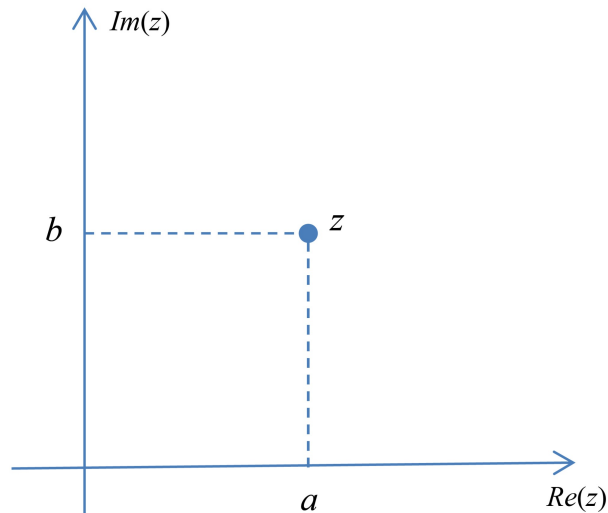


Figura 1.1: Representação de $z = (a, b)$ no plano de Argand-Gauss.

Definição 1.3. Dados $z = (a, b) \in \mathbb{C}$ e $z_1 = (a_1, b_1) \in \mathbb{C}$ utilizaremos $+$ e \cdot para simbolizar as operações de adição e produto, respectivamente, entre os complexos $z = (a, b)$ e $z_1 = (a_1, b_1)$ definidas como segue:

(i) $z + z_1 = (a + a_1, b + b_1)$

(ii) $z \cdot z_1 = (aa_1 - bb_1, ab_1 + ba_1)$

Veja que as operações definidas desta forma nos levam a concluir que tanto a adição como o produto são operações fechadas em \mathbb{C} , pois estamos adicionando e fazendo o produto entre números reais (as coordenadas de cada complexo) que é um corpo. A proposição a seguir nos garante que \mathbb{C} também é um corpo.

Proposição 1.1. As operações de adição e produto entre complexos é fechada. Para $z, z_1, z_2 \in \mathbb{C}$, quaisquer, são válidas as seguintes propriedades:

Na adição:

(A1) *Associatividade:* $z + (z_1 + z_2) = (z + z_1) + z_2$;

(A2) *Comutativa:* $z + z_1 = z_1 + z$;

(A3) *Elemento neutro:* $\exists \theta \in \mathbb{C}$ tal que $z + \theta = \theta + z = z$;

(A4) *Inverso aditivo (simétrico):* $\forall z \in \mathbb{C}, \exists -z \in \mathbb{C}$ tal que $z + (-z) = -z + z = \theta$.

No Produto:

(P1) *Associatividade:* $z \cdot (z_1 \cdot z_2) = (z \cdot z_1) \cdot z_2$;

(P2) Comutativa: $z \cdot z_1 = z_1 \cdot z$;

(P3) Elemento neutro: $\exists \mu \in \mathbb{C}$ tal que $z \cdot \mu = \mu \cdot z = z$;

(P4) Inverso: $\forall z \in \mathbb{C}, z \neq \theta, \exists z^{-1} \in \mathbb{C}$ tal que $z \cdot z^{-1} = z^{-1} \cdot z = \mu$;

(D) Distributiva: $z \cdot (z_1 + z_2) = z \cdot z_1 + z \cdot z_2$.

Demonstração. Para quaisquer $z, z_1, z_2 \in \mathbb{C}$ existem $a, a_1, a_2, b, b_1, b_2 \in \mathbb{R}$ tais que $z = (a, b)$, $z_1 = (a_1, b_1)$ e $z_2 = (a_2, b_2)$. Assim temos que $z + z_1 = (a + a_1, b + b_1) \in \mathbb{C}$, pois $a + a_1, b + b_1 \in \mathbb{R}$ e $z \cdot z_1 = (aa_1 - bb_1, ab_1 + a_1b) \in \mathbb{C}$, pois $aa_1 - bb_1, ab_1 + a_1b \in \mathbb{R}$, ou seja, as operações de adição e produto em \mathbb{C} são fechadas.

Vamos mostrar as propriedades da adição e do produto entre complexos.

(A1) Associativa: $z + (z_1 + z_2) = (z + z_1) + z_2$.

$$\begin{aligned}z + (z_1 + z_2) &= (a, b) + [(a_1, b_1) + (a_2, b_2)] \\ &= (a, b) + (a_1 + a_2, b_1 + b_2) \\ &= (a + (a_1 + a_2), b + (b_1 + b_2)) \\ &= ((a + a_1) + a_2, (b + b_1) + b_2) \\ &= (a + a_1, b + b_1) + (a_2, b_2) \\ &= [(a, b) + (a_1, b_1)] + (a_2, b_2) \\ &= (z + z_1) + z_2.\end{aligned}$$

(A2) Comutativa: $z + z_1 = z_1 + z$.

$$\begin{aligned}z + z_1 &= (a, b) + (a_1, b_1) \\ &= (a + a_1, b + b_1) \\ &= (a_1 + a, b_1 + b) \\ &= (a_1, b_1) + (a, b) \\ &= z_1 + z.\end{aligned}$$

(A3) Elemento neutro: $\exists \theta \in \mathbb{C}$ tal que $z + \theta = \theta + z = z$.

Como $\theta \in \mathbb{C}$ existem $x, y \in \mathbb{R}$ tal que $\theta = (x, y)$. Assim

$$z + \theta = z \Rightarrow (a, b) + (x, y) = (a, b) \Rightarrow (a + x, b + y) = (a, b).$$

Pela igualdade de pares ordenados temos $a + x = a$ e $b + y = b$, logo $x = 0$ e $y = 0$.
Portanto

$$\theta = (0, 0).$$

(A4) Inverso aditivo (simétrico): $\forall z \in \mathbb{C}, \exists -z \in \mathbb{C}$ tal que $z + (-z) = -z + z = \theta$.

Como $-z \in \mathbb{C}$ existem $x, y \in \mathbb{R}$ tal que $-z = (x, y)$. Assim

$$z + (-z) = \theta \Rightarrow (a, b) + (x, y) = (0, 0) \Rightarrow (a + x, b + y) = (0, 0).$$

Novamente pela igualdade de pares ordenados temos $a + x = 0$ e $b + y = 0$, logo $x = -a$ e $y = -b$. Portanto

$$-z = (-a, -b).$$

(P1) Associatividade: $z \cdot (z_1 \cdot z_2) = (z \cdot z_1) \cdot z_2$.

$$\begin{aligned} z \cdot (z_1 \cdot z_2) &= (a, b) \cdot [(a_1, b_1) \cdot (a_2, b_2)] \\ &= (a, b) \cdot (a_1 a_2 - b_1 b_2, a_1 b_2 + b_1 a_2) \\ &= (a(a_1 a_2 - b_1 b_2) - b(a_1 b_2 + b_1 a_2), a(a_1 b_2 + b_1 a_2) + b(a_1 a_2 - b_1 b_2)) \\ &= (aa_1 a_2 - ab_1 b_2 - ba_1 b_2 - bb_1 a_2, aa_1 b_2 + ab_1 a_2 + ba_1 a_2 - bb_1 b_2) \\ &= ((aa_1 - bb_1)a_2 - (ab_1 + ba_1)b_2, (aa_1 - bb_1)b_2 + (ab_1 + ba_1)a_2) \\ &= (aa_1 - bb_1, ab_1 + ba_1) \cdot (a_2, b_2) \\ &= [(a, b) \cdot (a_1, b_1)] \cdot (a_2, b_2) \\ &= (z \cdot z_1) \cdot z_2. \end{aligned}$$

(P2) Comutativa: $z \cdot z_1 = z_1 \cdot z$.

$$\begin{aligned} z \cdot z_1 &= (a, b) \cdot (a_1, b_1) \\ &= (aa_1 - bb_1, ab_1 + ba_1) \\ &= (a_1 a - b_1 b, a_1 b + b_1 a) \\ &= (a_1, b_1) \cdot (a, b) \\ &= z_1 \cdot z. \end{aligned}$$

(P3) Elemento neutro: $\exists \mu \in \mathbb{C}$ tal que $z \cdot \mu = \mu \cdot z = z$.

Como $\mu \in \mathbb{C}$ existem $x, y \in \mathbb{R}$ tal que $\mu = (x, y)$.

Note que para $z = \theta$ a propriedade já é válida, com efeito,

$$\theta \cdot \mu = (0, 0) \cdot (x, y) = (0 \cdot x - 0 \cdot y, 0 \cdot y + 0 \cdot x) = (0, 0) = \theta.$$

Vejamos agora o caso em que $z \neq \theta$, ou seja, $a \neq 0$ ou $b \neq 0$:

$$z \cdot \mu = z \Rightarrow (a, b) \cdot (x, y) = (a, b) \Rightarrow (ax - by, ay + bx) = (a, b)$$

pela igualdade de pares ordenados temos o sistema de duas equações e duas variáveis:

$$\begin{cases} ax - by = a \\ ay + bx = b \end{cases} .$$

Note que como a e b não são ao mesmo tempo nulos, o sistema possui solução única. Assim temos que $x = 1$ e $y = 0$. Portanto

$$\mu = (1, 0).$$

(P4) Inverso: $\forall z \in \mathbb{C}, z \neq \theta, \exists z^{-1} \in \mathbb{C}$ tal que $z \cdot z^{-1} = z^{-1} \cdot z = \mu$.

Como $z^{-1} \in \mathbb{C}$ existem $x, y \in \mathbb{R}$ tal que $z^{-1} = (x, y)$. Assim

$$z \cdot z^{-1} = \mu \Rightarrow (a, b) \cdot (x, y) = (1, 0) \Rightarrow (ax - by, ay + bx) = (1, 0)$$

novamente pela igualdade de pares ordenados temos

$$\begin{cases} ax - by = 1 \\ ay + bx = 0 \end{cases} \Rightarrow \begin{cases} x = \frac{a}{a^2 + b^2} \\ y = \frac{-b}{a^2 + b^2} \end{cases} ,$$

pois $z \neq \theta$, ou seja, $a \neq 0$ ou $b \neq 0$.

Portanto

$$z^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).$$

(D) Distributiva: $z \cdot (z_1 + z_2) = z \cdot z_1 + z \cdot z_2$.

$$\begin{aligned} z \cdot (z_1 + z_2) &= (a, b) \cdot [(a_1, b_1) + (a_2, b_2)] \\ &= (a, b) \cdot (a_1 + a_2, b_1 + b_2) \\ &= (a(a_1 + a_2) - b(b_1 + b_2), a(b_1 + b_2) + b(a_1 + a_2)) \\ &= ((aa_1 - bb_1) + (aa_2 - bb_2), (ab_1 + a_1b) + (ab_2 + a_2b)) \\ &= (aa_1 - bb_1, ab_1 + a_1b) + (aa_2 - bb_2, ab_2 + a_2b) \\ &= (a, b) \cdot (a_1, b_1) + (a, b) \cdot (a_2, b_2) \\ &= z \cdot z_1 + z \cdot z_2. \end{aligned}$$

□

O conjunto dos números complexos, \mathbb{C} , é um corpo. E antes de definirmos o que é um espaço vetorial vamos ver algumas propriedades e resultados que tornam o corpo dos complexos tão

importante.

Os itens *A4* e *P4* da proposição 1.1 nos permite definir a subtração e a divisão do complexo $z = (a, b)$ pelo complexo $z_1 = (a_1, b_1)$.

Para a subtração temos que:

$$z - z_1 = z + (-z_1) = (a - a_1, b - b_1),$$

para quaisquer $z, z_1 \in \mathbb{C}$.

E para a divisão temos que:

$$\frac{z}{z_1} = z \cdot z_1^{-1} = \left(\frac{aa_1 + bb_1}{a_1^2 + b_1^2}, \frac{a_1b - ab_1}{a_1^2 + b_1^2} \right),$$

para quaisquer $z, z_1 \in \mathbb{C}$, com $z_1 \neq (0, 0)$.

1.2.1 Forma algébrica dos números complexos

Mostraremos agora uma forma de identificar o subconjunto $A = \{(x, 0); x \in \mathbb{R}\}$ dos complexos com o conjunto dos números reais. Desta forma o complexo $(x, 0)$ será identificado pelo real x através de uma função que definiremos a seguir.

Proposição 1.2 (Função de identificação dos reais como subconjunto dos complexos). *Seja $\varphi : \mathbb{R} \rightarrow A$, onde $A = \{(x, 0); x \in \mathbb{R}\}$ e $\varphi(x) = (x, 0)$. Então φ é bijetiva e $\forall a, b \in \mathbb{R}$, temos:*

(i) $\varphi(a + b) = \varphi(a) + \varphi(b)$;

(ii) $\varphi(ab) = \varphi(a) \cdot \varphi(b)$.

Demonstração. Tomemos $u \in A$, qualquer. Pela definição de A , existe $a \in \mathbb{R}$ tal que $u = (a, 0)$. Assim temos que $\varphi(a) = (a, 0) = u$, logo φ é sobrejetiva.

Tomemos agora $a, b \in \mathbb{R}$, quaisquer, tal que $\varphi(a) = \varphi(b)$, ou seja, $(a, 0) = (b, 0)$ e pela igualdade entre pares ordenados temos $a = b$, logo φ é injetiva. Portanto φ é bijetiva.

Vamos mostrar agora que φ preserva a adição e o produto.

Sejam $a, b \in \mathbb{R}$, quaisquer. Temos que:

$$\varphi(a + b) = (a + b, 0) = (a, 0) + (b, 0) = \varphi(a) + \varphi(b)$$

e

$$\varphi(a) \cdot \varphi(b) = (a, 0) \cdot (b, 0) = (ab - 0 \cdot 0, a \cdot 0 + b \cdot 0) = (ab, 0) = \varphi(ab).$$

□

Vimos na proposição 1.2 que além de podermos identificar cada elemento do conjunto $A \subset \mathbb{C}$ com \mathbb{R} , ou seja, podemos considerar $\mathbb{R} \subset \mathbb{C}$, a função φ ainda preserva as operações de adição e produto.

Vamos denotar o complexo $(0, 1)$ por i , que chamaremos de unidade imaginária. E utilizando as propriedades que tornam os complexos um corpo temos:

$$i^2 = i \cdot i = (0, 1) \cdot (0, 1) = (0 - 1, 0) = (-1, 0) = -(1, 0) = -1,$$

ou seja,

$$i^2 = -1. \tag{1.1}$$

Façamos agora $i^0 = i \cdot i^{-1} = 1$, $i^1 = i$ e $i^n = \underbrace{i \cdot i \cdot \dots \cdot i}_{n\text{-vezes}}$, $\forall n \in \mathbb{N}$. Portanto:

$$\begin{array}{lll} i^0 = 1 & i^4 = 1 & i^8 = 1 \\ i^1 = i & i^5 = i & i^9 = i \\ i^2 = -1 & i^6 = -1 & i^{10} = -1 \\ i^3 = -i & i^7 = -i & i^{11} = -i \end{array}$$

Podemos supor que os resultados possíveis para as potências de i , com expoente em \mathbb{N} , estão no conjunto $\{1, i, -1, -i\}$, mais precisamente para todo $n \in \mathbb{N}$, existem $q, r \in \mathbb{N}$ tais que $n = 4q + r$ (*algoritmo da divisão*), com $r \in \{0, 1, 2, 3\}$. Com efeito temos que

$$i^n = i^{4q+r} = (i^4)^q \cdot i^r = 1^q \cdot i^r = 1 \cdot i^r = i^r.$$

Portanto

$$i^n = i^r,$$

onde r é o resto da divisão de n por 4, que confirma nossa suposição.

Temos então as condições necessárias para escrever o complexo $z = (a, b)$ na forma que é denominada algébrica:

$$z = (a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0) \cdot (0, 1) = a + bi.$$

portanto $z = a + bi$ é a *forma algébrica* do complexo $z = (a, b)$. O número real a é chamado de parte real e o número real b é chamado de parte imaginária do número complexo z , que indicamos por

$$Re(z) = a \quad \text{e} \quad Im(z) = b.$$

Quando $b = 0$, z é chamado de *número real* ($z \in \mathbb{R}$) e quando $a = 0$, z é chamado de *imaginário puro*.

Podemos reescrever a definição de \mathbb{C} utilizando a forma algébrica:

$$\mathbb{C} = \{a + bi; a, b \in \mathbb{R} \text{ e } i^2 = -1\}.$$

Para os complexos $z = a + bi$ e $z_1 = a_1 + b_1i$ as operações de adição e o produto podem ser reescritas como segue:

$$(i) z + z_1 = (a + a_1) + (b + b_1)i$$

$$(ii) z \cdot z_1 = (aa_1 - bb_1) + (ab_1 + ba_1)i$$

Deste modo vemos que a adição e o produto são operações fechadas em \mathbb{C} e as outras propriedades são também verificadas. Temos:

$$(i) \theta = (0, 0) = 0 + 0i = 0 \in \mathbb{R};$$

$$(ii) -z = -a - bi, \forall z = a + bi \in \mathbb{C};$$

$$(iii) z^{-1} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i$$

Definição 1.4 (Conjugado). *Seja $z = a + bi \in \mathbb{C}$. Dizemos que o conjugado de z é o complexo $\bar{z} = a - bi$.*

Esta definição nos permite observar algumas propriedades importantes dos conjugados dos números complexos que veremos a seguir.

Proposição 1.3. *Sejam $z_1, z_2 \in \mathbb{C}$, quaisquer. Então*

$$(i) \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2;$$

$$(ii) \overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2;$$

$$(iii) \overline{\bar{z}_1} = z_1;$$

$$(iv) \operatorname{Re}(\bar{z}_1 \cdot z_2) = \operatorname{Re}(z_1 \cdot \bar{z}_2) \text{ e } \operatorname{Im}(\bar{z}_1 \cdot z_2) = -\operatorname{Im}(z_1 \cdot \bar{z}_2).$$

Demonstração. Sejam $z_1, z_2 \in \mathbb{C}$, quaisquer. Então existem $a_1, a_2, b_1, b_2 \in \mathbb{R}$ com os quais podemos escrever $z_1 = a_1 + b_1i$ e $z_2 = a_2 + b_2i$. Sendo $\bar{z}_1 = a_1 - b_1i$ e $\bar{z}_2 = a_2 - b_2i$, temos:

$$(i) \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2:$$

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{(a_1 + b_1i) + (a_2 + b_2i)} \\ &= \overline{(a_1 + a_2) + (b_1 + b_2)i} \\ &= (a_1 + a_2) - (b_1 + b_2)i \\ &= (a_1 - b_1i) + (a_2 - b_2i) \\ &= \bar{z}_1 + \bar{z}_2; \end{aligned}$$

(ii) $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$:

$$\begin{aligned}\overline{z_1 \cdot z_2} &= \overline{(a_1 + b_1i) \cdot (a_2 + b_2i)} \\ &= \overline{(a_1a_2 - b_1b_2) + (a_1b_2 + a_2b_1)i} \\ &= (a_1a_2 - b_1b_2) - (a_1b_2 + a_2b_1)i \\ &= a_1a_2 + b_1ib_2i - a_1b_2i - a_2b_1i \\ &= a_1(a_2 - b_2i) - b_1i(a_2 - b_2i) \\ &= a_1\overline{z_2} - b_1i\overline{z_2} \\ &= (a_1 - b_1i)\overline{z_2} \\ &= \overline{z_1} \cdot \overline{z_2};\end{aligned}$$

(iii) $\overline{\overline{z_1}} = z_1$:

$$\overline{\overline{z_1}} = \overline{a_1 + b_1i} = \overline{a_1 - b_1i} = a_1 + b_1i = z_1;$$

(iv) $Re(\overline{z_1} \cdot z_2) = Re(z_1 \cdot \overline{z_2})$ e $Im(\overline{z_1} \cdot z_2) = -Im(z_1 \cdot \overline{z_2})$:

Vamos calcular inicialmente $\overline{z_1} \cdot z_2$ e $z_1 \cdot \overline{z_2}$:

$$\begin{aligned}\overline{z_1} \cdot z_2 &= \overline{a_1 + b_1i} \cdot (a_2 + b_2i) = (a_1 - b_1i) \cdot (a_2 + b_2i) \\ &= (a_1a_2 + b_1b_2) + (a_1b_2 - a_2b_1)i \\ z_1 \cdot \overline{z_2} &= (a_1 + b_1i) \cdot \overline{a_2 + b_2i} = (a_1 + b_1i) \cdot (a_2 - b_2i) \\ &= (a_1a_2 + b_1b_2) - (a_1b_2 - a_2b_1)i.\end{aligned}$$

Logo temos:

$$Re(\overline{z_1} \cdot z_2) = a_1a_2 + b_1b_2 = Re(z_1 \cdot \overline{z_2}),$$

e

$$Im(\overline{z_1} \cdot z_2) = a_1b_2 - a_2b_1 = -[-(a_1b_2 - a_2b_1)] = -Im(z_1 \cdot \overline{z_2}).$$

□

A figura [1.2](#) mostra a representação de $z = a + bi$ e seu conjugado $\overline{z} = a - bi$ no plano de Argand-Gaus. Vemos que z e \overline{z} são simétricos em relação ao eixo real.

