



Universidade Federal de Goiás
Instituto de Matemática e Estatística
Programa de Mestrado Profissional em
Matemática em Rede Nacional



Introdução à Análise Convexa

Conjuntos e Funções Convexas

Ronan Gomes de Amorim

Goiânia

2013

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR ELETRONICAMENTE OS TRABALHOS DE CONCLUSÃO DE CURSO NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico:

Trabalho de Conclusão de Curso de Mestrado Profissional em Matemática

2. Identificação do Trabalho

Autor:	Ronan Gomes de Amorim				
E-mail:	ronanamorim@bol.com.br				
Seu e-mail pode ser disponibilizado na página? <input checked="" type="checkbox"/> Sim <input type="checkbox"/> Não					
Vínculo empregatício do autor	Colégio Militar Dom Pedro II - CMDPII				
Agência de fomento:	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior			Sigla:	CAPES
País:	Brasil	UF:	DF	CNPJ:	05.509.077/0001-05
Título:	Introdução à Análise Convexa: Conjuntos e Funções Convexas				
Palavras-chave:	Conjuntos Convexos, Funções Convexas, Convexidade				
Título em outra língua:	Introduction to Convex Analysis: Convex Sets and Functions				
Palavras-chave em outra língua:	Convex Sets, Convex Functions, Convexity				
Área de concentração:	Matemática do Ensino Básico				
Data defesa: (dd/mm/aaaa)	18 de março de 2013				
Programa de Pós-Graduação:	Mestrado Profissional em Matemática em rede nacional				
Orientador:	Professor Doutor José Yunier Bello Cruz				
E-mail:	Yunier.bello@gmail.com				
Co-orientador(a):*					
E-mail:					

*Necessita do CPF quando não constar no SisPG

3. Informações de acesso ao documento:

Concorda com a liberação total do documento SIM NÃO¹

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF ou DOC do trabalho de conclusão de curso.

O sistema da Biblioteca Digital de Teses e Dissertações garante aos autores, que os arquivos contendo eletronicamente as teses, dissertações ou trabalhos de conclusão de curso, antes de sua disponibilização, receberão procedimentos de segurança, criptografia (para não permitir cópia e extração de conteúdo, permitindo apenas impressão fraca) usando o padrão do Acrobat.

Data: 18 de março de 2013.

Assinatura do autor

¹ Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

Ronan Gomes de Amorim

Introdução à Análise Convexa

Conjuntos e Funções Convexas

Trabalho de Conclusão de Curso apresentado ao Instituto de Matemática e Estatística da Universidade Federal de Goiás, como parte dos requisitos para obtenção do grau de Mestre em Matemática.

Área de Concentração: Matemática do Ensino Básico

Orientador: Prof. Dr. José Yunier Bello Cruz

Goiânia

2013

**Dados Internacionais de Catalogação na Publicação (CIP)
GPT/BC/UFG**

A524i Amorim, Ronan Gomes de.
Introdução à análise convexa [manuscrito]: conjuntos e funções convexas / Ronan Gomes de Amorim. – 2013.
79 f. : il., figs., tabs.

Orientador: Prof. Dr. José Yunier Bello Cruz.
Dissertação (Mestrado) – Universidade Federal de Goiás,
Instituto de Matemática e Estatística, 2013.

Bibliografia.
Inclui lista de figuras.

1. Conjuntos Convexos 2. Funções Convexas 3. Convexidade.
I. Título.

CDU: 517.5

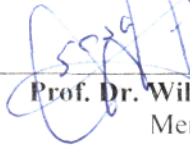
Ronan Gomes de Amorim

Introdução à Análise Convexa: Conjuntos e Funções Convexas

Trabalho de Conclusão de Curso defendido no Programa de Mestrado Profissional em Matemática em Rede Nacional – PROFMAT/UFG, do Instituto de Matemática e Estatística da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em Matemática, área de concentração Matemática do Ensino Básico, aprovado no dia 18 de março de 2013, pela Banca Examinadora constituída pelos professores:



Prof. Dr. José Yunier Bello Cruz
Instituto de Matemática e Estatística-UFG
Presidente da Banca