Proposição 1.4. *Seja $z \in \mathbb{C}$, qualquer. Então*

(i) $Re(z) = \frac{z + \overline{z}}{2}$;

(ii) $Im(z) = \frac{z - \overline{z}}{2i}$;

(iii) $z = \overline{z} \Leftrightarrow z \in \mathbb{R}$;

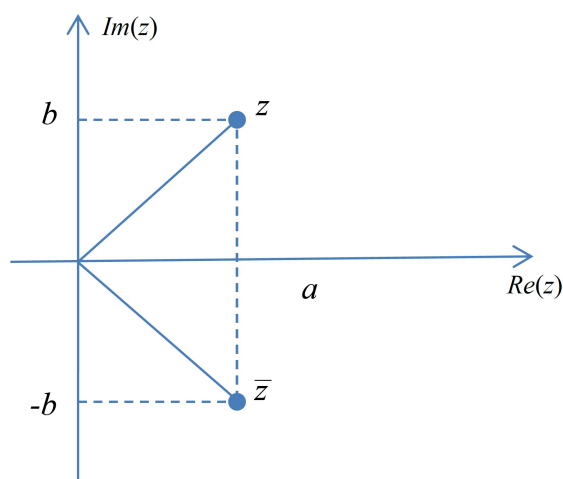


Figura 1.2: Representação de $z = a + bi$ e $\bar{z} = a - bi$ no plano de Argand-Gauss.

(iv) $z \cdot \bar{z} = a^2 + b^2$.

Demonstração. Seja $z \in \mathbb{C}$, qualquer, então existem $a, b \in \mathbb{R}$ tal que $z = a + bi$. Sendo $\bar{z} = a - bi$, temos:

(i) $Re(z) = \frac{z + \bar{z}}{2}$:

$$\frac{z + \bar{z}}{2} = \frac{(a + bi) + (a - bi)}{2} = \frac{2a}{2} = a = Re(z);$$

(ii) $Im(z) = \frac{z - \bar{z}}{2i}$:

$$\frac{z - \bar{z}}{2i} = \frac{(a + bi) - (a - bi)}{2} = \frac{2bi}{2i} = b = Im(z);$$

(iii) $z = \bar{z} \Leftrightarrow z \in \mathbb{R}$:

$$z = \bar{z} \Leftrightarrow a + bi = a - bi \Leftrightarrow b = -b \Leftrightarrow b = 0 \Leftrightarrow z \in \mathbb{R};$$

(iv) $z \cdot \bar{z} = a^2 + b^2$:

$$z \cdot \bar{z} = (a + bi) \cdot (a - bi) = a^2 - (-b^2) + (-ab + ab)i = a^2 + b^2.$$

□

Anteriormente definimos a divisão do complexo $z = (a, b)$ pelo complexo $z_1 = (a_1, b_1)$, com $a_1 \neq 0$ ou $b_1 \neq 0$:

$$\frac{z}{z_1} = z \cdot z_1^{-1} = \left(\frac{aa_1 + bb_1}{a_1^2 + b_1^2}, \frac{a_1b - ab_1}{a_1^2 + b_1^2} \right).$$

Usando a forma algébrica para z e z_1 temos:

$$\frac{z}{z_1} = z \cdot z_1^{-1} = (a + bi) \cdot \left(\frac{a_1}{a_1^2 + b_1^2} - \frac{b_1}{a_1^2 + b_1^2}i \right) = \frac{aa_1 + bb_1}{a_1^2 + b_1^2} + \frac{(a_1b - ab_1)}{a_1^2 + b_1^2}i,$$

para quaisquer $z, z_1 \in \mathbb{C}$, com $z_1 \neq 0$.

Podemos ainda ter:

$$\begin{aligned} \frac{z}{z_1} &= z \cdot z_1^{-1} \\ &= (a + bi) \cdot \left(\frac{a_1}{a_1^2 + b_1^2} - \frac{b_1}{a_1^2 + b_1^2}i \right) \\ &= \frac{1}{a_1^2 + b_1^2} \cdot [(a + bi) \cdot (a_1 - b_1i)] \\ &= \frac{1}{z_1 \cdot \bar{z}_1} \cdot (z \cdot \bar{z}_1) \\ &= \frac{z \cdot \bar{z}_1}{z_1 \cdot \bar{z}_1}. \end{aligned}$$

Portanto

$$\frac{z}{z_1} = \frac{z \cdot \bar{z}_1}{z_1 \cdot \bar{z}_1} \quad (1.2)$$

1.2.2 Módulo de um número complexo

Definição 1.5. Seja $z \in \mathbb{C}$, existem $a, b \in \mathbb{R}$, tais que $z = a + bi$. O módulo de z , $|z|$, é o número real

$$|z| = \sqrt{a^2 + b^2}$$

.

Geometricamente o módulo do complexo z é a distância *euclídeana* da origem do plano de Argand-Gauss ao ponto representado por z , conforme vemos na figura [1.3](#).

Como $z \cdot \bar{z} = a^2 + b^2$ e pela definição [1.5](#), temos:

$$|z|^2 = z \cdot \bar{z} \quad (1.3)$$

Proposição 1.5 (Propriedades do módulo de z). Seja $z \in \mathbb{C}$. Então:

- (i) $|z| \geq 0$ e $|z| = 0 \Leftrightarrow z = 0$;
- (ii) $|z| = |\bar{z}|$;
- (iii) $Re(z) \leq |Re(z)| \leq |z|$ e $Im(z) \leq |Im(z)| \leq |z|$.

Demonstração. Seja $z \in \mathbb{C}$, existem $a, b \in \mathbb{R}$ tais que $z = a + bi$ e $\bar{z} = a - bi$. Assim temos:

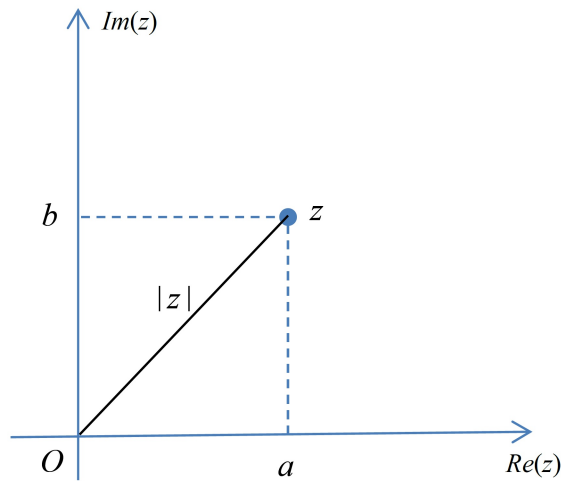


Figura 1.3: Módulo do complexo z .

(i) $a^2 \geq 0$ e $b^2 \geq 0 \Rightarrow a^2 + b^2 \geq 0 \Rightarrow |z|^2 \geq 0 \Rightarrow |z| \geq 0$. $|z| = 0 \Leftrightarrow a^2 + b^2 = 0 \Leftrightarrow a^2 = b^2 = 0 \Leftrightarrow a = b = 0 \Leftrightarrow z = 0$

(ii) $|\bar{z}| = \sqrt{a^2 + (-b)^2} = \sqrt{a^2 + b^2} = |z|$.

(iii) Como $a \in \mathbb{R}$ e $b \in \mathbb{R}$, temos:

$$\operatorname{Re}(z) = a \leq |a| = |\operatorname{Re}(z)| = \sqrt{a^2} \leq \sqrt{a^2 + b^2} = |z|,$$

e

$$\operatorname{Im}(z) = b \leq |b| = |\operatorname{Im}(z)| = \sqrt{b^2} \leq \sqrt{a^2 + b^2} = |z|.$$

□

Teorema 1.1. *Sejam $z_1, z_2 \in \mathbb{C}$, quaisquer. Então:*

(i) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$;

(ii) $\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|}$, $z_2 \neq 0$;

(iii) $|z_1 + z_2| \leq |z_1| + |z_2|$ (*desigualdade triangular*);

(iv) $|z_1 - z_2| \geq ||z_1| - |z_2||$.

Demonstração. Sejam $z_1, z_2 \in \mathbb{C}$, quaisquer. Temos que:

(i) Por 1.3 temos que

$$\begin{aligned} |z_1 \cdot z_2|^2 &= (z_1 \cdot z_2) \cdot \overline{(z_1 \cdot z_2)} \\ &= (z_1 \cdot z_2) \cdot (\bar{z}_1 \cdot \bar{z}_2) \\ &= (z_1 \cdot \bar{z}_1) \cdot (z_2 \cdot \bar{z}_2) \\ &= |z_1|^2 \cdot |z_2|^2. \end{aligned}$$

Portanto

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|$$

(ii) Para $z_2 \neq 0$ e por [1.2](#):

$$z_2^{-1} = 1 \cdot z_2^{-1} = \frac{1}{z_2} = \frac{1 \cdot \overline{z_2}}{z_2 \cdot \overline{z_2}} = \frac{\overline{z_2}}{|z_2|^2}.$$

Assim,

$$|z_2^{-1}| = \frac{|\overline{z_2}|}{|z_2|^2} = \frac{|z_2|}{|z_2|^2} = \frac{1}{|z_2|}.$$

Agora temos que

$$\left| \frac{z_1}{z_2} \right| = |z_1 \cdot z_2^{-1}| = |z_1| \cdot |z_2^{-1}| = |z_1| \cdot \frac{1}{|z_2|} = \frac{|z_1|}{|z_2|}.$$

Portanto

$$\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|}.$$

(iii)

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2) \cdot \overline{(z_1 + z_2)} \\ &= (z_1 + z_2) \cdot (\overline{z_1} + \overline{z_2}) \\ &= z_1 \cdot \overline{z_1} + z_1 \cdot \overline{z_2} + z_2 \cdot \overline{z_1} + z_2 \cdot \overline{z_2} \\ &= |z_1|^2 + z_1 \cdot \overline{z_2} + \overline{\overline{z_2} \cdot z_1} + |z_2|^2 \\ &= |z_1|^2 + 2\operatorname{Re}(z_1 \cdot \overline{z_2}) + |z_2|^2 \\ &\leq |z_1|^2 + 2|z_1 \cdot \overline{z_2}| + |z_2|^2 \\ &= |z_1|^2 + 2|z_1| \cdot |\overline{z_2}| + |z_2|^2 \\ &= |z_1|^2 + 2|z_1| \cdot |z_2| + |z_2|^2 \\ &= (|z_1| + |z_2|)^2. \end{aligned}$$

Portanto

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

(iv)

$$\begin{aligned} |z_1 - z_2|^2 &= (z_1 - z_2) \cdot \overline{(z_1 - z_2)} \\ &= (z_1 - z_2) \cdot (\overline{z_1} - \overline{z_2}) \\ &= z_1 \cdot \overline{z_1} - z_1 \cdot \overline{z_2} - z_2 \cdot \overline{z_1} + z_2 \cdot \overline{z_2} \\ &= |z_1|^2 - z_1 \cdot \overline{z_2} - \overline{z_2} \cdot z_1 + |z_2|^2 \\ &= |z_1|^2 - 2\operatorname{Re}(z_1 \cdot \overline{z_2}) + |z_2|^2 \\ &\geq |z_1|^2 - 2|z_1 \cdot \overline{z_2}| + |z_2|^2 \\ &= |z_1|^2 - 2|z_1| \cdot |\overline{z_2}| + |z_2|^2 \\ &= |z_1|^2 - 2|z_1| \cdot |z_2| + |z_2|^2 \\ &= (|z_1| - |z_2|)^2 \\ &= ||z_1| - |z_2||^2 \end{aligned}$$

Portanto

$$|z_1 - z_2| \geq ||z_1| - |z_2||.$$

□

Vimos até aqui como realizar operações com números complexos. O plano complexo torna-se mais útil e nos dá uma compreensão ainda melhor das operações aritméticas com números complexos, bem como do módulo de um número complexo. A figura 1.4 ilustra a interpretação geométrica da adição do complexo z com o complexo z_1 e o simétrico (inverso aditivo) do complexo z .

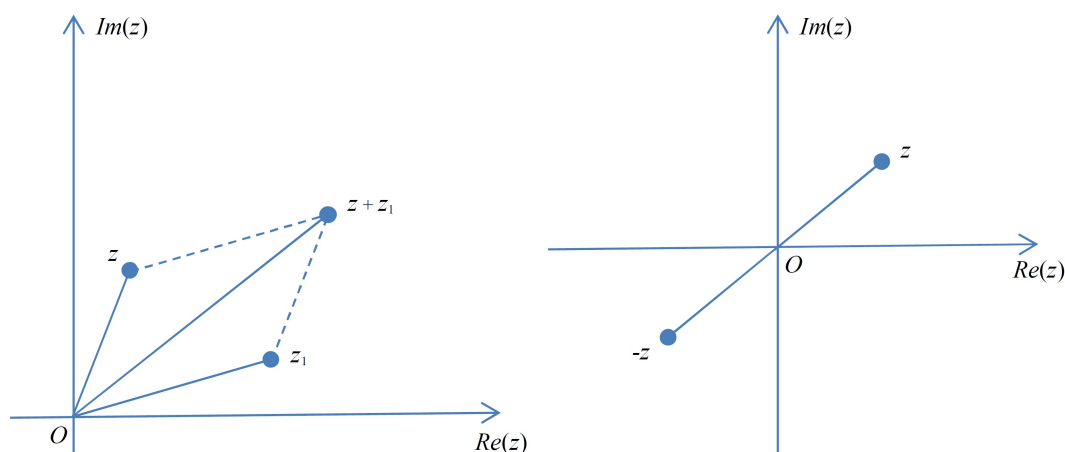


Figura 1.4: A adição $z + z_1$ e o simétrico $-z$ do complexo z .

A figura 1.5 mostra a desigualdade triangular e a distância do número complexo z (ponto z) ao complexo z_1 (ponto z_1).

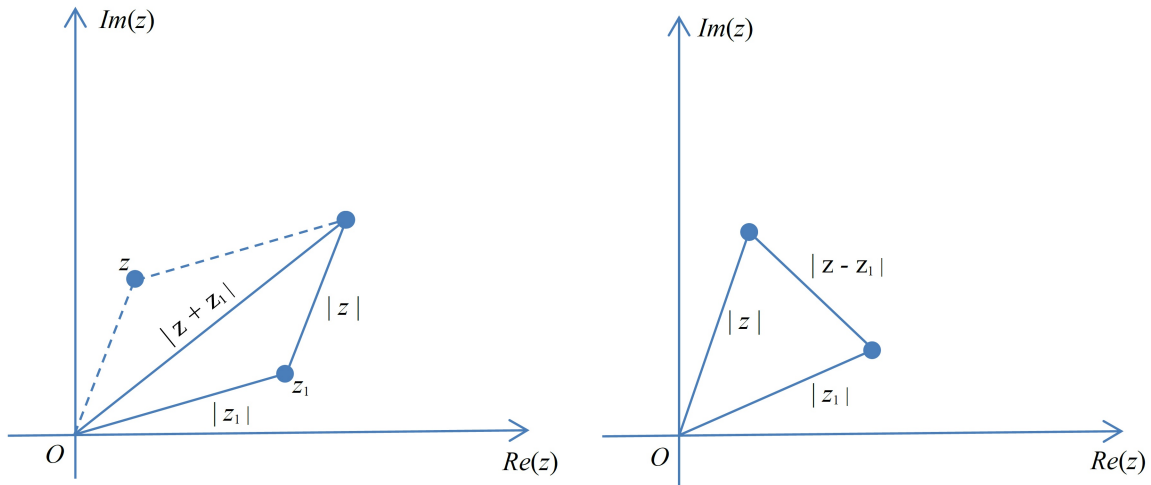


Figura 1.5: Desigualdade triangular e a distância do ponto z ao ponto z_1 .

A próxima seção nos dá uma outra maneira de determinamos no plano de Argand-Gauss o complexo $z = a + bi$.

1.2.3 Forma polar de um número complexo

Apresentaremos agora a forma polar do número complexo. A forma polar nos permite, além de facilitar os cálculos de produtos com muitos fatores complexos, nos permite calcular raízes de números complexo. Apresentaremos ainda a forma polar através da exponencial complexa.

Definição 1.6. *Seja $z = a + bi \in \mathbb{C}$, qualquer. A forma polar do complexo z é dada por*

$$z = r(\cos\theta + i\sin\theta),$$

onde $r = \sqrt{a^2 + b^2}$ e $\theta = \operatorname{arg}z = \arctan \frac{b}{a}$ (*argumento de z*)

Geometricamente r é a distância da origem ao complexo z (ponto z), $r = |z|$ e θ é chamado de argumento de z , e θ é o ângulo medido do semi-eixo real positivo ao segmento Oz , onde O é a origem do plano de Argand-Gauss, conforme vemos na figura [1.6](#).