Prof. Dr. Wilfredo Sosa Sandoval
Membro/UCB



Prof. Dr. Luis Román Lucambio Pérez
Instituto de Matemática e Estatística-UFG

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Ronan Gomes de Amorim

Graduou-se em Matemática na UCB - Universidade Católica de Brasília. Durante sua graduação, foi monitor no Departamento de Matemática da UCB nas disciplinas Introdução à Teoria dos Números, Álgebra I e Álgebra II. Atualmente trabalha com turmas do Ensino Médio no CMDPII - Colégio Militar Dom Pedro II, instituição ligada ao CBMDF - Corpo de Bombeiros Militar do Distrito Federal.

Dedico este trabalho à minha amada Esposa Lia, à minha linda Filha Ana Gabriela e ao meu Filhinho João Gabriel que nasce no mês de julho próximo.

Agradecimentos

Agradeço primeiramente a Deus, por me conceder saúde e sabedoria para conduzir meus planos com discernimento necessário para evitar o fracasso.

À minha esposa Lia, à minha filhinha Ana Gabriela e ao meu bebê que está a caminho, por terem compreendido meus momentos de ausência e por terem aberto mão do lazer e descanso ao meu lado. Obrigado, amo vocês.

Aos meus pais, Sebastião e Hilda, por terem me educado e moldado meu caráter, ensinando-me sempre o respeito ao próximo e o agir com honestidade. Aos meus sete irmãos, especialmente ao Professor Doutor Ronni G. G. de Amorim, pelas palavras de incentivo e pelo exemplo da busca diária pelo conhecimento.

Aos companheiros de mestrado, em especial aos amigos do transporte solidário: Carlos, Hélder, Pedro e Wálter, pelo companheirismo e momentos de descontração, tão necessários em caminhadas longas como estas. Agradeço também às amigas Karina e Vanessa pela generosidade e carinho sempre demonstrados.

Aos professores, tutores e coordenadores da Universidade Federal de Goiás que nos ajudaram ao longo do curso, em especial ao meu orientador, Professor José Yunier Bello Cruz, pelo desprendimento e atenção com que sempre me tratou. Também queria estender meu agradecimento ao Professor Mário José de Souza, que transbordando paciência participou de todos os semestres com nossa turma. A todos vocês, meus sinceros agradecimentos.

À CAPES pelo apoio financeiro.

Resumo

Amorim, Ronan Gomes de. **Introdução à Análise Convexa**. Goiânia, 2013. 79p. Dissertação de Mestrado. Instituto de Matemática e Estatística, Universidade Federal de Goiás.

Neste trabalho, apresentamos as principais ideias concernentes aos conjuntos convexos e às funções convexas. Nosso principal foco é tratar, de forma didática, os principais tópicos envolvidos na convexidade, bem como a consequente exploração dos conceitos matemáticos envolvidos. Nesse sentido, realizamos uma revisão bibliográfica que contemplou teoremas, lemas, corolários e proposições relevantes a um primeiro leitor e a todos que pretendem trabalhar com as aplicações decorrentes da convexidade. Assim, esperamos que este material constitua uma importante fonte de pesquisa a estudantes, professores e pesquisadores que almejem estudar conteúdos relacionados aos conjuntos convexos.

Palavras-chave

Conjuntos Convexos, Funções Convexas, Convexidade

Abstract

Amorim, Ronan Gomes de. **Introdução à Análise Convexa**. Goiânia, 2013. 79p. MSc. Dissertation. Instituto de Matemática e Estatística, Universidade Federal de Goiás.

This paper presents the main ideas concerning convex sets and functions. Our aim is to deal, didactically, with the main topics concerning convexity, as well as the consequent exploitation of the involved mathematical concepts. In this sense, we have made a bibliographic revision approaching important theorems, lemmas, corollaries and propositions designed both to first readers and to those who want to work with applications arising from convexity. We hope that this study may constitute an important research source either for students, teachers or researchers who wish to learn more about convex sets.