Relação de Euler

Vamos agora definir a exponencial com expoente imaginário puro, $e^{\theta i}$, $\theta \in \mathbb{R}$. Vamos usar para atingir tal objetivo a série de Taylor. As séries de Taylor para as funções seno, cosseno e exponencial são:

$$\operatorname{sen} t = \frac{t}{1!} - \frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} + \frac{t^9}{9!} + \dots + (-1)^n \frac{t^{2n+1}}{(2n+1)!} + \dots \quad (1.4)$$

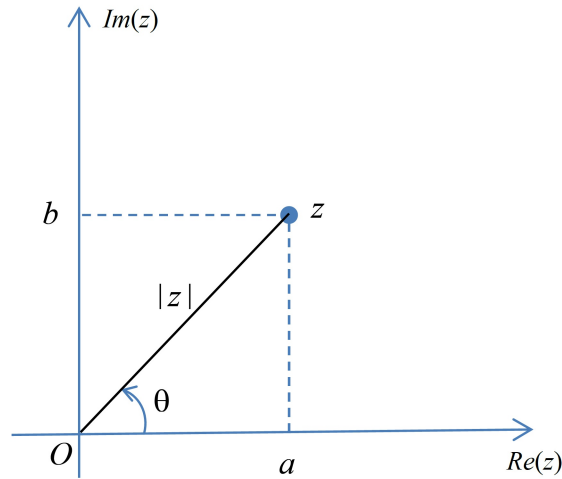


Figura 1.6: Forma trigonométrica do complexo z .

$$\cos t = 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \frac{t^6}{6!} + \frac{t^8}{8!} + \dots + (-1)^n \frac{t^{2n}}{(2n)!} + \dots \quad (1.5)$$

$$e^t = 1 + \frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \frac{t^5}{5!} + \frac{t^6}{6!} + \frac{t^7}{7!} + \frac{t^8}{8!} + \frac{t^9}{9!} + \dots + \frac{t^n}{n!} + \dots \quad (1.6)$$

Estas séries estão definidas para qualquer $t \in \mathbb{R}$ como vemos em [6]. Fazendo $t = \theta i$, para qualquer $\theta \in \mathbb{R}$, e substituindo em [1.6], temos:

$$e^{\theta i} = 1 + \frac{\theta i}{1!} + \frac{(\theta i)^2}{2!} + \frac{(\theta i)^3}{3!} + \frac{(\theta i)^4}{4!} + \frac{(\theta i)^5}{5!} + \frac{(\theta i)^6}{6!} + \frac{(\theta i)^7}{7!} + \frac{(\theta i)^8}{8!} + \frac{(\theta i)^9}{9!} + \dots$$

$$e^{\theta i} = 1 + \frac{\theta i}{1!} - \frac{\theta^2}{2!} - \frac{\theta^3 i}{3!} + \frac{\theta^4}{4!} + \frac{\theta^5 i}{5!} - \frac{\theta^6}{6!} - \frac{\theta^7 i}{7!} + \frac{\theta^8}{8!} + \frac{\theta^9 i}{9!} - \dots$$

Reorganizando e colocando i em evidência, temos:

$$e^{\theta i} = \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \frac{\theta^8}{8!} - \dots \right) + i \left(\frac{\theta}{1!} - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \frac{\theta^9}{9!} - \dots \right)$$

Por [1.4] e [1.5], temos a **Relação de Euler**:

$$e^{\theta i} = \cos \theta + i \operatorname{sen} \theta. \quad (1.7)$$

E para $-\theta$:

$$e^{-\theta i} = \cos(-\theta) + i \operatorname{sen}(-\theta) = \cos \theta - i \operatorname{sen} \theta$$

pois como sabemos o cosseno é função par e o seno é função ímpar, e também periódicas, veja [6].

Pelo fato de serem periódicas, o seno e o cosseno, observamos que:

$$\begin{aligned}
 e^{(\theta-2\pi)i} &= \cos(\theta - 2\pi) + i\text{sen}(\theta - 2\pi) = \cos \theta + i\text{sen} \theta = e^{\theta i} \\
 e^{(\theta+2\pi)i} &= \cos(\theta + 2\pi) + i\text{sen}(\theta + 2\pi) = \cos \theta + i\text{sen} \theta = e^{\theta i} \\
 e^{(\theta+4\pi)i} &= \cos(\theta + 4\pi) + i\text{sen}(\theta + 4\pi) = \cos \theta + i\text{sen} \theta = e^{\theta i} \\
 &\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\
 e^{(\theta+n\pi)i} &= \cos(\theta + n\pi) + i\text{sen}(\theta + n\pi) = \cos \theta + i\text{sen} \theta = e^{\theta i},
 \end{aligned}$$

com $n \in \mathbb{Z}$. Portanto a função exponencial com expoente imaginário é periódica de período 2π .

Logo a exponencial com expoente imaginário herda as propriedades da periodicidade do seno e do cosseno, o período p de $e^{(a\theta)i}$, com $a, \theta \in \mathbb{R}$, $a > 0$ é dado por

$$p = \frac{2\pi}{a}.$$

Pela definição [1.6](#), o complexo $z = a + bi$ é dado por $z = |z|(\cos\theta + i\text{sen} \theta)$, com $\theta \in \mathbb{R}$, argumento de z . Pela relação de Euler podemos reescrever a forma trigonométrica da seguinte forma:

$$z = |z|e^{\theta i}, \text{ com } \theta \in \mathbb{R}. \tag{1.8}$$

Vamos agora definir o produto entre dois complexos na forma trigonométrica.

Sejam $z = a + bi$, $z_1 = a_1 + b_1i \in \mathbb{C}$. Então existem $\theta, \theta_1 \in \mathbb{R}$ tais que

$$z = |z|e^{\theta i} \quad \text{e} \quad z_1 = |z_1|e^{\theta_1 i}.$$

O produto entre z e z_1 é dado por

$$z \cdot z_1 = (|z|e^{\theta i})(|z_1|e^{\theta_1 i}) = |z||z_1|e^{\theta i}e^{\theta_1 i} = |z||z_1|e^{(\theta+\theta_1)i}.$$

Portanto

$$z \cdot z_1 = |z||z_1|e^{(\theta+\theta_1)i} \tag{1.9}$$

Vamos calcular z^n , para $n \in \mathbb{N}$:

$$z^n = (|z|e^{\theta i})^n = |z|^n(e^{\theta i})^n = |z|^n e^{(n\theta)i}.$$

Portanto

$$z^n = |z|^n e^{(n\theta)i}. \tag{1.10}$$

A equação [1.10](#) é chamada de 1ª **Fórmula de Moivre**. Vamos determinar a 2ª fórmula de Moivre que trata do cálculo das raízes de números complexos.

Vamos calcular agora $\sqrt[n]{z}$, com $z \in \mathbb{C}$ qualquer e $n \in \mathbb{N}$, $n \geq 2$. Seja $\omega \in \mathbb{C}$, tal que $\omega = \sqrt[n]{z}$. Assim temos

$$\omega = \sqrt[n]{z} \Leftrightarrow \omega^n = z.$$

Como $\omega, z \in \mathbb{C}$, existem α e $\theta \in \mathbb{R}$ tais que $\omega = |\omega|e^{\alpha i}$ e $z = |z|e^{\theta i}$. Assim temos ainda que

$$\omega^n = z \Leftrightarrow |\omega|^n e^{(n\alpha)i} = |z|e^{\theta i}.$$

Então

$$|\omega|^n = |z| \Rightarrow |\omega| = \sqrt[n]{|z|}$$

e

$$n\alpha = \theta + 2k\pi \Rightarrow \alpha = \frac{\theta + 2k\pi}{n},$$

com $k \in \mathbb{Z}$.

Portanto,

$$\omega = \sqrt[n]{|z|} e^{\frac{\theta + 2k\pi}{n} i}.$$

Como para cada $k \in \mathbb{Z}$ teremos uma raiz de z , vamos escrever ω da seguinte forma

$$\sqrt[n]{z} = \omega_k = \sqrt[n]{|z|} e^{\frac{\theta + 2k\pi}{n} i}. \quad (1.11)$$

Pela periodicidade da exponencial com expoente imaginário, temos que

$$k \in \{0, 1, 2, \dots, n-1\}.$$

A expressão [1.11](#) é chamada de de 2ª **Fórmula de Moivre**.

Exemplo 1.1. Para calcular a \sqrt{i} , sabemos que $|i| = 1$ e o seu argumento é $\theta = \frac{\pi}{2}$. Logo temos que:

$$\omega_k = |i| e^{\frac{\frac{\pi}{2} + 2k\pi}{2}} = e^{\frac{\pi}{4} + k\pi}, \text{ com } k \in \{0, 1\}.$$

Assim temos:

$$\omega_0 = e^{\frac{\pi}{4}} = \frac{\sqrt{2}}{2} + i \frac{\sqrt{2}}{2} \quad e \quad \omega_1 = e^{\frac{5\pi}{4}} = -\frac{\sqrt{2}}{2} - i \frac{\sqrt{2}}{2}.$$

Observamos que no exemplo acima as raízes quadradas de i são extremidades de um diâmetro da circunferência de raio 1, como mostra a figura [1.7](#).

Observamos que para raízes cúbicas, quádras ou mais geralmente para raízes onde $n \geq 3$, de $z \neq 0$, as raízes obtidas são vértices de um polígono regular inscrito em uma circunferência de raio $\sqrt[n]{|z|}$. A figura [1.8](#) ilustra o resultado para $\sqrt[3]{z}$ e $\sqrt[5]{z}$. Vejamos:

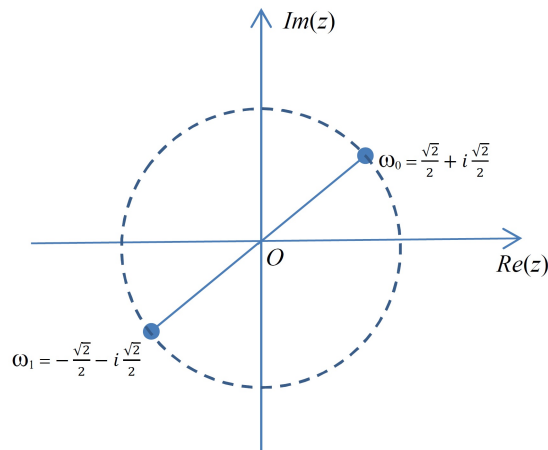


Figura 1.7: As raízes $\omega_0 = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$ e $\omega_1 = -\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}$ no plano de Argand-Gauss.

- Como vimos em [1.11](#), para $\sqrt[3]{z}$ temos:

$$\begin{aligned}\omega_0 &= \sqrt[3]{|z|} e^{\frac{\theta}{3}i} \\ \omega_1 &= \sqrt[3]{|z|} e^{\left(\frac{\theta}{3} + \frac{2\pi}{3}\right)i} \\ \omega_2 &= \sqrt[3]{|z|} e^{\left(\frac{\theta}{3} + \frac{4\pi}{3}\right)i}.\end{aligned}$$

Os argumentos das raízes cúbicas de z , ω_0 , ω_1 e ω_2 , são respectivamente $\frac{\theta}{3}$, $\frac{\theta}{3} + \frac{2\pi}{3}$ e $\frac{\theta}{3} + \frac{4\pi}{3}$. Note que o ângulo entre $O\omega_0$ e $O\omega_1$ é $\frac{2\pi}{3}$, entre $O\omega_1$ e $O\omega_2$ é $\frac{2\pi}{3}$ e entre $O\omega_2$ e $O\omega_0$ é $\frac{2\pi}{3}$ portanto as raízes cúbicas são vértices de um triângulo equilátero, como vemos na figura [1.8](#).

- Novamente como vimos em [1.11](#), para $\sqrt[5]{z}$ temos:

$$\begin{aligned}\omega_0 &= \sqrt[5]{|z|} e^{\frac{\theta}{5}i} \\ \omega_1 &= \sqrt[5]{|z|} e^{\left(\frac{\theta}{5} + \frac{2\pi}{5}\right)i} \\ \omega_2 &= \sqrt[5]{|z|} e^{\left(\frac{\theta}{5} + \frac{4\pi}{5}\right)i} \\ \omega_3 &= \sqrt[5]{|z|} e^{\left(\frac{\theta}{5} + \frac{6\pi}{5}\right)i} \\ \omega_4 &= \sqrt[5]{|z|} e^{\left(\frac{\theta}{5} + \frac{8\pi}{5}\right)i}.\end{aligned}$$

Vemos na figura [1.8](#) que os polígonos formados pelos vértices das raízes cúbicas e quinta de um complexo qualquer não nulo são vértices de triângulo equilátero e de um pentágono regular, respectivamente.

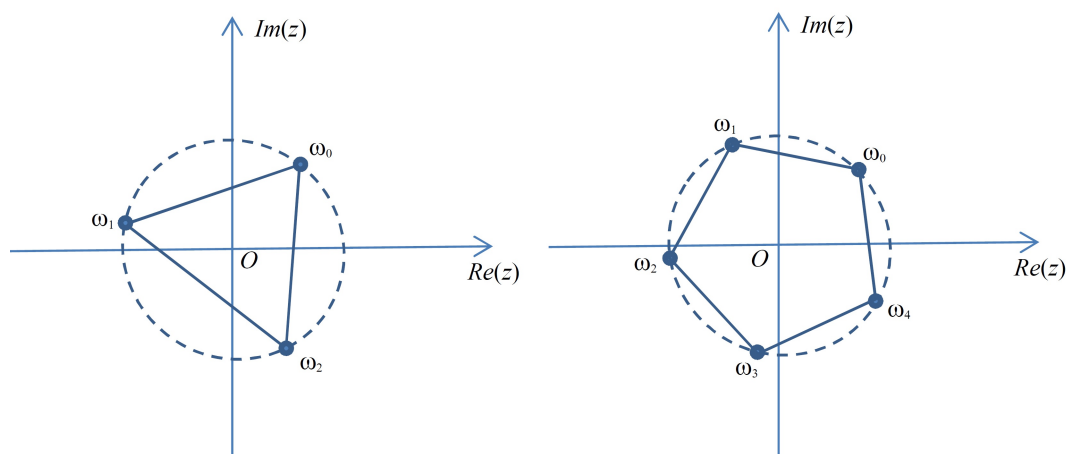


Figura 1.8: Representação das raízes $\sqrt[3]{z}$ e $\sqrt[5]{z}$ no plano de Argand-Gauss.

1.3 Espaço Vetorial

Definição 1.7. Sejam V um conjunto não vazio, \mathbb{K} um corpo, uma operação de adição (+) entre os elementos de V e uma operação de multiplicação (\cdot) entre os elementos de \mathbb{K} e os elementos de V . O conjunto V é chamado de **espaço vetorial** sobre o corpo \mathbb{K} , e simbolizamos por $V(\mathbb{K})$, quando para $u, v, w \in V$ e $\alpha, \beta \in \mathbb{K}$ temos:

$$\begin{array}{ll}
 (AE) \mathbf{u} + \mathbf{v} \in V; & (M) \alpha \cdot \mathbf{u} \in \mathbb{K}; \\
 (AE_1) (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}); & (M_1) (\alpha \cdot \beta) \cdot \mathbf{u} = \alpha \cdot (\beta \cdot \mathbf{u}); \\
 (AE_2) \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}; & (M_2) (\alpha + \beta) \cdot \mathbf{u} = \alpha \cdot \mathbf{u} + \beta \cdot \mathbf{u}; \\
 (AE_3) \exists \mathbf{0} \in V \text{ tal que } \mathbf{u} + \mathbf{0} = \mathbf{u}; & (M_3) \alpha \cdot (\mathbf{u} + \mathbf{v}) = \alpha \cdot \mathbf{u} + \alpha \cdot \mathbf{v}; \\
 (AE_4) \forall \mathbf{u}, \exists -\mathbf{u} \in V \text{ tal que } \mathbf{u} + (-\mathbf{u}) = \mathbf{0}. & (M_4) \exists \mu \in \mathbb{K}, \text{ tal que } \mu \cdot \mathbf{u} = \mathbf{u}.
 \end{array}$$

Os elementos de um espaço vetorial são chamados de vetores, não importando sua natureza. Também é comum na multiplicação ou produto omitirmos o seu símbolo, $\alpha \cdot \mathbf{u} = \alpha \mathbf{u}$.

Como exemplo de espaço vetorial temos o conjunto dos polinômios (sobre \mathbb{R}), o conjunto das matrizes (sobre \mathbb{R}) com suas operações de adição usuais e seus produtos por um escalar real.

O conjunto dos números reais sobre si mesmo é um espaço vetorial, o plano cartesiano (chamado de $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$) sobre \mathbb{R} é um espaço vetorial. Mais geralmente o conjunto

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}}_{n\text{-vezes}},$$

onde para $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, existem reais $u_1, u_2, \dots, u_n, v_1, v_2, \dots, v_n$ tais que $\mathbf{u} = (u_1, u_2, \dots, u_n)$,

$\mathbf{v} = (v_1, v_2, \dots, v_n)$ e $\alpha \in \mathbb{R}$ com a adição:

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$$

e a multiplicação:

$$\alpha \cdot \mathbf{u} = (\alpha \cdot u_1, \alpha \cdot u_2, \dots, \alpha \cdot u_n)$$

é um espaço vetorial, $\mathbb{R}^n(\mathbb{R})$, mas em geral representamos apenas por \mathbb{R}^n . Para cada $n \in \mathbb{N}$ temos um espaço vetorial diferente. Os espaços assim obtidos são chamados de espaços vetoriais reais.

Da modo análogo, ao que fizemos para os reais podemos fazer para os complexos, obtendo assim uma família denominada espaços vetoriais complexos. Mesmo para os polinômios e matrizes, se os definirmos em \mathbb{C} e como corpo também tomarmos os complexos, ainda teremos espaços vetoriais.

Trabalharemos a solução do *problema de Steiner* utilizando o espaço vetorial $\mathbb{C}(\mathbb{R}) = \mathbb{C}$ que como vimos na seção [1.2.1](#), se confunde com o \mathbb{R}^2 .

1.4 Produto interno

Definição 1.8. *Sejam V um espaço vetorial, $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ e $\alpha \in \mathbb{R}$. Definimos o produto interno entre \mathbf{u} e \mathbf{v} , e simbolizamos por $\langle \mathbf{u}, \mathbf{v} \rangle$, como sendo uma operação tal que $\langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{R}$ e além disso:*

i) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ e $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$;

ii) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$;

iii) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$;

iv) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$;

v) $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \alpha \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$;

Exemplo 1.2 (Produto interno canônico). *Sejam $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ vetores quaisquer e $\alpha \in \mathbb{R}$ um escalar qualquer. Assim $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$ e $\mathbf{z} = (z_1, z_2, \dots, z_n)$, com $x_j, y_j, z_j \in \mathbb{R}$ e $j \in \{1, 2, \dots, n\}$. Vamos definir uma operação de produto entre \mathbf{x} e \mathbf{y} como segue:*

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n \in \mathbb{R}.$$

Veja que para esta operação temos:

i) $\mathbf{x} \cdot \mathbf{x} = (x_1)^2 + (x_2)^2 + \dots + (x_n)^2 \geq 0$ e $\mathbf{x} \cdot \mathbf{x} = 0 \Leftrightarrow (x_1)^2 + (x_2)^2 + \dots + (x_n)^2 = 0 \Leftrightarrow x_1 = x_2 = \dots = x_n = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$;

ii) $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = y_1x_1 + y_2x_2 + \cdots + y_nx_n = \mathbf{y} \cdot \mathbf{x};$

iii)

$$\begin{aligned} (\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \cdot (z_1, z_2, \dots, z_n) \\ &= ((x_1 + y_1)z_1, (x_2 + y_2)z_2, \dots, (x_n + y_n)z_n) \\ &= (x_1z_1 + y_1z_1, x_2z_2 + y_2z_2, \dots, x_nz_n + y_nz_n) \\ &= (x_1z_1, x_2z_2, \dots, x_nz_n) + (y_1z_1, y_2z_2, \dots, y_nz_n) \\ &= \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}; \end{aligned}$$

iv)

$$\begin{aligned} \mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) &= (x_1, x_2, \dots, x_n) \cdot (z_1 + y_1, z_2 + y_2, \dots, z_n + y_n) \\ &= (x_1(y_1 + z_1), x_2(y_2 + z_2), \dots, x_n(y_n + z_n)) \\ &= (x_1y_1 + x_1z_1, x_2y_2 + x_2z_2, \dots, x_ny_n + x_nz_n) \\ &= (x_1y_1, x_2y_2, \dots, x_ny_n) + (x_1z_1, x_2z_2, \dots, x_nz_n) \\ &= \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}; \end{aligned}$$

v)

$$\begin{aligned} (\alpha\mathbf{x}) \cdot \mathbf{y} &= (\alpha x_1, \alpha x_2, \dots, \alpha x_n) \cdot (y_1, y_2, \dots, y_n) \\ &= ((\alpha x_1)y_1, (\alpha x_2)y_2, \dots, (\alpha x_n)y_n) \\ &= (x_1(\alpha y_1), x_2(\alpha y_2), \dots, x_n(\alpha y_n)) \\ &= (x_1, x_2, \dots, x_n) \cdot (\alpha y_1, \alpha y_2, \dots, \alpha y_n) \\ &= \mathbf{x} \cdot (\alpha\mathbf{y}) \end{aligned}$$

$$\begin{aligned} \alpha(\mathbf{x} \cdot \mathbf{y}) &= \alpha(x_1y_1 + x_2y_2 + \cdots + x_ny_n) \\ &= (\alpha(x_1y_1) + \alpha(x_2y_2) + \cdots + \alpha(x_ny_n)) \\ &= ((\alpha x_1)y_1 + (\alpha x_2)y_2 + \cdots + (\alpha x_n)y_n) \\ &= (\alpha x_1, \alpha x_2, \dots, \alpha x_n) \cdot (y_1, y_2, \dots, y_n) \\ &= (\alpha\mathbf{x}) \cdot \mathbf{y} \end{aligned}$$

Assim temos $\alpha(\mathbf{x} \cdot \mathbf{y}) = (\alpha\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\alpha\mathbf{y})$.

Assim temos que $\mathbf{x} \cdot \mathbf{y}$, com $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, define um produto interno no espaço vetorial \mathbb{R}^n e é chamado de produto interno canônico (ou usual).

1.5 Norma

Definição 1.9. Sejam V um espaço vetorial real. A função $\| \cdot \| : V \rightarrow \mathbb{R}$ que associa a cada vetor $\mathbf{u} \in V$ o número real $\|\mathbf{u}\|$, denominado **norma do vetor \mathbf{u}** , e para quaisquer $\mathbf{u}, \mathbf{v} \in V$ e $\lambda \in \mathbb{R}$ satisfaz as seguintes condições :

- i) $\|\mathbf{u}\| \geq 0$ e $\|\mathbf{u}\| = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$;
- ii) $\|\lambda \cdot \mathbf{u}\| = |\lambda| \cdot \|\mathbf{u}\|$;
- iii) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Vale salientar que a *norma* do vetor \mathbf{u} nos dá o seu comprimento. E através dela podemos introduzir algumas noções topológicas importantes nos espaços reais, como veremos mais adiante.

Exemplo 1.3. O produto interno nos dá naturalmente uma norma. Sejam V um espaço vetorial, $\mathbf{u}, \mathbf{v} \in V$ e $\alpha \in \mathbb{R}$. Definimos $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$. Vejamos:

i) Pelo item (i) da definição 1.4 temos:

- $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0 \Rightarrow \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} \geq 0 \Rightarrow \|\mathbf{u}\| \geq 0$
- $\mathbf{u} = \mathbf{0} \Leftrightarrow \langle \mathbf{u}, \mathbf{u} \rangle = 0 \Leftrightarrow \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = 0 \Leftrightarrow \|\mathbf{u}\| = 0$

ii) Pelo item (v) da definição 1.4 temos:

$$\|\alpha \mathbf{u}\| = \sqrt{\langle \alpha \mathbf{u}, \alpha \mathbf{u} \rangle} = \sqrt{\alpha^2 \langle \mathbf{u}, \mathbf{u} \rangle} = |\alpha| \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = |\alpha| \|\mathbf{u}\|$$

iii) Para mostrarmos que essa norma satisfaz o item (iii) da definição 1.5, vamos demonstrar que dados os vetores $\mathbf{u}, \mathbf{v} \in V$ temos que $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$. Veja que para $\mathbf{u} = \mathbf{0}$ a desigualdade é claramente satisfeita. Assim para $\mathbf{u} \neq \mathbf{0}$, façamos $\alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}$. Temos então para $\mathbf{w} = \mathbf{v} - \alpha \mathbf{u}$:

$$\langle \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v} - \alpha \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle - \alpha \langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \langle \mathbf{u}, \mathbf{u} \rangle = 0 \quad (1.12)$$

Para $\mathbf{v} = \mathbf{w} + \alpha \mathbf{u}$, e por 1.12 temos:

$$\langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{w} + \alpha \mathbf{u}, \mathbf{w} + \alpha \mathbf{u} \rangle = \langle \mathbf{w}, \mathbf{w} \rangle + 2\alpha \langle \mathbf{w}, \mathbf{u} \rangle + \alpha^2 \langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{w}, \mathbf{w} \rangle + \alpha^2 \langle \mathbf{u}, \mathbf{u} \rangle,$$

ou seja,

$$\langle \mathbf{v}, \mathbf{v} \rangle \geq \alpha^2 \langle \mathbf{u}, \mathbf{u} \rangle \Rightarrow \|\mathbf{v}\|^2 \geq \alpha^2 \|\mathbf{u}\|^2.$$

Por outro lado,

$$\alpha^2 \|\mathbf{u}\|^2 = \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\langle \mathbf{u}, \mathbf{u} \rangle^2} \|\mathbf{u}\|^2 = \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\|\mathbf{u}\|^4} \|\mathbf{u}\|^2 = \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\|\mathbf{u}\|^2}.$$

Assim temos que

$$\|\mathbf{v}\|^2 \geq \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\|\mathbf{u}\|^2} \Rightarrow |\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\ &\leq \langle \mathbf{u}, \mathbf{u} \rangle + 2|\langle \mathbf{u}, \mathbf{v} \rangle| + \langle \mathbf{v}, \mathbf{v} \rangle \\ &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\| \|\mathbf{v}\| + \|\mathbf{v}\|^2 \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2. \end{aligned}$$

Portanto $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

O produto interno definido no exemplo [1.2](#) nos dá a chamada norma euclidiana:

$$\begin{aligned} \|\mathbf{x}\| &= \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \\ &= \sqrt{\mathbf{x} \cdot \mathbf{x}} \\ &= \sqrt{(x_1)^2 + (x_2)^2 + \dots + (x_n)^2}, \end{aligned}$$

$\forall \mathbf{x} \in \mathbb{R}^n$.

Exemplo 1.4 (Norma do máximo e da soma). Seja $\mathbf{x} \in \mathbb{R}^n$ um vetor qualquer, ou seja, $\mathbf{x} = (x_1, x_2, \dots, x_n)$.

Definimos a norma do máximo por

$$\|\mathbf{x}\|_M = \max\{|x_1|, |x_2|, \dots, |x_n|\}.$$

E a norma da soma por

$$\|\mathbf{x}\|_S = |x_1| + |x_2| + \dots + |x_n|.$$

Que claramente satisfazem a definição [1.5](#).

1.6 Bolas

No estudo de otimização convexa é de fundamental importância a noção de bolas abertas e fechadas.

Definição 1.10. Seja $x_0 \in \mathbb{R}^n$ um ponto qualquer e $r \in \mathbb{R}_+$ um escalar qualquer. A **bola aberta** de centro x_0 e raio r é o conjunto $B(x_0, r)$ formado pelos pontos do \mathbb{R}^n cuja distância ao ponto x_0 é menor que r :

$$B(x_0, r) = \{x \in \mathbb{R}^n; \|x - x_0\| < r\}.$$

Definição 1.11. Seja $x_0 \in \mathbb{R}^n$ um ponto qualquer e $r \in \mathbb{R}_+$ um escalar qualquer. A **bola fechada** de centro x_0 e raio r é o conjunto $B[x_0, r]$ dos pontos do \mathbb{R}^n cuja distância ao ponto x_0 é menor ou igual a r :

$$B[x_0, r] = \{x \in \mathbb{R}^n; \|x - x_0\| \leq r\}.$$

Note que a diferença da bola aberta para a bola fechada, é que na última estão incluídos os pontos da fronteira (ver [4]) da primeira.

Exemplo 1.5. Em \mathbb{R}^n para $n = 1$ temos que a bola aberta de centro x_0 e raio r é o intervalo aberto $(x_0 - r, x_0 + r)$ e a bola fechada de mesmos centro e raio é o intervalo fechado $[x_0 - r, x_0 + r]$.

Exemplo 1.6. Em \mathbb{R}^n para $n = 2$ temos que a bola aberta ou fechada de centro x_0 e raio r depende da norma que tomarmos para defini-las. A figura 1.9 ilustra a bola fecha na norma euclídeana, na norma do máximo e na norma da soma.

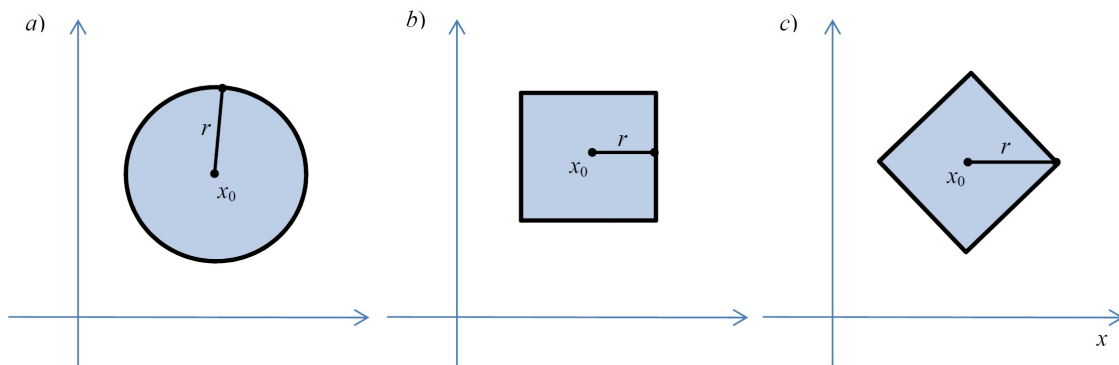


Figura 1.9: Representação de a) norma euclídeana, b) norma do máximo e c) norma da soma.

Capítulo 2

Conjuntos Convexos

Neste capítulo trataremos dos principais resultados sobre funções convexas que usaremos na solução do problema de Steiner. Toda a teoria foi desenvolvida baseada em [2].

2.1 História

O estudo da convexidade não é recente. Já em "Os Elementos" de Euclides [3] (300 a.C.) temos várias contribuições ao assunto, onde encontramos as principais propriedades de convexidade dos polígonos e poliedros. Arquimedes (287-212 a.C.), em seu livro "Sobre a Esfera e o Cilindro", foi o primeiro a definir o que se entendia por uma curva ou superfície convexa. Entre as propriedades obtidas por Arquimedes sobre convexidade, destacamos os postulados e resultados referentes ao centro de gravidade de conjuntos planos e a descrição dos 13 poliedros semirregulares, conhecidos como sólidos arquimedianos. Um poliedro convexo se diz semirregular se suas faces são polígonos regulares de pelo menos dois tipos e o grupo de isometrias é transitivo sobre os vértices. O alemão Johannes Kepler (1571 - 1630) em seu livro "Harmônicas Mundi" (1619) demonstrou que só existem 13 tipos diferentes de sólidos arquimedianos. Resultados referentes ao famoso problema isoperimétrico, cuja origem data de cerca de 810 a.C., e devido a Zenodorus (200 a. C.) parece ser o mais importante: mostra que entre todos os n -ângulos convexos de mesmo perímetro, o regular é o que possui maior área. No final do século XIX apareceram diversos resultados importantes em convexidade, graças a matemáticos como Brunn e Minkowski. A geometria convexa é relativamente recente, pois um primeiro estudo sistemático foi encontrado em 1934, no livro de Bonnesen e Fenchel "Theorie der Konvexen Körper". Nos anos 40 e 50 se descobriram várias aplicações importantes para conjuntos convexos, principalmente no campo da Otimização Geométrica, onde o interesse por esta teoria cresceu.

2.2 Conjuntos Convexos

Definição 2.1. Sejam $D \subset \mathbb{R}^n$ um conjunto. O conjunto D é chamado de convexo se para x, y pontos em D e $\alpha \in [0; 1]$ temos $\alpha x + (1 - \alpha)y \in D$. O ponto $\alpha x + (1 - \alpha)y$ é chamado combinação convexa de parâmetro α .

Esta afirmação nos diz que um conjunto é convexo quando quaisquer dois pontos de um dado conjunto podem ser conectados por uma trajetória retilínea contida inteiramente no conjunto, ou seja, os pontos que formam a trajetória são pontos do conjunto. Como consequência imediata da definição o conjunto vazio, o conjunto com um único ponto e o espaço \mathbb{R}^n são convexos e qualquer conjunto que não satisfaça a definição é chamado de não convexo.

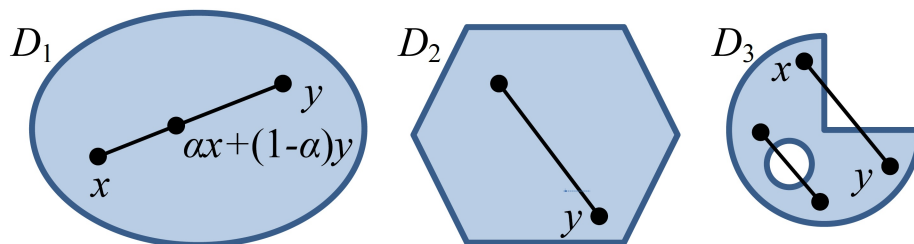


Figura 2.1: Os conjuntos D_1 e D_2 são convexos e o D_3 é não convexo

Exemplo 2.1. Toda reta no plano é um conjunto convexo, com efeito, para quaisquer dois pontos de uma reta o segmento de reta que os une está contido na mesma reta.

Exemplo 2.2. Todo subespaço S do \mathbb{R}^n é convexo. Para mostrarmos basta tomarmos $u, v \in S$ quaisquer e sabemos que $\alpha u + \beta v \in S$, para todo α e β reais. Fazendo $\alpha \in [0, 1]$ e $\beta = 1 - \alpha$ temos que S é convexo.

Exemplo 2.3. Seja $S_a = \{x \in \mathbb{R}^n; \langle a, x \rangle \leq c\}$, com $a \in \mathbb{R}^n$ e $c \in \mathbb{R}$ fixos, um semi-espaço em \mathbb{R}^n . Veja que dados $x, y \in S_a$ temos que

$$\langle a, \alpha x + (1 - \alpha)y \rangle = \langle a, \alpha x \rangle + \langle a, (1 - \alpha)y \rangle \leq \alpha c + (1 - \alpha)c = c.$$

Portanto S_a é convexo.

Exemplo 2.4. Seja $H = \{x \in \mathbb{R}^n; \langle a, x \rangle = c\}$ com $a \in \mathbb{R}^n$ e $c \in \mathbb{R}$ fixos, um hiperplano em \mathbb{R}^n , é convexo. Com efeito dados $x, y \in H$ temos que

$$\langle a, \alpha x + (1 - \alpha)y \rangle = \langle a, \alpha x \rangle + \langle a, (1 - \alpha)y \rangle \leq \alpha c + (1 - \alpha)c = c.$$

Portanto H é convexo.

Exemplo 2.5. Uma bola $B[0, c] \in \mathbb{R}^n$, ou seja, $B[0, c] = \{x \in \mathbb{R}^n, \|x\| \leq c\}$, com $c \in \mathbb{R}_+$ fixo, é convexo. Com efeito dados $x, y \in B[0, c]$ temos que

$$\|\alpha x + (1 - \alpha)y\| \leq \alpha\|x\| + (1 - \alpha)\|y\| \leq \alpha c + (1 - \alpha)c = c.$$

Portanto $B[0, c]$ é convexo.

O conjunto $B(0, c) = \{x \in \mathbb{R}^n, \|x\| < c\}$ é chamado de *bola aberta* e é evidentemente convexo.

Exemplo 2.6. O conjunto $S_c \in \mathbb{R}^n$ tal que $S_c = \{x \in \mathbb{R}^n, \|x\| = c\}$, com $c \geq 0$ fixo, é convexo se, e somente se, $c = 0$. Com efeito, se tivermos $S_c = \{0\}$ temos que S_c é convexo. Assim

$$S_c = \{0\} \Leftrightarrow x = 0 \Leftrightarrow \|x\| = 0 \Leftrightarrow c = 0.$$

Em particular qualquer esfera não trivial não é um conjunto convexo.

Exemplo 2.7. Seja $Q \in \mathbb{R}(n, n)$ uma matriz simétrica semidefinida positiva. O conjunto $A = \{x \in \mathbb{R}^n; \langle Qx, x \rangle \leq c\}$, com $c \in \mathbb{R}_+$, é convexo. Com efeito dados $x, y \in A$ e $\alpha \in [0, 1]$ temos

$$\begin{aligned} \langle Q[\alpha x + (1 - \alpha)y], \alpha x + (1 - \alpha)y \rangle &= \\ &= \alpha^2 \langle Qx, x \rangle + \alpha(1 - \alpha) \langle Qx, y \rangle + \alpha(1 - \alpha) \langle Qy, x \rangle + (1 - \alpha)^2 \langle Qy, y \rangle \\ &= \alpha^2 \langle Qx, x \rangle + \alpha(1 - \alpha) \langle Qx + Qy, y + x \rangle + (1 - \alpha)^2 \langle Qy, y \rangle \\ &= \alpha^2 \langle Qx, x \rangle + \alpha(1 - \alpha) \langle Q(x + y), x + y \rangle + (1 - \alpha)^2 \langle Qy, y \rangle \\ &\leq \alpha^2 c + \alpha(1 - \alpha)c + (1 - \alpha)^2 c = (1 - \alpha + \alpha^2)c, \end{aligned}$$

como $\alpha \in [0, 1]$ temos então

$$0 \leq \alpha \leq 1 \Rightarrow 0 \leq \alpha^2 \leq \alpha \Rightarrow -\alpha \leq -\alpha + \alpha^2 \leq 0 \Rightarrow 1 - \alpha \leq 1 - \alpha + \alpha^2 \leq 1.$$

Portanto $\langle Q[\alpha x + (1 - \alpha)y], \alpha x + (1 - \alpha)y \rangle \leq c$.

Proposição 2.1. Sejam $D_i \subset \mathbb{R}^n$, $i \in \mathfrak{I}$, conjuntos convexos, onde \mathfrak{I} é um conjunto qualquer. Então a interseção

$$D = \bigcap_{i \in \mathfrak{I}} D_i$$

é um conjunto convexo.