Keywords

Convex Sets, Convex Functions, Convexity

Sumário

Lista de Figuras	10
1 Introdução	12
2 Preliminares	15
2.1 Tópicos de Álgebra Linear	15
2.1.1 Espaços Vetoriais Reais	15
2.1.2 Exemplos de Espaços Vetoriais	17
2.2 Subespaços Vetoriais	18
2.2.1 Exemplos de Subespaços Vetoriais	18
2.3 Combinação Linear	19
2.4 Dependência e Independência Linear	20
2.5 Base de um Espaço Vetorial	20
2.6 Dimensão de um Espaço Vetorial	21
2.7 Produto Interno	21
2.8 Desigualdades	22
2.9 Tópicos de Análise	24
2.9.1 Algumas Noções Topológicas: O espaço euclidiano n -dimensional	24
2.9.2 Tipos Especiais de Conjuntos	24
2.9.3 Princípio de Indução Matemática	27
3 Conjuntos Convexos	30
3.1 Um Pouco de História	30
3.2 Conjuntos Convexos	31
3.3 Propriedades dos Conjuntos Convexos	33
3.4 Fecho Convexo	40
3.5 Ponto Extremo	41
3.6 Hiperplano	42
3.6.1 Hiperplano Suporte e Hiperplano Separador	43
3.7 Semi-espaços	46
3.7.1 Politopo	48
3.8 O Operador de Projeção	48
3.9 Teoremas da Separação	50
4 Funções Convexas	57
4.1 Caracterização da Função Convexa	57
4.2 Propriedades das Funções Convexas	63
4.3 Funções Convexas Diferenciáveis	73

5	Considerações Finais e Perspectivas	77
	Referências Bibliográficas	78

Lista de Figuras

2.1	$X = \{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$, $p \in \text{int}(X)$.	26
3.1	(a) e (b) conjuntos convexos; (c) Conjunto não-convexo.	31
3.2	Combinação convexa	32
3.3	União de conjuntos convexos	34
3.4	Um exemplo que mostra que a soma de conjuntos convexos fechados pode não ser fechada: $D_1 + D_2 = \{x \in \mathbb{R}^2 \mid x_2 > 0\}$.	39
3.5	fecho convexo	41
3.6	$\Omega := \{(-1, 0)\} \cup \{(x_1, x_2) \mid x_1^2 + x_2^2 = 1, x_1 \geq 0\}$	42
3.7	Exemplo de ponto extremo de um conjunto	42
3.8	hiperplano suporte	44
3.9	hiperplano separador	44
3.10	Representação Gráfica do \mathbb{R}^2 .	45
3.11	Hiperplano $r: 2x_1 + 5x_2 = 10$ no \mathbb{R}^2 .	45
3.12	Semi-espaço $2x_1 + 5x_2 \leq 10$ no \mathbb{R}^2 .	46
3.13	(a) Poliedro limitado; (b) Poliedro ilimitado	48
3.14	\bar{x} é a projeção de x sobre D ; tem-se que para todo $y \in D$, $\langle x - \bar{x}, y - \bar{x} \rangle \leq 0$ ou, equivalentemente, $x - \bar{x} \in N_D(\bar{x})$.	49
3.15	Os conjuntos D_1 e D_2 são separáveis; os conjuntos C_1 e C_2 são estritamente separáveis.	51
3.16	Os conjuntos D_1 e D_2 não são separáveis.	52
3.17	Ilustração do Lema de Minkowski. Quando $x \neq \text{cl}(D)$, o ponto x pode ser separado de D estritamente.	52
3.18	Tem-se que $0 \in \text{cl}(D)$, onde D é dado por (3-6); mas o zero não é estritamente separável de D .	53
3.19	Separação de ponto x do conjunto D quando x pertence a fronteira de D ; hiperplanos de apoio de D em $x \in \text{fr}(D)$.	54
3.20	Separação de D_1 e D_2 quando $D_1 \cap D_2 = \emptyset$.	55
4.1	Gráfico de uma Função Convexa	58
4.2	Convexidade da função $f \iff$ convexidade do epígrafo de f .	61
4.3	O epígrafo de supremo de funções convexas é um conjunto convexo (no desenho, $f(x) = \max\{f_1(x), f_2(x)\}$). Portanto, o supremo de funções convexas é uma função convexa.	64
4.4	Conjuntos de nível de uma função convexa são convexos	66
4.5	Um exemplo de função quase-convexa.	67
4.6	Ilustração de prova do Teorema 4.16.	69
4.7	A função f é convexa em $D = \{x \in \mathbb{R} \mid x \geq -2\}$, mas não é contínua no ponto $x = -2$ da fronteira de D .	70