Demonstração. Tomemos $x, y \in D$. Pela definição de interseção $x, y \in D_i$, para todo $i \in \mathfrak{I}$. Com cada D_i é convexo, temos que $\alpha x + (1 - \alpha)y \in D_i$, para todo $\alpha \in [0, 1]$ e todo $i \in \mathfrak{I}$. Novamente pela definição de interseção $\alpha x + (1 - \alpha)y \in D$, ou seja, D é convexo. □

Apesar dos cones nem sempre serem convexos, apresentaremos algumas propriedades úteis dos cones para a convexidade.

Definição 2.2. Um conjunto $K \subset \mathbb{R}^n$ é chamado de cone quando para $d \in K$ temos que $td \in K$, para qualquer $t \in \mathbb{R}_+$.

Veja que a definição de cone nos dá que se $K \neq \emptyset$ então $0 \in K$. Essa definição também nos mostra que um cone é também um conjunto de direções.

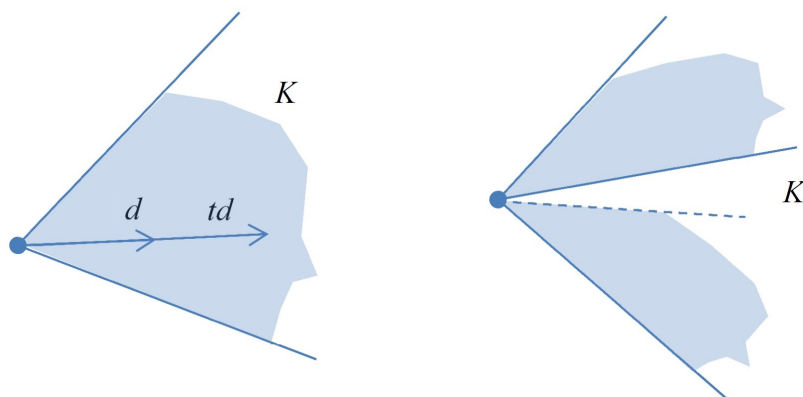


Figura 2.2: Cone convexo e não convexo respectivamente.

O espaço \mathbb{R}^n e qualquer um de seus subespaços são exemplos de cones. O octante não negativo \mathbb{R}_+^n também é um cone.

Observe na figura 2.2 que um cone nem sempre é um conjunto convexo. Entretanto a dualização de um cone (cone dual ou cone polar) é um conjunto convexo, como definimos a seguir:

Definição 2.3. O cone dual de um cone $K \subset \mathbb{R}^n$ é o conjunto definido por

$$K^* = \{y \in \mathbb{R}^n; \langle y, d \rangle \leq 0, \forall d \in K\}.$$

A proposição a seguir mostra que a dualização de um cone qualquer (mesmo não convexo) produz um cone convexo como vemos na figura 2.4.

Proposição 2.2. Dado $\emptyset \neq K \subset \mathbb{R}^n$, o cone dual K^* é convexo e fechado.

Demonstração. Sejam $x, y \in K^*$, ou seja, $\langle x, d \rangle \leq 0$ e $\langle y, d \rangle \leq 0, \forall d \in K$. Tomemos $\alpha \in [0, 1]$ e $d \in K$, assim temos

$$\langle \alpha x + (1 - \alpha)y, d \rangle = \alpha \langle x, d \rangle + (1 - \alpha) \langle y, d \rangle \leq 0,$$

ou seja $\alpha x + (1 - \alpha)y \in K^*$. Logo, K^* é convexo.

Seja $\{y_n\} \subset K^*$ uma sequência com $y_n \rightarrow y$ ($n \rightarrow \infty$). Seja $d \in K$, fixo (qualquer). Temos então

$$\langle y, d \rangle = \langle y + y_n - y_n, d \rangle = -\langle y_n - y, d \rangle + \langle y_n, d \rangle \leq -\langle y_n - y, d \rangle$$

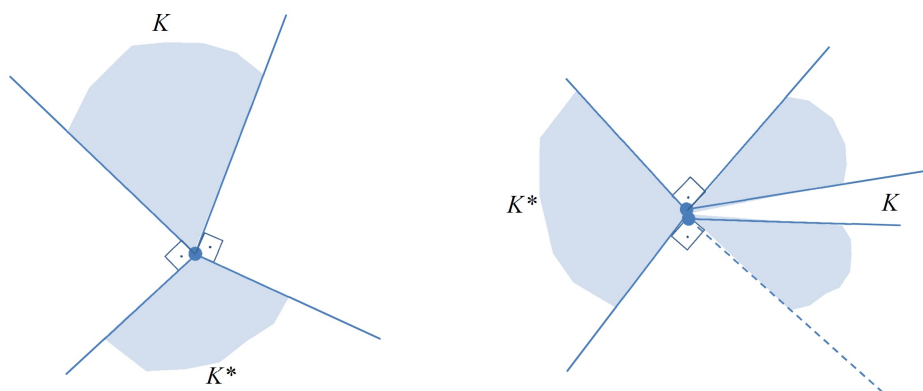


Figura 2.3: O cone Dual K^* é convexo para qualquer cone K .

fazendo $n \rightarrow \infty$ temos $\langle y_n - y, d \rangle = 0$, ou seja $\langle y, d \rangle \leq 0$. Como d é qualquer e $y \in K^*$, temos que K^* é fechado.

□

Definição 2.4. Seja $D \subset \mathbb{R}^n$ um conjunto. O fecho cônico de D , denotado por cone D , é o menor cone convexo em \mathbb{R}^n que contém D , ou equivalentemente, sendo \mathcal{L} um conjunto e $K_i \subset \mathbb{R}^n$, $i \in \mathcal{L}$, a família de cones que contém D então

$$\text{cone}D = \bigcap_{i \in \mathcal{L}} K_i$$

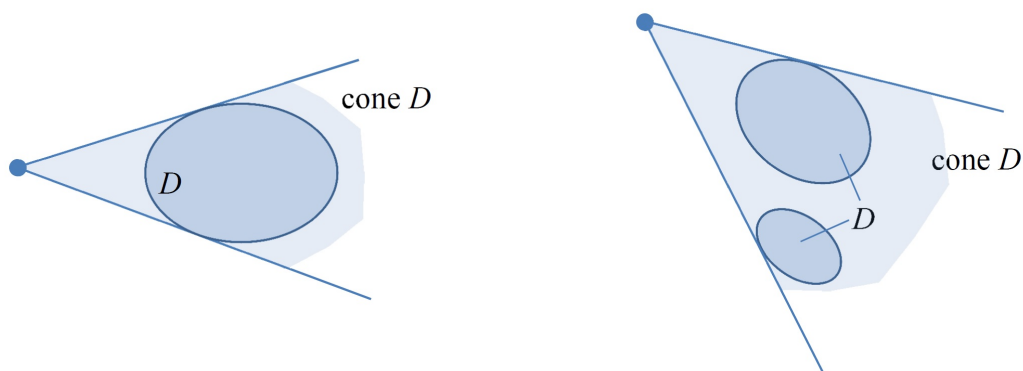


Figura 2.4: Exemplos de fecho cônico de um conjunto.

Teorema 2.1 (Fecho cônico de um conjunto convexo). Seja $D \subset \mathbb{R}^n$ um conjunto convexo. Então

$$\text{cone}D = \{d \in \mathbb{R}^n; d = \alpha x, x \in D \text{ e } \alpha \in \mathbb{R}_+\}.$$

Vamos agora definir distância de um ponto $x \in \mathbb{R}^n$ a um conjunto $D \subset \mathbb{R}^n$ como

$$\text{dist}(x, D) = \inf_{y \in D} \|x - y\|.$$

Definição 2.5. Sejam $D \subset \mathbb{R}^n$ um conjunto e $\bar{x} \in D$. Dizemos que $d \in \mathbb{R}^n$ é uma direção tangente a D no ponto \bar{x} quando

$$o(t) = \text{dist}(\bar{x} + td, D), t \in \mathbb{R}_+.$$

E o conjunto de todas as direções tangentes a D em \bar{x} é denotado por $\mathcal{T}_D(\bar{x})$.

A figura 2.5 ilustra a definição 2.5.

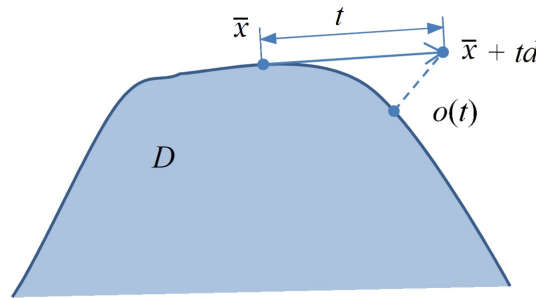


Figura 2.5: Ilustração da definição 2.5, onde d é uma direção tangente a D em \bar{x} . No desenho, consideramos $\|d\| = 1$, portanto a distancia entre os pontos \bar{x} e $\bar{x} + td$ é t . Quando $t \rightarrow 0_+$, $\text{dist}(\bar{x} + td, D)$ é de ordem menor que t , o comprimento do passo na direção d a partir de \bar{x} .

Proposição 2.3 (Cone tangente de um conjunto convexo). Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in D$. Então o cone $\mathcal{T}_D(\bar{x})$ é convexo e fechado.

Veremos agora que um cone K e seu fecho $\text{cl } K$ tem o mesmo cone dual.

Proposição 2.4. Seja $K \subset \mathbb{R}^n$ um cone qualquer. Então $\text{cl } K$ é um cone e

$$K^* = (\text{cl } K)^*.$$

Em particular, se $D \subset \mathbb{R}^n$ é um conjunto convexo e $\bar{x} \in D$, temos

$$(\mathcal{T}_D(\bar{x}))^* = (\text{cone}(D - \{\bar{x}\}))^*.$$

Definição 2.6. Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in D$. O cone normal (cone das direções normais) no ponto \bar{x} em relação a D é dado por

$$\mathcal{N}_D(\bar{x}) = \{d \in \mathbb{R}^n; \langle d, x - \bar{x} \rangle \leq 0, \forall x \in D\}.$$

A definição 2.6 de cone normal nos dá que ele é o dual do cone tangente conforme o próximo teorema.

Teorema 2.2 (Cone normal é dual ao cone tangente). *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in D$. Então*

$$(\mathcal{T}_D(\bar{x}))^* = (\text{cone}(D - \{\bar{x}\}))^* = \mathcal{N}_D(\bar{x}).$$

Teorema 2.3 (Condição necessária de primeira ordem). *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável no ponto $x_0 \in D$. Se x_0 é um minimizador local de f em D , então*

$$\langle f'(x_0), x - x_0 \rangle \geq 0,$$

ou equivalentemente,

$$-f'(x_0) \in \mathcal{N}_D(x_0).$$

Mais adiante veremos que a condição do teorema 2.3 também é suficiente para funções convexas.

Definição 2.7. *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo, $f : D \rightarrow \mathbb{R}$ uma função. A função f é dita convexa quando dados $x, y \in D$ e $\alpha \in [0, 1]$ tem-se*

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

Caso se tenha

$$f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y)$$

para $x \neq y$ e $\alpha \in (0, 1)$ dizemos que f é *estritamente convexa*.

A função f é dita *fortemente convexa com módulo $\beta > 0$* , quando para quaisquer $x, y \in D$ e $\alpha \in [0, 1]$, tem-se

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) - \beta \alpha(1 - \alpha) \|x - y\|^2$$

Note que uma função fortemente convexa é estritamente convexa e uma função estritamente convexa é convexa.

Exemplo 2.8. *Seja $f : \mathbb{R} \rightarrow \mathbb{R}$, a função definida por $f(x) = x^2$. Tomemos $x, y \in \mathbb{R}$ e $\alpha \in [0, 1]$. Vamos calcular $f(\alpha x + (1 - \alpha)y)$:*

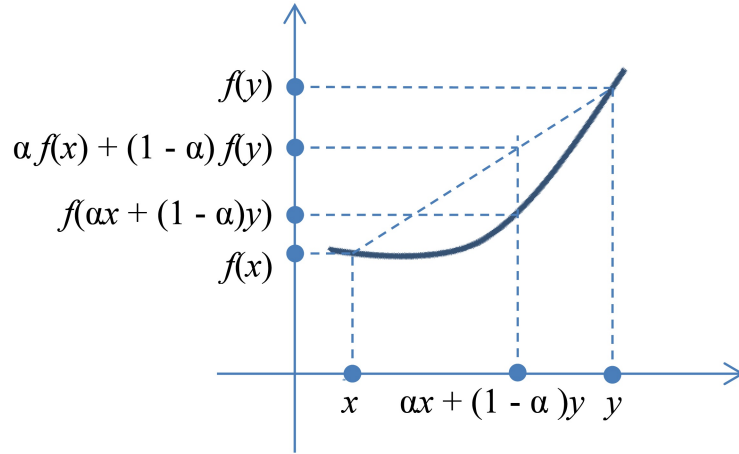


Figura 2.6: Gráfico de uma Função convexa

$$\begin{aligned}
 f(\alpha x + (1 - \alpha)y) &= (\alpha x + (1 - \alpha)y)^2 \\
 &= (\alpha x)^2 + ((1 - \alpha)y)^2 + 2\alpha(1 - \alpha)xy \\
 &= \alpha^2 x^2 + (1 - \alpha)(1 - \alpha)y^2 + 2\alpha(1 - \alpha)xy + \alpha x^2 - \alpha x^2 \\
 &= \alpha x^2 + (1 - \alpha)y^2 - \alpha(1 - \alpha)y^2 + 2\alpha(1 - \alpha)xy - \alpha(1 - \alpha)x^2 \\
 &= \alpha f(x) + (1 - \alpha)f(y) - \alpha(1 - \alpha)(x - y)^2 \\
 &= \alpha f(x) + (1 - \alpha)f(y) - \alpha(1 - \alpha)|x - y|^2.
 \end{aligned}$$

Assim tem-se que a função $f(x) = x^2$ é fortemente convexa, com $\beta = 1$.

Exemplo 2.9. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a função norma definida por $f(x_1, x_2, \dots, x_n) = \|(x_1, x_2, \dots, x_n)\|$. Tomemos $(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ e $\alpha \in [0, 1]$. Vamos calcular $f(\alpha(x_1, x_2, \dots, x_n) + (1 - \alpha)(y_1, y_2, \dots, y_n))$:

$$\begin{aligned}
 f(\alpha(x_1, x_2, \dots, x_n) + (1 - \alpha)(y_1, y_2, \dots, y_n)) &= \|\alpha(x_1, x_2, \dots, x_n) + (1 - \alpha)(y_1, y_2, \dots, y_n)\| \\
 &\leq \|\alpha(x_1, x_2, \dots, x_n)\| + \|(1 - \alpha)(y_1, y_2, \dots, y_n)\| \\
 &= \alpha\|(x_1, x_2, \dots, x_n)\| + (1 - \alpha)\|(y_1, y_2, \dots, y_n)\| \\
 &= \alpha f(x_1, x_2, \dots, x_n) + (1 - \alpha)f(y_1, y_2, \dots, y_n).
 \end{aligned}$$

Portanto,

$$f(\alpha(x_1, x_2, \dots, x_n) + (1 - \alpha)(y_1, y_2, \dots, y_n)) \leq \alpha f(x_1, x_2, \dots, x_n) + (1 - \alpha)f(y_1, y_2, \dots, y_n),$$

ou seja, a função $f(x_1, x_2, \dots, x_n) = \|(x_1, x_2, \dots, x_n)\|$ é convexa. Note que para $(x_1, x_2, \dots, x_n) \neq$

(y_1, y_2, \dots, y_n) , $\alpha \neq 0$ e $\alpha \neq 1$ a igualdade não se verifica, ou seja, a norma é estritamente convexa.

A definição e o teorema a seguir dão uma forma de relacionar funções convexas e conjuntos convexos.

Definição 2.8. *Seja $f : D \rightarrow \mathbb{R}$ uma função. O conjunto*

$$E_f = \{(x, c) \in D \times \mathbb{R} : f(x) \leq c\}$$

é denominado epígrafo da função f .

Teorema 2.4. *Seja $D \subset \mathbb{R}^n$ um conjunto convexo. Uma função $f : D \rightarrow \mathbb{R}$ é convexa em D se, e somente se, o epígrafo de f é um conjunto convexo em $\mathbb{R}^n \times \mathbb{R}$.*

Demonstração. Inicialmente vamos supor que E_f é um conjunto convexo. Sejam $x, y \in D$ arbitrários. Pela definição de epígrafo $(x, f(x)), (y, f(y)) \in E_f$. Como supomos E_f convexo, para qualquer $\alpha \in [0, 1]$ temos que

$$(\alpha x + (1 - \alpha)y, \alpha f(x) + (1 - \alpha)f(y)) = \alpha(x, f(x)) + (1 - \alpha)(y, f(y)) \in E_f.$$

Novamente pela definição de epígrafo

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

ou seja, f é convexa.

Agora vamos supor f convexa. Sejam $(x, a), (y, b) \in E_f$, ou seja, $f(x) \leq a$ e $f(y) \leq b$. Assim pela convexidade de f e para qualquer $\alpha \in [0, 1]$ tem-se

$$\begin{aligned} f(\alpha x + (1 - \alpha)y) &\leq \alpha f(x) + (1 - \alpha)f(y) \\ &\leq \alpha a + (1 - \alpha)b, \end{aligned}$$

que nos dá

$$\alpha(x, a) + (1 - \alpha)(y, b) = (\alpha x + (1 - \alpha)y, \alpha a + (1 - \alpha)b) \in E_f,$$

ou seja E_f é convexo. □

O teorema [2.4](#) nos permite classificar as funções convexas como funções cujo epígrafo é convexo.

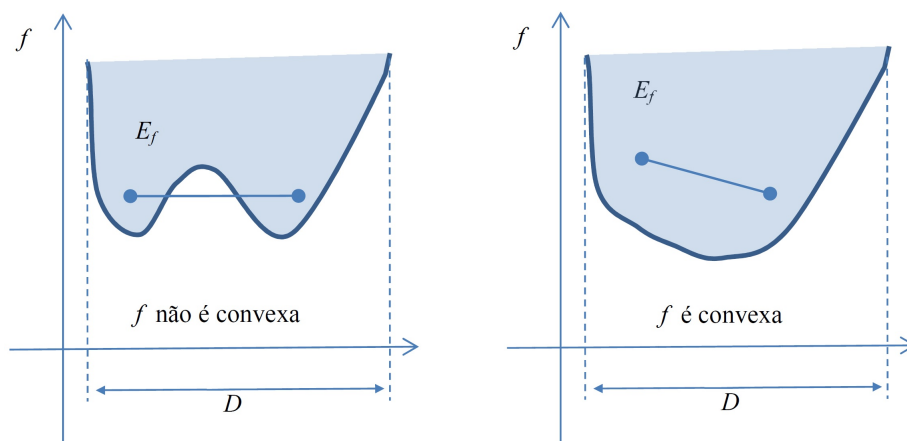


Figura 2.7: A função f é convexa \Leftrightarrow O epígrafo de f é convexo.

A teoria de funções convexas é utilizada principalmente na solução de problemas de otimização, como o problema que apresentamos a seguir.

$$\min f(x) \text{ sujeito a } x \in D. \quad (2.1)$$

é chamado de *problema minimização* e se $D \subset \mathbb{R}^n$ é um conjunto convexo e $f : D \rightarrow \mathbb{R}$ é uma função convexa em D de *problema minimização convexa*.

Apresentaremos agora algumas definições, proposições e teoremas necessários para a solução do problema [2.1](#) em sua forma convexa.

Teorema 2.5 (Teorema de Weierstrass). *Seja $D \subset \mathbb{R}^n$ um conjunto compacto não vazio e $f : D \rightarrow \mathbb{R}$ uma função contínua. Então, o problema*

$$\min f(x) \text{ sujeito a } x \in D$$

e

$$\max f(x) \text{ sujeito a } x \in D$$

tem soluções globais.

Demonstração. Por [2.1](#) e pela definição [2.10](#), basta mostrar a existência de um minimizador ou maximizador. Assim vamos mostrar a existência de um minimizador. Como D é compacto, temos que $f(D)$ também o é. Logo $f(D)$ possui ínfimo. Seja \bar{y} o ínfimo de $f(D)$. Pela definição de ínfimo podemos obter $x_k \in D$, com $k \in \mathbb{N}$, tal que

$$\bar{y} \leq f(x_k) \leq \bar{y} + \frac{1}{k}.$$

Donde temos que

$$\lim_{k \rightarrow \infty} f(x_k) = \bar{y}. \quad (2.2)$$

Como a sequência $x_k \in D$ e D é compacto, então a sequência (x_k) possui uma subsequência (x_{k_i}) convergente a um ponto de D , ou seja

$$\lim_{i \rightarrow \infty} x_{k_i} = \bar{x} \in D.$$

Temos assim que

$$\lim_{i \rightarrow \infty} f(x_{k_i}) = f(\bar{x}),$$

pois f é contínua.

Por [2.2](#), temos que $f(\bar{x}) = \bar{y}$, ou seja, f assume o seu valor mínimo em $\bar{x} \in D$. Portanto \bar{x} é um minimizador global de f .

□

A compacidade de D pode ser eliminada pelas definições e resultados que já vimos e alguns que ainda apresentaremos.

Definição 2.9. *Sejam D um conjunto convexo, $f : D \rightarrow \mathbb{R}$ uma função convexa e $c \in \mathbb{R}$, fixo. Definimos o conjunto de nível da função f relativo a c por*

$$L_{f,D}(c) = \{x \in D; f(x) \leq c\}.$$

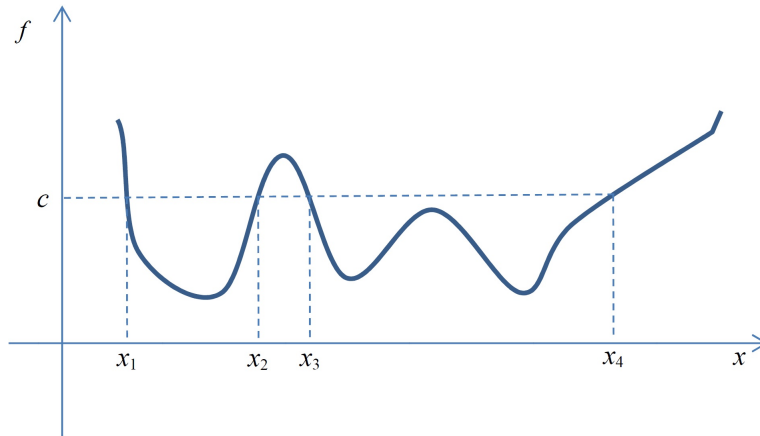


Figura 2.8: Conjunto de nível $L_{f,\mathbb{R}}(c) = [x_1, x_2] \cup [x_3, x_4]$.

Corolário 2.1. *Sejam $D \subset \mathbb{R}^n$ e $f : D \rightarrow \mathbb{R}$ uma função contínua em D . Se existe $c \in \mathbb{R}$ tal que $L_{f,D}(c) \neq \emptyset$ e compacto, então o problema [2.1](#) possui uma solução global.*

Demonstração. Por hipótese $L_{f,D}(c)$ é compacto, logo pelo teorema de Weierstrass (teorema [2.5](#)) o problema

$$\min f(x) \text{ sujeito a } x \in L_{f,D}(c)$$

tem solução global, digamos \bar{x} . Veja que se $x \in D - L_{f,D}(c)$, então $f(x) > c \geq f(\bar{x})$, ou seja, \bar{x} é minimizador global de $L_{f,D}(c)$ e também de D .

□

O teorema a seguir já demonstra a importância da convexidade para problemas de otimização.

Teorema 2.6 (Teorema da minimização convexa). *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f : D \rightarrow \mathbb{R}$ uma função convexa em D . Então todo minimizador local do problema 2.1 é global. Além disso o conjunto de minimizadores é convexo, e se f é estritamente convexa o minimizador é único.*

Demonstração. Mostraremos inicialmente que todo minimizador local é global:

Suponha que $\bar{x} \in D$ seja um minimizador local de f . Assim existe $y \in D$ tal que $f(y) < f(\bar{x})$. Façamos $x(\alpha) = \alpha y + (1 - \alpha)\bar{x}$. Como D é convexo $x(\alpha) \in D$, $\forall \alpha \in [0, 1]$, pela convexidade de f e para $\alpha \in (0, 1]$ temos que

$$\begin{aligned} f(x(\alpha)) &= f(\alpha y + (1 - \alpha)\bar{x}) \leq \alpha f(y) + (1 - \alpha)f(\bar{x}) \\ &= \alpha f(y) + f(\bar{x}) - \alpha f(\bar{x}) \\ &= f(\bar{x}) + \alpha(f(y) - f(\bar{x})) < f(\bar{x}). \end{aligned}$$

Veja que para α suficientemente pequeno $x(\alpha)$ está arbitrariamente próximo de \bar{x} . Como $f(x(\alpha)) < f(\bar{x})$ e $x(\alpha) \in D$ temos uma contradição pois \bar{x} é minimizador local do problema 2.1. Logo qualquer solução local é uma solução global.

Mostraremos que o conjunto dos minimizadores é convexo:

Sejam $S \subset D$ o conjunto das soluções (minimizadores) do problema 2.1 e $\bar{v} \in \mathbb{R}$ seu valor ótimo, ou seja, $f(x) = \bar{v}$, $\forall x \in S$. Tomemos agora $x, \bar{x} \in S$ e $\alpha \in [0, 1]$. Como f é convexa temos

$$f(\alpha x + (1 - \alpha)\bar{x}) \leq \alpha f(x) + (1 - \alpha)f(\bar{x}) = \alpha \bar{v} + (1 - \alpha)\bar{v} = \bar{v},$$

veja que como $\alpha x + (1 - \alpha)\bar{x} \in D$ (convexidade de D), temos então

$$f(\alpha x + (1 - \alpha)\bar{x}) \geq f(x) = f(\bar{x}) = \bar{v}.$$

Assim $f(\alpha x + (1 - \alpha)\bar{x}) = \bar{v}$, logo $\alpha x + (1 - \alpha)\bar{x} \in S$. Portanto S é convexo.

Para mostrar que o minimizador é único, suponha f estritamente convexa, que existam $x, \bar{x} \in S$, $x \neq \bar{x}$, e $\alpha \in (0, 1)$. Sabemos que $\alpha x + (1 - \alpha)\bar{x} \in D$, pela convexidade de D . Assim temos

$$f(\alpha x + (1 - \alpha)\bar{x}) \geq f(x) = f(\bar{x}) = \bar{v}.$$

Mas f estritamente convexa, ou seja,

$$f(\alpha + (1 - \alpha)\bar{x}) < \alpha f(x) + (1 - \alpha)f(\bar{x}) = \alpha\bar{v} + (1 - \alpha)\bar{v} = \bar{v},$$

que nos dá uma contradição. Portanto o minimizador é único. □

A necessidade de resolver problemas de otimização não passa só por encontrar minimizadores, passa também por encontrar maximizadores. Vamos dar uma definição a seguir que torna o teorema 2.6 válido também neste outro caso, bastando substituir os termos minimização por maximização e convexa por côncava.

Definição 2.10. A função $f : D \rightarrow \mathbb{R}$, com D conjunto convexo, é chamada de função côncava se $(-f)$ for convexa em D .

2.3 Operador Projeção

Definição 2.11. Sejam $D \subset \mathbb{R}^n$ um conjunto e $y \in \mathbb{R}^n$ um ponto. A projeção (ortogonal) de y sobre D é uma solução global do problema

$$\min \|x - y\|, \text{ sujeito a } x \in D. \quad (2.3)$$

Esta definição nos diz que a projeção de y sobre D é um dos pontos de D mais próximos de y . Veremos a seguir um resultado que garante a existência da projeção em um conjunto não vazio e fechado.

Corolário 2.2. Seja $D \subset \mathbb{R}^n$ um conjunto fechado e não vazio. Então a projeção de y sobre D existe para todo $y \in \mathbb{R}^n$

Demonstração. Sejam $y \in \mathbb{R}^n$, fixo e arbitrário, $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ a função $f(x) = \|x - y\|$. Sabemos que f é contínua e $L_{f,D}(C) = D \cap B[y, c]$, onde $B[y, c] = \{x \in \mathbb{R}^n; \|x - y\| \leq c\} = \{x \in \mathbb{R}^n; f(x) \leq c\}$.

O conjunto de nível $L_{f,D}(C)$ é a intersecção do fechado D com o compacto $B[y, c]$, ou seja, $L_{f,D}(C)$ é compacto e é não vazio para $c > 0$ suficientemente grande. Assim pelo corolário 2.1 o problema 2.3 tem solução. □

Teorema 2.7 (Teorema da projeção). Seja $D \subset \mathbb{R}^n$ um conjunto convexo e fechado. Então para todo $x \in \mathbb{R}^n$, a projeção de x sobre D , denotada por $P_D(x)$, existe e é única.

Além disso, $\bar{x} = P_D(x)$ se, e somente se,

$$\bar{x} \in D, \quad \langle x - \bar{x}, y - \bar{x} \rangle \leq 0 \quad \forall y \in D. \quad (2.4)$$

Demonstração. Como D é fechado pelo corolário 2.2 existe $\bar{x} \in D$ que é solução do problema 2.3. Seja $y \in D$ qualquer. Definimos $x(\alpha) = (1 - \alpha)\bar{x} + \alpha y$, com $\alpha \in [0, 1)$, pela convexidade de D temos que $x(\alpha) \in D$. Assim temos que $\|x - \bar{x}\| \leq \|x - x(\alpha)\|$, portanto

$$\begin{aligned} 0 &\geq \|x - \bar{x}\|^2 - \|x - x(\alpha)\|^2 \\ &= 2\langle x - \bar{x}, x(\alpha) - \bar{x} \rangle - \|x(\alpha) - \bar{x}\|^2 \\ &= 2\alpha\langle x - \bar{x}, y - \bar{x} \rangle - \alpha^2\|y - \bar{x}\|^2. \end{aligned}$$

Dividindo a desigualdade acima por $2\alpha > 0$ e fazendo $\alpha \rightarrow 0_+$ obtemos

$$\langle x - \bar{x}, y - \bar{x} \rangle \leq 0,$$

sendo que $y \in D$ é arbitrário.

Vamos supor que existe \bar{x} que satisfaça 2.4. Então para todo $y \in D$,

$$\langle x - \bar{x}, y - \bar{x} \rangle \leq 0.$$

Por outro lado temos que

$$\begin{aligned} \frac{1}{2}(\|x - \bar{x}\|^2 - \|x - y\|^2) &\leq \frac{1}{2}(\|x - \bar{x}\|^2 - \|x - y\|^2 + \|y - \bar{x}\|^2) \\ &= \frac{1}{2}(2\langle x - \bar{x}, y - \bar{x} \rangle) \\ &= \langle x - \bar{x}, y - \bar{x} \rangle \leq 0. \end{aligned}$$

Temos então que $\|x - \bar{x}\| \leq \|x - x(\alpha)\|$, para todo $y \in D$, ou seja, \bar{x} é a projeção de x sobre D .

Mostraremos agora que \bar{x} é único. Suponhamos que \hat{x} também seja solução do problema 2.3. Aplicando \bar{x} em 2.4 com $y = \hat{x}$ e também aplicando \hat{x} com $y = \bar{x}$, obtemos

$$\langle x - \bar{x}, \hat{x} - \bar{x} \rangle \leq 0 \quad \text{e} \quad \langle x - \hat{x}, \bar{x} - \hat{x} \rangle \leq 0.$$

Somando-se as inequações acima membro a membro temos

$$\begin{aligned} 0 &\geq \langle x - \bar{x}, \hat{x} - \bar{x} \rangle + \langle x - \hat{x}, \bar{x} - \hat{x} \rangle \\ &= \langle x, \bar{x} - \hat{x} \rangle - \langle \hat{x}, \bar{x} - \hat{x} \rangle + \langle x, \hat{x} - \bar{x} \rangle - \langle \bar{x}, \hat{x} - \bar{x} \rangle \\ &= \langle \hat{x}, \hat{x} - \bar{x} \rangle - \langle \bar{x}, \hat{x} - \bar{x} \rangle \\ &= \langle \hat{x} - \bar{x}, \hat{x} - \bar{x} \rangle \\ &= \|\hat{x} - \bar{x}\|^2. \end{aligned}$$

Portanto $\hat{x} = \bar{x}$. □

Verificamos agora que pela definição [2.6](#) de cone normal, temos que [2.4](#) é equivalente a

$$\bar{x} \in D, \quad x - \bar{x} \in \mathcal{N}_D(\bar{x}). \quad (2.5)$$

NA demonstração do teorema da projeção, supomos que na definição [2.3](#), a norma usada é a euclidiana. Se usarmos uma norma não euclideana a projeção sobre um conjunto convexo ainda existe, mas pode não ser única como vemos no exemplo a seguir.

Exemplo 2.10. *Seja $D = \{(x, y) \in \mathbb{R}_+^2; x + y = 1\}$. Veja que a projeção do ponto $(0, 0)$ sobre D não é única se utilizarmos a norma da soma*

$$\|(x, y)\|_S = |x| + |y|.$$

Observe que se $(x, y) \in D$, temos que $\|(x, y)\|_S = |x| + |y| = 1$, ou seja, a projeção de $(0, 0)$ sobre D , é qualquer de seus pontos conforme vemos na figura [2.9](#)

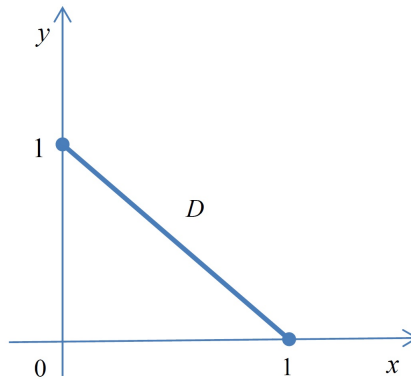


Figura 2.9: A projeção de $(0, 0)$ sobre D não é única, pois $\|(x, y)\|_S = 1$, é qualquer ponto de D .

A condição necessária de otimalidade dada no teorema [2.3](#) pode ser reescrita numa forma mais útil para o desenvolvimento de métodos computacionais utilizando-se o operador projeção, conforme é dado no teorema [2.8](#) a seguir.

Teorema 2.8 (Condição necessária de primeira ordem). *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e fechado, e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável no ponto $\bar{x} \in D$.*

Se \bar{x} é um minimizador local de f no conjunto D , então

$$P_D(\bar{x} - \alpha f'(\bar{x})) = \bar{x}, \quad \forall \alpha \in \mathbb{R}_+. \quad (2.6)$$

Demonstração. Seja $\alpha \in \mathbb{R}_+$, fixo, arbitrário. Por hipótese \bar{x} é um minimizador local de f em

D . Assim pelo teorema da projeção (teorema 2.7), com $x = \bar{x} - \alpha f'(\bar{x})$ e $y = \bar{x} \in D$, temos

$$\begin{aligned} 0 &\geq \langle \bar{x} - \alpha f'(\bar{x}) - P_D(\bar{x} - \alpha f'(\bar{x})), \bar{x} - P_D(\bar{x} - \alpha f'(\bar{x})) \rangle \\ &= \langle \bar{x} - P_D(\bar{x} - \alpha f'(\bar{x})), \bar{x} - P_D(\bar{x} - \alpha f'(\bar{x})) \rangle + \langle -\alpha f'(\bar{x}), \bar{x} - P_D(\bar{x} - \alpha f'(\bar{x})) \rangle \\ &= \|\bar{x} - P_D(\bar{x} - \alpha f'(\bar{x}))\|^2 - \alpha \langle f'(\bar{x}), \bar{x} - P_D(\bar{x} - \alpha f'(\bar{x})) \rangle. \end{aligned}$$

$$\|\bar{x} - P_D(\bar{x} - \alpha f'(\bar{x}))\|^2 \leq \alpha \langle f'(\bar{x}), \bar{x} - P_D(\bar{x} - \alpha f'(\bar{x})) \rangle. \quad (2.7)$$

Como $P_D(\bar{x} - \alpha f'(\bar{x})) \in D$ e pelo teorema 2.3 temos

$$\langle f'(\bar{x}), P_D(\bar{x} - \alpha f'(\bar{x})) - \bar{x} \rangle \geq 0,$$

ou seja,

$$\langle f'(\bar{x}), \bar{x} - P_D(\bar{x} - \alpha f'(\bar{x})) \rangle \leq 0.$$

Por 2.7 e como $\alpha \geq 0$, segue que

$$\|\bar{x} - P_D(\bar{x} - \alpha f'(\bar{x}))\|^2 \leq 0$$

e portanto $\bar{x} = P_D(\bar{x} - \alpha f'(\bar{x}))$.

□

Observamos ainda que por 2.5 a relação 2.6 é equivalente a dizer que, para todo $\alpha \in \mathbb{R}_+$, temos que

$$\bar{x} - \alpha f'(\bar{x}) - \bar{x} = -\alpha f'(\bar{x}) \in \mathcal{N}_D(\bar{x}).$$

Ou seja, $-f'(\bar{x}) \in \mathcal{N}_D(\bar{x})$, pois $\mathcal{N}_D(\bar{x})$ é um cone, que é a condição dada no teorema 2.3.

Para o caso em que a função f é convexa, a condição necessária de otimalidade de 2.6 se torna suficiente, como veremos mais adiante.

2.4 Funções Convexas

2.4.1 Propriedades fundamentais das funções convexas

Mostraremos inicialmente que uma combinação linear de funções convexas cujos coeficientes são positivos é uma função convexa.

Proposição 2.5 (Convexidade da soma de funções convexas). *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f_i : D \rightarrow \mathbb{R}$, $i = 1, 2, \dots, p$, funções convexas em D . Então para quaisquer $\mu_i \in \mathbb{R}_+$, $i = 1, 2, \dots, p$, a função*

$$f : D \rightarrow \mathbb{R}, \quad f(x) = \sum_{i=1}^p \mu_i f_i(x)$$

é convexa em D

Demonstração. Sejam $x, y \in D$ e $\alpha \in [0, 1]$ quaisquer, temos que

$$f(\alpha x + (1 - \alpha)y) = \sum_{i=1}^p \mu_i f_i(\alpha x + (1 - \alpha)y)$$

Sabemos que f_i é convexa e $\mu_i \geq 0$ para $i = 1, 2, \dots, p$, assim

$$\begin{aligned} \sum_{i=1}^p \mu_i f_i(\alpha x + (1 - \alpha)y) &\leq \sum_{i=1}^p \mu_i (\alpha f_i(x) + (1 - \alpha)f(y)) \\ &= \alpha \sum_{i=1}^p \mu_i f_i(x) + (1 - \alpha) \sum_{i=1}^p \mu_i f(y) \\ &= \alpha f(x) + (1 - \alpha)f(y). \end{aligned}$$

Portanto $f(\alpha x + (1 - \alpha)y) = \alpha f(x) + (1 - \alpha)f(y)$.