4.8 Ilustração do Teorema 4.20(2). Para todo y , tem-se que $f(y) \geq f(x) + \langle f'(x), y - x \rangle$, onde x é fixo, porém arbitrário.

Introdução

A constante ameaça da 2ª Guerra Mundial fez o governo da Grã-Bretanha recrutar um grupo de cientistas formado por matemáticos, químicos, físicos e engenheiros, para realizarem um estudo sistemático de problemas estratégicos e táticos associados à defesa aérea, terrestre e de guerra anti-submarina. O objetivo era determinar a melhor utilização dos recursos militares um tanto quanto limitados na época. O desenvolvimento desses estudos deu origem ao termo Pesquisa Operacional, cujos métodos visam auxiliar na seleção da melhor maneira de se operar um sistema que exijam a utilização de recursos escassos. Mesmo com o fim do período de guerra, muitos destes cientistas continuaram suas pesquisas com objetivos militares e também não militares. Um nome que merece destaque neste trabalho é o do matemático americano George Bernard Dantzig (1914 - 2005), considerado o pai da programação linear, um ramo da Otimização, e responsável pela criação do algoritmo Simplex, que é um método iterativo, muito eficaz, usado para encontrar, algebricamente, a solução ótima de um Problema de Programação Linear. Com o avanço computacional, o método Simplex de Dantzig, tornou-se a principal ferramenta para a solução de Problemas de Programação Linear cada vez mais complexos [1, 15, 16].

Atualmente, em diversos setores da sociedade, a Pesquisa Operacional está presente por meio de sistemas de suporte de tomadas de decisão. Sistemas estes que se baseiam na modelagem matemática e métodos quantitativos que auxiliam na tomada de decisão. Esses modelos matemáticos são, em geral, modelos de Otimização que buscam obter decisões ótimas, com custo mínimo ou lucro máximo. Para cada problema de planejamento existe um modelo de Otimização. Ao conjunto de modelos e métodos de Otimização chamaremos Programação Matemática. A modelagem matemática de um problema real começa pela adoção adequada de uma notação que representará as quantidades envolvidas na definição do problema. Na sequência, construímos uma fórmula matemática envolvendo tais quantidades e chamamos esta fórmula de função-objetivo. Com a modelagem matemática concluída, busca-se determinar os valores das variáveis envolvidas que conduzem ao menor ou maior valor da função-objetivo, isto é, valores capazes de minimizar ou maximizar a função-objetivo.

Esquemáticamente, um problema de otimização, considerando os conjuntos

$D \subset \mathbb{R}^n$ e $\Omega \subset \mathbb{R}^n$, e uma função $f : \Omega \rightarrow \mathbb{R}$, seria resumido da seguinte maneira [8]:

$$\begin{cases} \text{Minimizar } f(x) \\ \text{sujeito a } x \in D, \end{cases}$$

onde o conjunto D é o conjunto viável ou factível do problema, os pontos de D serão chamados pontos viáveis, e f será a função-objetivo. Com isso, teríamos um modelo matemático da realidade, pois estamos empregando símbolos matemáticos para representar as variáveis do sistema real. Tipicamente, o conjunto viável de um problema é definido por um sistema de igualdades e/ou desigualdades e/ou inclusão, como por exemplo,

$$D = \{x \in \Omega \mid h_i(x) = 0, \quad i = 1, \dots, l, \quad g_j(x) \leq 0, \quad j = 1, \dots, m\},$$

ou, em notação mais compacta,

$$D = \{x \in \Omega \mid h(x) = 0, \quad g(x) \leq 0\},$$

onde $\Omega \subset \mathbb{R}^n$, $h : \Omega \rightarrow \mathbb{R}^l$ e $g : \Omega \rightarrow \mathbb{R}^m$ e também supomos que a função objetivo f é definida no conjunto Ω . O conjunto Ω representa o que se chama restrições diretas e as restrições de igualdade e desigualdades se chamam restrições funcionais. Quando $D = \mathbb{R}^n$, dizemos que o problema é irrestrito, e quando $D \neq \mathbb{R}^n$ falamos de otimização com restrições. O problema irrestrito é um caso particular de problemas de minimização se tomamos $\Omega = \mathbb{R}^n$ e eliminamos as restrições funcionais.