□

Vamos mostrar que o supremo de funções convexas é uma função convexa. A demonstração baseia-se no fato de que o epígrafo do supremo é a interseção dos epígrafos das funções que definem o supremo, como vemos na figura [2.10](#)

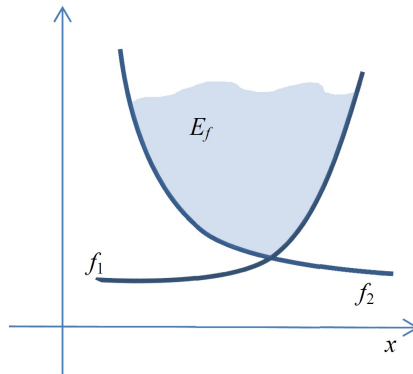


Figura 2.10: Ilustração do epígrafo da função $f(x) = \max\{f_1(x), f_2(x)\}$ que é convexo. Assim o supremo de funções convexas é uma função convexa.

Proposição 2.6 (Convexidade do supremo de funções convexas). *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f_i : D \rightarrow \mathbb{R}$, $i \in \mathfrak{I}$, funções convexas em D , com \mathfrak{I} um conjunto qualquer. Se existe*

$\beta \in \mathbb{R}$ tal que $f_i(x) \leq \beta$ para todo $x \in D$ e $i \in \mathfrak{I}$. Então

$$f : D \rightarrow \mathbb{R}, \quad f(x) = \sup_{i \in \mathfrak{I}} f_i(x)$$

é convexa em D .

Proposição 2.7. Sejam $g : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa e $h : \mathbb{R} \rightarrow \mathbb{R}$ uma função convexa e não decrescente. Então a função

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(x) = h(g(x))$$

é convexa.

Demonstração. Sejam $x, y \in \mathbb{R}^n$ e $\alpha \in [0, 1]$, quaisquer. Vamos mostrar que

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

para tanto, sabemos que

$$g(\alpha x + (1 - \alpha)y) \leq \alpha g(x) + (1 - \alpha)g(y),$$

pois g é convexa, e

$$h(g(\alpha x + (1 - \alpha)y)) \leq h(\alpha g(x) + (1 - \alpha)g(y)) \leq \alpha h(g(x)) + (1 - \alpha)h(g(y))$$

pois h é não decrescente e convexa. Sabemos também que

$$f(\alpha x + (1 - \alpha)y) = h(g(\alpha x + (1 - \alpha)y)) \quad \text{e} \quad \alpha f(x) + (1 - \alpha)f(y) = \alpha h(g(x)) + (1 - \alpha)h(g(y)),$$

ou seja,

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

□

Exemplo 2.11. Sejam $g : [\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}] \rightarrow \mathbb{R}$ e $h : \mathbb{R} \rightarrow \mathbb{R}$, as funções $g(x) = x^2 - 1$ e $h(x) = e^{-x}$. Observamos na figura [2.11](#) que g é convexa e h é convexa decrescente. A função $f : [\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}] \rightarrow \mathbb{R}$, tal que $f(x) = h(g(x)) = e^{-x^2+1}$ não é convexa. Isto mostra que a afirmação da proposição [2.7](#) pode ser falsa, caso função h não seja não decrescente.

2.4.2 Funções convexas diferenciáveis

As funções diferenciáveis nos permitem muitas formas úteis de caracterizar sua convexidade.

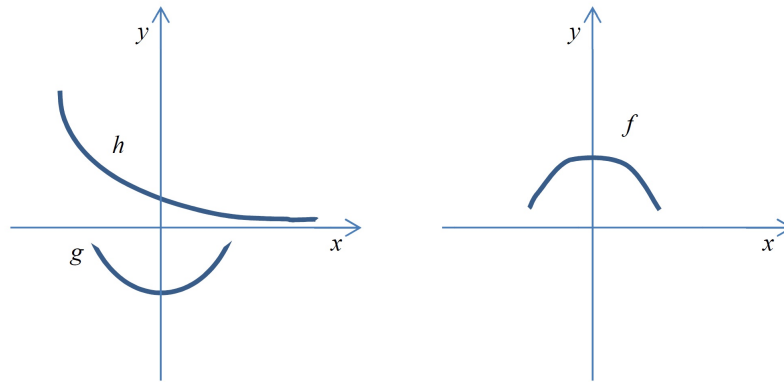


Figura 2.11: Gráficos das funções g convexa, h convexa não decrescente e pela proposição 2.7 f é não convexa.

Teorema 2.9 (Caracterização de funções convexas diferenciáveis). *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e aberto, e $f : D \rightarrow \mathbb{R}$ uma função diferenciável em D . Então as seguintes propriedades são equivalentes:*

- (a) *A função f é convexa em D .*
- (b) *Para todo $x \in D$ e todo $y \in D$, $f(y) \geq f(x) + \langle f'(x), y - x \rangle$.*
- (c) *Para todo $x \in D$ e todo $y \in D$, $\langle f'(y) - f'(x), y - x \rangle \geq 0$.*

Quando f é duas vezes diferenciável em D , as propriedades acima também são equivalentes a

- (d) *A matriz Hessiana de f é semidefinida positiva em todo ponto de D :*

$$\langle f''(x)d, d \rangle \geq 0 \quad \forall x \in D, \quad \forall d \in \mathbb{R}^n.$$

Demonstração. Mostraremos inicialmente que (a) \Rightarrow (b).

Por hipótese f é convexa. Assim para $x, y \in D$, $\alpha \in [0, 1]$, quaisquer e fazendo $d = y - x$, temos:

$$\begin{aligned} f(x + \alpha d) &= f(x + \alpha(y - x)) = f((1 - \alpha)x + \alpha y) \\ &\leq (1 - \alpha)f(x) + \alpha f(y) \\ &= f(x) - \alpha(f(x) - f(y)). \end{aligned}$$

Donde

$$f(x + \alpha d) - f(x) \leq \alpha(f(x) - f(y))$$

Dividindo ambos os membros da desigualdade acima por $\alpha > 0$, temos

$$\frac{f(x + \alpha d) - f(x)}{\alpha} \leq f(y) - f(x)$$

fazendo $\alpha \rightarrow 0_+$, temos

$$f(y) - f(x) \geq \lim_{\alpha \rightarrow 0_+} \frac{f(x + \alpha d) - f(x)}{\alpha}.$$

O segundo membro da desigualdade acima é a derivada de f em x na direção de d . Assim

$$f(y) - f(x) \geq \langle f'(x), d \rangle = \langle f'(x), y - x \rangle$$

Portanto,

$$f(y) \geq f(x) + \langle f'(x), y - x \rangle. \quad (2.8)$$

Mostraremos agora $(b) \Rightarrow (c)$.

Vamos inverter as posições de x e y em [2.8](#) e obtemos

$$f(x) \geq f(y) + \langle f'(y), x - y \rangle.$$

Somando a desigualdade acima com [2.8](#), temos

$$\begin{aligned} 0 &\geq \langle f'(x), y - x \rangle + \langle f'(y), x - y \rangle \\ &= \langle f'(x), y - x \rangle - \langle f'(y), y - x \rangle \\ &= \langle f'(x) - f'(y), y - x \rangle. \end{aligned}$$

Portanto

$$\langle f'(y) - f'(x), y - x \rangle \geq 0.$$

Mostraremos $(c) \Rightarrow (b)$.

Vamos tomar $x, y \in D$ e pelo teorema do valor médio como vemos em [4](#) existe $\alpha \in (0, 1)$ tal que

$$f(y) - f(x) = \langle f'(x + \alpha(y - x)), y - x \rangle. \quad (2.9)$$

Pela hipótese (c) com os pontos $x + \alpha(y - x)$ e x , temos

$$\begin{aligned} 0 &\leq \langle f'(x + \alpha(y - x)) - f'(x), \alpha(y - x) \rangle \\ &= \langle f'(x + \alpha(y - x)), \alpha(y - x) \rangle - \langle f'(x), \alpha(y - x) \rangle \end{aligned}$$

$$\begin{aligned}\langle f'(x + \alpha(y - x)), \alpha(y - x) \rangle &\geq \langle f'(x), \alpha(y - x) \rangle \\ \alpha \langle f'(x + \alpha(y - x)), (y - x) \rangle &\geq \alpha \langle f'(x), y - x \rangle.\end{aligned}$$

Dividindo ambos os membros da desigualdade por α e usando [2.9](#), temos então

$$f(y) \geq f(x) + \langle f'(x), y - x \rangle.$$

Mostraremos $(b) \Rightarrow (a)$.

Pela hipótese (b), para os pontos x e $x + \alpha d$, y e $x + \alpha d$, respectivamente e com $\alpha \in [0, 1]$. Temos então

$$\begin{aligned}f(x) &\geq f(x + \alpha d) - \alpha \langle f'(x + \alpha d), d \rangle \\ f(y) &\geq f(x + \alpha d) + (1 - \alpha) \langle f'(x + \alpha d), d \rangle\end{aligned}$$

Vamos multiplicar a primeira desigualdade por $1 - \alpha \geq 0$ e a segunda por $\alpha \geq 0$ e somá-las, obtendo

$$\begin{aligned}(1 - \alpha)f(x) + \alpha f(y) &\geq (1 - \alpha)f(x + \alpha d) + \alpha f(x + \alpha d) \\ &= f(x + \alpha d) \\ &= f(x + \alpha(y - x)) \\ &= f((1 - \alpha)x + \alpha y).\end{aligned}$$

Portanto f é convexa.

Para mostrar a última equivalência vamos supor que f é duas vezes diferenciável em D e provaremos que $(b) \Leftrightarrow (d)$.

Tomemos $x \in D$ e $d \in \mathbb{R}^n$, fixos e arbitrários. Como D é aberto, então existe $\alpha \geq 0$ tal que $x + \alpha d \in D$. Assim pela hipótese (b) temos

$$\begin{aligned}f(x + \alpha d) &\geq f(x) + \langle f'(x), \alpha d \rangle \\ 0 &\leq f(x + \alpha d) - f(x) - \alpha \langle f'(x), d \rangle \\ &\leq \alpha \langle f'(x + \alpha d), d \rangle - \alpha \langle f'(x), d \rangle \\ &= \alpha (\langle f'(x + \alpha d), d \rangle - \langle f'(x), d \rangle) \\ &= \frac{\alpha^2}{2} \langle f''(x) d, d \rangle + \theta(\alpha^2).\end{aligned}$$

Agora dividindo por α^2 e fazendo $\alpha \rightarrow 0_+$, obtemos

$$\langle f''(x) d, d \rangle \geq 0$$

Sejam $x, y \in D$, quaisquer. Pelo teorema do valor médio, existe $\alpha \in (0, 1)$ tal que

$$\begin{aligned} f(y) - f(x) - \langle f'(x), y - x \rangle &= \langle f'(x + \alpha(y - x)), y - x \rangle - \langle f'(x), y - x \rangle \\ &= \langle f'(x + \alpha(y - x)) - f'(x), y - x \rangle \\ &= \frac{1}{2} \langle f''(x + \alpha(y - x))(y - x), y - x \rangle. \end{aligned}$$

Sabemos que pela hipótese de (d) temos

$$\langle f''(x + \alpha(y - x))(y - x), y - x \rangle \geq 0,$$

logo obtemos (b). □

Exemplo 2.12. A função $f : D \rightarrow \mathbb{R}$, com $D = \{(x, y) \in \mathbb{R}^2; y > 0\}$ e $f(x, y) = \frac{x}{y}$, é convexa. Para provar que é convexa, vamos mostrar que f satisfaz o item (d) do teorema 2.9, pois D sendo o semiplano onde y é positivo é convexo e aberto.

Obviamente f é diferenciável ao menos duas vezes em D . O gradiente de f é o vetor

$$\nabla f(x, y) = \left(\frac{2x}{y}, -\frac{x^2}{y^2} \right), \quad (x, y) \in D,$$

e o sua matriz Hessiana em $(x, y) \in D$ é dada por

$$H_{\nabla f(x,y)} = \begin{bmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix}.$$

Para $d = (d_1, d_2) \in \mathbb{R}^n$, qualquer, o Hessiano de f em d é dado por

$$\langle H_{\nabla f(x,y)} d, d \rangle = \frac{2(d_1 y - d_2 x)^2}{y^3}$$

Como $y > 0$, temos que $\langle J_{\nabla f(x,y)} d, d \rangle \geq 0$.

Portanto f é convexa.

Exemplo 2.13. Sejam $\alpha \in \mathbb{R}$, fixo, qualquer e $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função $f(x, y) = (a+6)x + 2axy + y^2$. Vamos determinar a para os quais f é convexa. O $\nabla f(x, y) = (2(a+6)x + 2ay, 2ax + 2y)$, $(x, y) \in \mathbb{R}^2$. Assim seu Hessiano em $(x, y) \in D$ é dada por

$$H_{\nabla f(x,y)} = \begin{bmatrix} 2(a+6) & 2a \\ 2a & 2 \end{bmatrix}.$$

Para $d = (d_1, d_2) \in \mathbb{R}^n$, qualquer, o Hessiano de f em d é dado por

$$\langle H_{\nabla f(x,y)} d, d \rangle = (4a + 6)d_1^2 + (2a + 2)d_2^2.$$

Neste caso $\langle H_{\nabla f(x,y)} d, d \rangle \geq 0$ quando $a \geq -1$.

Observe que nos exemplos 2.12 e 2.13 os domínios de ambas as funções são convexos. Pelos métodos do cálculo diferencial podemos determinar minimizadores locais das funções desses exemplos e o teorema 2.6 garante que esses minimizadores serão globais.

O próximo teorema garante que as condições necessárias para otimalidade, no caso convexo, também são suficientes.

Teorema 2.10 (Condição necessária e suficiente para um problema de minimização convexa). *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f : C \rightarrow \mathbb{R}$ uma função convexa e diferenciável no aberto C que contém D . Então \bar{x} é um minimizador de f em D se, e somente se,*

$$\langle f'(\bar{x}), x - \bar{x} \rangle \geq 0, \quad \forall x \in D, \quad (2.10)$$

ou, equivalentemente,

$$-f'(\bar{x}) \in \mathcal{N}_D(\bar{x}).$$

Ainda a condição 2.10 é equivalente a

$$\langle f'(x), x - \bar{x} \rangle \geq 0, \quad \forall x \in D. \quad (2.11)$$

Se D é fechado, 2.10 e 2.11 também são equivalentes à seguinte condição:

$$\bar{x} = P_D(\bar{x} - \alpha f'(\bar{x})) \quad \text{para algum } \alpha > 0.$$

O teorema a seguir garante a invariância do gradiente da função objetivo no conjunto das soluções de um problema de minimização convexa.

Teorema 2.11. *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa duas vezes diferenciável e $D \subset \mathbb{R}^n$ um conjunto convexo fechado. Seja \bar{x} qualquer minimizador de f em D .*

Então $\hat{x} \in D$ é minimizador de f em D se, e somente se,

$$f'(\hat{x}) = f'(\bar{x}), \quad \langle f'(\bar{x}), \hat{x} - \bar{x} \rangle = 0.$$

2.4.3 Funções convexas não diferenciáveis

Apresentaremos agora resultados para funções convexas não diferenciáveis. Estas funções tem grande importância, tendo em vista que em grande parte dos casos de otimização a função objetivo ou não é diferenciável, ou per a diferenciabilidade em alguns pontos do seu domínio.

Nosso objetivo é mostrar que tais funções quando convexas possuem derivadas direcionais e ainda uma derivada generalizada, que é chamada de subdiferencial.

Teorema 2.12 (Derivada direcional de uma função convexa). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa. Então para todo $x \in \mathbb{R}^n$, f é diferenciável em cada direção $d \in \mathbb{R}^n$. Além disso,*

$$f(x + \alpha d) \geq f(x) + \alpha f'(x; d) \quad \forall \alpha \in \mathbb{R}_+.$$

Observamos ainda que para f convexa temos:

$$f'(x; d) = \lim_{\alpha \rightarrow 0^+} \frac{f(x + \alpha d) - f(x)}{\alpha} = \inf_{\alpha > 0} \frac{f(x + \alpha d) - f(x)}{\alpha} \quad (2.12)$$

Definição 2.12. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa. Dizemos que $y \in \mathbb{R}^n$ é um subgradiente de f no ponto $x \in \mathbb{R}^n$ se*

$$f(z) \geq f(x) + \langle y, z - x \rangle \quad \forall z \in \mathbb{R}^n.$$

O conjunto ∂f dos subgradientes de f em x se chama subdiferencial de f em x .

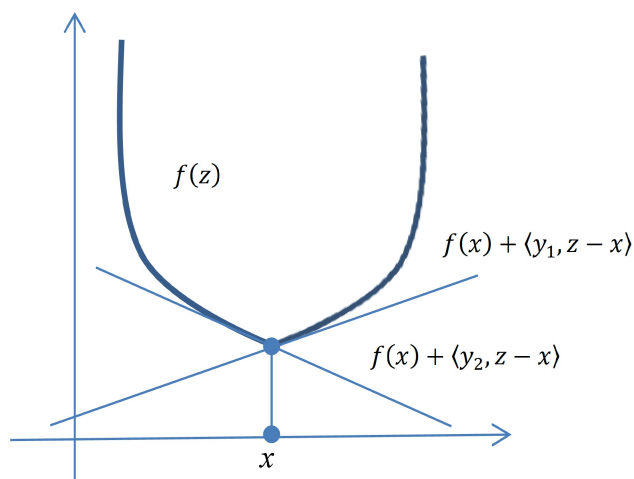


Figura 2.12: y_1 e y_2 são subgradientes de f em x , ou seja, $f(z) \geq f(x) + \langle y_i, z - x \rangle$, $i = 1, 2$.

Assim o subgradiente nos dá uma aproximação linear da função f cujo gráfico está abaixo de f e seu valor coincide com f no ponto x , veja figura 2.12. Note que esta propriedade já foi vista no teorema 2.9. Vale ressaltar que no caso em que f é diferenciável o subgradiente é único: $\partial f = \{f'(x)\}$. Isto não ocorre no caso de funções não diferenciáveis como mostra a figura 2.12 e veremos no teorema 2.13.

Vamos considerar na definição 2.12 $z = x + \alpha d$, onde $d \in \mathbb{R}^n$ e $\alpha > 0$. Assim temos

$$y \in \partial f(x) \Leftrightarrow \frac{f(x + \alpha d) - f(x)}{\alpha} \geq \langle y, d \rangle \quad \forall d \in \mathbb{R}^n, \forall \alpha > 0. \quad (2.13)$$

É fácil ver que 2.13 é equivalente a

$$f'(x; d) = \lim_{\alpha \rightarrow 0^+} \frac{f(x + \alpha d) - f(x)}{\alpha} \geq \langle y, d \rangle \quad \forall d \in \mathbb{R}^n, \quad (2.14)$$

ou seja,

$$\begin{aligned} \partial f(x) &= \{y \in \mathbb{R}^n; f(z) \geq f(x) + \langle y, z - x \rangle, \quad \forall z \in \mathbb{R}^n\} \\ &= \{y \in \mathbb{R}^n; f'(x; d) \geq \langle y, d \rangle, \quad \forall d \in \mathbb{R}^n\}. \end{aligned}$$

Veremos agora que para uma função convexa definida em todo o espaço seu subdiferencial, em cada ponto, é um conjunto convexo, compacto e não vazio e em seguida um teorema que associa o subdiferencial à diferenciabilidade das funções convexas.

Teorema 2.13 (O subdiferencial da função convexa). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa. Então para cada $x \in \mathbb{R}^n$, o conjunto $\partial f(x)$ é convexo, compacto e não vazio. além disso para qualquer $d \in \mathbb{R}^n$, tem-se*

$$f'(x; d) = \max_{y \in \partial f(x)} \langle y, d \rangle$$

Teorema 2.14 (O subdiferencial da função diferenciável). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa. A função f é diferenciável no ponto $x \in \mathbb{R}^n$ se, e somente se, o conjunto $\partial f(x)$ contém um só elemento. Neste caso, $\partial f(x) = \{f'(x)\}$.*

O próximo resultado mostra a importância do subdiferencial na teoria de otimização.

Teorema 2.15 (Condição de otimalidade para minimização de uma função convexa num conjunto convexo). *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa e $D \subset \mathbb{R}^n$ um conjunto convexo. Então $\bar{x} \in \mathbb{R}^n$ é um minimizador de f em D se, e somente se,*

$$\exists y \in \partial f(\bar{x}) \text{ tal que } \langle y, x - \bar{x} \rangle \geq 0, \quad \forall x \in D,$$

ou, equivalentemente,

$$0 \in \partial f(\bar{x}) + \mathcal{N}_D(\bar{x}).$$

Em particular, \bar{x} é um minimizador de f em \mathbb{R}^n se, e somente se,

$$0 \in \partial f(\bar{x}).$$

Teorema 2.16 (O subdiferencial da soma de funções convexas). *Sejam $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i =$*

$1, \dots, p$, funções convexas. Então

$$\partial \left(\sum_{i=1}^p f_i(x) \right) = \left(\sum_{i=1}^p \partial f_i(x) \right)$$

Exemplo 2.14. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função $f(x, y) = \|(x, y)\|_M = \max\{|x|, |y|\}$. Vamos calcular o subdiferencial de f em $(0, 0)$.

Sejam $(d_1, d_2) \in \mathbb{R}^2$ e $\alpha > 0$. Para $(x, y) \in \partial f(0, 0)$ e usando [2.14](#) temos

$$\begin{aligned} f'((0, 0); (d_1, d_2)) &= \lim_{\alpha \rightarrow 0^+} \frac{f((0, 0) + \alpha(d_1, d_2)) - f(0, 0)}{\alpha} \geq \langle (x, y), (d_1, d_2) \rangle \\ &= \lim_{\alpha \rightarrow 0^+} \frac{\alpha \cdot \max\{|d_1|, |d_2|\}}{\alpha} \geq xd_1 + yd_2 \\ &= \max\{|d_1|, |d_2|\} \geq xd_1 + yd_2. \end{aligned}$$

$$f'((0, 0); (d_1, d_2)) = \max\{|d_1|, |d_2|\} \geq xd_1 + yd_2 \quad (2.15)$$

Analisando esta última desigualdade, temos:

- (i) $d_1 = d_2 = 1 \Rightarrow x + y \leq 1$;
- (ii) $d_1 = -d_2 = 1 \Rightarrow x - y \leq 1$;
- (iii) $-d_1 = d_2 = 1 \Rightarrow -x + y \leq 1$; e
- (iv) $-d_1 = -d_2 = 1 \Rightarrow -x - y \leq 1$.

Desta forma concluímos que $(x, y) \in B[0, 1]_S$, ou seja, $\partial f(0, 0) \subset B[0, 1]_S$.

Vamos mostrar que $B[0, 1]_S \subset \partial f(0, 0)$.

Tomemos $(x, y) \in B[0, 1]_S$ e $(d_1, d_2) \in \mathbb{R}^2$, assim temos

$$\langle (x, y), (d_1, d_2) \rangle = xd_1 + yd_2 \leq |xd_1 + yd_2|$$

$$\langle (x, y), (d_1, d_2) \rangle \leq |x| \cdot |d_1| + |y| \cdot |d_2|$$

Como $\max\{|d_1|, |d_2|\}$ é maior ou igual a d_1 e d_2 , temos

$$\langle (x, y), (d_1, d_2) \rangle \leq (|x| + |y|) \cdot \max\{|d_1|, |d_2|\}$$

Como $\max\{|x|, |y|\} \leq 1$, temos

$$\langle (x, y), (d_1, d_2) \rangle \leq \max\{|d_1|, |d_2|\}$$

Por [2.15](#) temos $\langle (x, y), (d_1, d_2) \rangle \leq f'((0, 0); (d_1, d_2))$. Assim $(x, y) \in \partial f(0, 0)$, logo $B[0, 1]_S \subset \partial f(0, 0)$.

Portanto $\partial f(0, 0) = B[0, 1]_S$.

Exemplo 2.15. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função $f(x, y) = \|(x, y)\|_S = |x| + |y|$. Vamos calcular a subdiferencial de f em $(0, 0)$.

Sejam $(d_1, d_2) \in \mathbb{R}^2$ e $\alpha > 0$. Para $(x, y) \in \partial f(0, 0)$ e usando [2.14](#) temos

$$\begin{aligned} f'((0, 0); (d_1, d_2)) &= \lim_{\alpha \rightarrow 0^+} \frac{f((0, 0) + \alpha(d_1, d_2)) - f(0, 0)}{\alpha} \geq \langle (x, y), (d_1, d_2) \rangle \\ &= \lim_{\alpha \rightarrow 0^+} \frac{\alpha \cdot (|x| + |y|)}{\alpha} \geq xd_1 + yd_2 \\ &= |x| + |y| \geq xd_1 + yd_2. \end{aligned}$$

$$f'((0, 0); (d_1, d_2)) = |x| + |y| \geq xd_1 + yd_2 \quad (2.16)$$

Analisando esta última desigualdade, temos:

$$\begin{aligned} (i) \text{ Para } d_1 \neq 0 \text{ e } d_2 = 0 \Rightarrow 1 \geq \frac{xd_1}{|d_1|} &\Rightarrow \begin{cases} d_1 > 0 \Rightarrow 1 \geq x \\ d_1 < 0 \Rightarrow -1 \leq x \end{cases} \Rightarrow -1 \leq x \leq 1; \\ (ii) \text{ Para } d_1 = 0 \text{ e } d_2 \neq 0 \Rightarrow 1 \geq \frac{yd_2}{|d_2|} &\Rightarrow \begin{cases} d_2 > 0 \Rightarrow 1 \geq y \\ d_2 < 0 \Rightarrow -1 \leq y \end{cases} \Rightarrow -1 \leq y \leq 1. \end{aligned}$$

Desta forma concluímos que $(x, y) \in B[0, 1]_M$, ou seja, $\partial f(0, 0) \subset B[0, 1]_M$.

Vamos mostrar que $B[0, 1]_M \subset \partial f(0, 0)$. Tomemos $(x, y) \in B[0, 1]_M$ e $(d_1, d_2) \in \mathbb{R}^2$, assim temos

$$\langle (x, y), (d_1, d_2) \rangle = xd_1 + yd_2 \leq |xd_1 + yd_2|$$

$$\langle (x, y), (d_1, d_2) \rangle \leq |x| \cdot |d_1| + |y| \cdot |d_2|$$

Como $\max\{|x|, |y|\}$, é maior ou igual a $|x|$ e $|y|$, e é menor que 1 pois $(x, y) \in B[0, 1]_M$, temos

$$\langle (x, y), (d_1, d_2) \rangle \leq |d_1| + |d_2|$$

Por [2.16](#) temos $\langle (x, y), (d_1, d_2) \rangle \leq f'((0, 0); (d_1, d_2))$. Assim $(x, y) \in \partial f(0, 0)$, logo $B[0, 1]_M \subset \partial f(0, 0)$.

Portanto $\partial f(0, 0) = B[0, 1]_M$.

Exemplo 2.16. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função $f(x, y) = \|(x, y)\| = \sqrt{x^2 + y^2}$. Vamos calcular a subdiferencial de f em $(0, 0)$.

Sejam $(d_1, d_2) \in \mathbb{R}^2$ e $\alpha > 0$. Para $(x, y) \in \partial f(0, 0)$ e usando [2.14](#) temos

$$\begin{aligned} f'((0, 0); (d_1, d_2)) &= \lim_{\alpha \rightarrow 0^+} \frac{f((0, 0) + \alpha(d_1, d_2)) - f(0, 0)}{\alpha} \geq \langle (x, y), (d_1, d_2) \rangle \\ &= \lim_{\alpha \rightarrow 0^+} \frac{\alpha \cdot \sqrt{d_1^2 + d_2^2}}{\alpha} \geq xd_1 + yd_2 \\ &= \sqrt{d_1^2 + d_2^2} \geq xd_1 + yd_2. \end{aligned}$$

$$f'((0, 0); (d_1, d_2)) = \sqrt{d_1^2 + d_2^2} = \|(d_1, d_2)\| \geq xd_1 + yd_2 \quad (2.17)$$

Fazendo $d_1 = x$ e $d_2 = y$ na última desigualdade, temos

$$\begin{aligned} \sqrt{x^2 + y^2} &\geq x^2 + y^2 \\ 1 &\geq \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} \\ 1 &\geq \sqrt{x^2 + y^2} \\ 1 &\geq x^2 + y^2. \end{aligned}$$

Desta forma concluímos que $(x, y) \in B[0, 1]$, ou seja, $\partial f(0, 0) \subset B[0, 1]$.

Vamos mostrar que $B[0, 1] \subset \partial f(0, 0)$. Tomemos $(x, y) \in B[0, 1]$ e $(d_1, d_2) \in \mathbb{R}^2$, assim temos

$$\langle (x, y), (d_1, d_2) \rangle = |\langle (x, y), (d_1, d_2) \rangle| \leq \|(x, y)\| \cdot \|(d_1, d_2)\|,$$

Como $\|(x, y)\| \leq 1$ pois $(x, y) \in B[0, 1]$, temos

$$\langle (x, y), (d_1, d_2) \rangle \leq \|(d_1, d_2)\|$$

Por [2.17](#) temos $\langle (x, y), (d_1, d_2) \rangle \leq f'((0, 0); (d_1, d_2))$. Assim $(x, y) \in \partial f(0, 0)$, logo $B[0, 1] \subset \partial f(0, 0)$.

Portanto $\partial f(0, 0) = B[0, 1]$.

2.5 O Problema de Steiner

Vejam agora a solução do problema de Steiner que consiste em determinar um ponto do plano cuja soma das distâncias a três pontos dados, também bem do plano, é a menor possível.

Dividimos a solução em dois casos, no primeiro vamos supor que a solução é um dos pontos dados, e no segundo caso um ponto qualquer que não seja um dos pontos dados.

Então vejamos:

Sejam $z_1, z_2, z_3 \in \mathbb{C}$, fixados, e $f : \mathbb{C} \rightarrow \mathbb{R}$ definida por $f(z) = |z - z_1| + |z - z_2| + |z - z_3|$, onde $|\cdot|$ é a norma euclidiana.

2.5.1 A solução é um dos pontos dados

Vamos supor que um dos vértices é a solução do Problema de Steiner, digamos z_1 .

Vimos anteriormente que a função $|z|$ é convexa e pela proposição 2.5 a soma de funções convexas é convexa e como

$$\lim_{|z| \rightarrow \infty} |f(z)| = \infty$$

a função f tem um único mínimo global.

Sabemos também, como vimos no exemplo 2.16, que o subdiferencial de $|z - z_1|$ em z_1 é $B[0, 1]$ (a bola euclidiana unitária). Assim pelas condições de otimalidade z_1 é mínimo se, e somente se,

$$0 \in \partial f(z_1)$$

Pelo teorema 2.16, temos que

$$\partial f(z_1) = B[0, 1] + \frac{z_1 - z_2}{|z_1 - z_2|} + \frac{z_1 - z_3}{|z_1 - z_3|}.$$

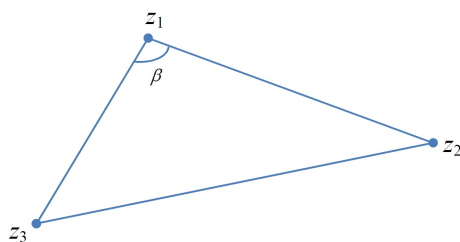


Figura 2.13: Triângulo formado pelos pontos z_1, z_2 e z_3 , destacando o ângulo interno β de z_1 .

Desta forma existe $\omega \in \mathbb{C}$ e $r \in \mathbb{R}$, com $0 \leq r \leq 1$ tais que

$$r\omega + \frac{z_1 - z_2}{|z_1 - z_2|} + \frac{z_1 - z_3}{|z_1 - z_3|} = 0,$$

por rotação podemos supor $\frac{z_1 - z_3}{|z_1 - z_3|} = 1$, $\frac{z_1 - z_2}{|z_1 - z_2|} = e^{\beta i}$ e $\omega = e^{\mu i}$, onde β é o ângulo interno do vértice z_1 conforme vemos na figura 2.13.

Assim temos

$$re^{\mu i} + e^{\beta i} + 1 = 0$$

Pela fórmula de Euler temos então

$$r(\cos \mu + i \operatorname{sen} \mu) + (\cos \beta + i \operatorname{sen} \beta) + 1 = 0,$$

donde obtemos o sistema

$$\begin{cases} r \cos \mu + \cos \beta + 1 = 0 \\ r \operatorname{sen} \mu + \operatorname{sen} \beta = 0 \end{cases}$$

$$\begin{cases} r \cos \mu = -1 - \cos \beta \\ r \operatorname{sen} \mu = -\operatorname{sen} \beta \end{cases} .$$

Elevando cada membro das equações ao quadrado e somando as equações resultantes:

$$r^2 = 2(1 + \cos \beta) = 4 \cos^2 \frac{\beta}{2}.$$

Como $r \leq 1$, temos:

$$4 \cos^2 \frac{\beta}{2} \leq 1 \rightarrow \left| \cos \frac{\beta}{2} \right| \leq \frac{1}{2},$$

ou seja,

$$120^\circ \leq \beta \leq 240^\circ,$$

mas como β é ângulo interno de um triângulo fica restrito à

$$120^\circ \leq \beta \leq 180^\circ.$$

Concluimos assim que para termos como solução em um dos pontos dados é necessário e suficiente que um dos ângulos internos do triângulo formado por estes pontos seja maior que 120° e seu vértice é a solução.

2.5.2 A solução é um ponto diferente dos pontos dados

Vejamos agora o caso em que nenhum vértice é solução:

Pelo que vimos no caso anterior o triângulo formado pelos pontos dados z_1 , z_2 e z_3 têm ângulos internos menores que 120° .

Sabemos que f é diferenciável para qualquer z em \mathbb{C} , exceto para z_1 , z_2 e z_3 . A solução do problema de otimização é um ponto z_0 onde f é diferenciável. Um candidato natural é um z_0 que seja ponto crítico de f , ou seja, $\nabla f(z_0) = 0$

$$\frac{z_0 - z_1}{|z_0 - z_1|} + \frac{z_0 - z_2}{|z_0 - z_2|} + \frac{z_0 - z_3}{|z_0 - z_3|} = 0.$$

O primeiro membro da equação acima é a soma de vetores unitários que podem ser representados por números complexos e z_0 é o vértice dos ângulos formados por estes vetores. Por uma rotação podemos supor

$$\frac{z_0 - z_3}{|z_0 - z_3|} = 1$$

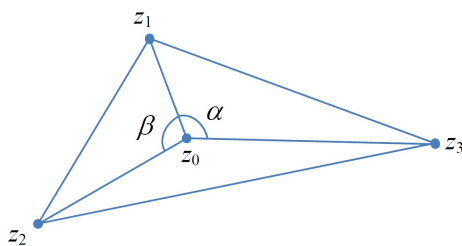


Figura 2.14: Triângulo formado pelos pontos z_1 , z_2 e z_3 , destacando os ângulo α e β formados pelos pontos z_1 , z_2 , z_3 e z_0 .

e existem $\alpha, \beta \in \mathbb{R}$, conforme mostrado na figura 2.14, tais que

$$\frac{z_0 - z_1}{|z_0 - z_1|} = e^{\alpha i}$$

e

$$\frac{z_0 - z_2}{|z_0 - z_2|} = e^{\beta i}.$$

Assim temos

$$e^{\alpha i} + e^{\beta i} + 1 = 0.$$

Novamente pela fórmula de Euler temos

$$\cos \alpha + i \operatorname{sen} \alpha + (\cos \beta + i \operatorname{sen} \beta) + 1 = 0,$$

donde obtemos o sistema

$$\begin{cases} \cos \alpha + \cos \beta + 1 = 0 \\ \operatorname{sen} \alpha + \operatorname{sen} \beta = 0 \end{cases} \quad \begin{cases} \cos \alpha = -1 - \cos \beta \\ \operatorname{sen} \alpha = -\operatorname{sen} \beta \end{cases}.$$

Procedendo novamente como no caso anterior, obtemos

$$\cos \alpha = \cos \beta = -\frac{1}{2}$$

que nos dá $\alpha = 120^\circ$ e $\beta = 240^\circ$, ou seja, os ângulos no vértice z_0 são todos iguais a 120°

Considerações Finais

Um dos principais ramos da pesquisa em matemática aplicada é a otimização. A teoria da Análise Convexa torna-se um item de fundamental importância para resolver questões relacionadas à este assunto. Vemos neste trabalho que após o estudo das principais definições e teoremas temos um ambiente bastante favorável ao desenvolvimentos e soluções das questões relacionadas à otimização.

Apresentamos o problema de Steiner como um exemplo de aplicação da teoria da análise convexa. Fizemos, para uso da teoria de Análise Convexa, a modelagem do problema através da função norma que mostramos ser uma função convexa e não diferenciável em todo seu domínio, onde se fez necessário utilizar a definição de *subdiferencial* para chegarmos a solução do problema.

O estudo dos temas apresentados são proposto não só para a solução do problema apresentado mas também como referência para trabalhos posteriores que envolvam otimização e suas aplicações. Vemos aqui que a modelagem através de funções convexas abrem muitas possibilidades de aplicações sejam elas em matemática ou geometria, como o problema de Steiner, o em outras ciências ou no desenvolvimento de novas tecnologias.

Referências Bibliográficas

- [1] Richard Courant and Herbert Robbins. *O que é Matemática?* Editora Ciência Moderna Ltda., 2000.
- [2] Alexey Izmailov, Alexey Izmailov, Mikhail Solodov, and Mikhail Solodov. *Otimização*. IMPA, 2005.
- [3] E.L. Lima and Instituto de Matematica Pura e Aplicada (Brasil). *Álgebra linear*. Coleção matemática universitária. IMPA, 2008.
- [4] Elon L. Lima and Instituto de Matematica Pura e Aplicada (Brasil). *Análise Real, Vol. 2*. Projeto Euclides. IMPA, 1981.
- [5] L. Robbiano. *Álgebra Linear: Para todos*. SpringerLink : Bücher. Springer, 2011.
- [6] Earl William Swokowski. *Cálculo com Geometria Analítica, Vol. II*. Makron Books, 1994.