Uma classe especial de problemas de otimização se refere ao caso em que D é um conjunto poliedral. Se além disso f for quadrática, trata-se de um problema de programação quadrática, e se f for linear, o problema é de programação linear. A solução destes problemas está em encontrar valores adequados das variáveis de decisão que otimizem o desenvolvimento do sistema. Nesse panorama, o desenvolvimento dos métodos de otimização encontram sustentação na teoria de conjuntos convexos e funções convexas que se encontram nos Capítulos 3 e 4 respectivamente.

Dessa forma, tendo a vista a relevância do assunto, estudaremos neste trabalho conjuntos e funções convexas. Convexidade é uma noção muito importante na teoria de otimização. A otimização convexa é um ramo na área de otimização matemática que trata de problemas nos quais a função objetivo e o conjunto factível são ambos convexas. Com hipóteses de convexidade, as condições necessária de otimalidade passam a ser suficientes, isto é, todo ponto estacionário torna-se uma solução do problema. Particularmente, qualquer minimizador local é global. Problemas de otimização convexa

surtem em diversas aplicações e já possuem uma teoria rica e bastante desenvolvida, o que possibilitou a criação de métodos robustos e eficientes para resolvê-los. Além disso, no caso convexo podemos desenvolver a teoria da dualidade em sua forma mais completa, isto é, associar a um problema original (primal) um outro problema, chamado dual, que sob certas hipóteses é equivalente ao original e muitas vezes, mais fácil de ser resolvido [2].

Embora a teoria matemática em que se baseia esta área seja relativamente antiga, foi a partir do final do século XX que houve um grande desenvolvimento no estudo de aplicações e na construção de métodos computacionais específicos para essa classe de problemas. A concepção dos métodos de pontos interiores para problemas de programação linear criou uma nova classe de algoritmos que foram generalizados e que hoje resolvem de forma rápida e confiável os problemas de otimização convexa [19].

As aplicações da convexidade estão em diversas áreas do conhecimento, como a engenharia, a física, a administração, a economia, dentre outras [4, 6].

Esta dissertação está organizada em cinco capítulos. Sendo que o primeiro destina-se apenas para a introdução do trabalho. No Capítulo 2, apresentamos algumas definições e noções básicas sobre Álgebra Linear e Análise, necessárias para a apresentação dos dois capítulos seguintes. O Capítulo 3 é dedicado exclusivamente à definição de Conjuntos Convexos, suas propriedades, exemplos, o Operador de Projeção e os Teoremas da Separação. No Capítulo 4, inspirados no capítulo anterior, apresentamos as definições de Função Convexa, Estritamente Convexa, Fortemente Convexa, Quase-convexa, bem como suas propriedades básicas. Além disso, ainda no Capítulo 4, fizemos um breve comentário sobre Funções Convexas Diferenciáveis afim de finalizar o capítulo e despertar o leitor a novas pesquisas. Por fim, no Capítulo 5, apresentamos nossas considerações finais e perspectivas.

Preliminares

Neste capítulo, daremos atenção especial a alguns resultados que serão utilizados nos capítulos posteriores. Revisaremos alguns temas relacionados às disciplinas de Álgebra Linear e Análise necessários para a compreensão dos resultados deste trabalho. Os conceitos e resultados apresentadas neste capítulo são baseadas nas referências [3, 9, 11, 12, 13, 14, 18].

2.1 Tópicos de Álgebra Linear

2.1.1 Espaços Vetoriais Reais

Vamos ilustrar o estudo de espaços vetoriais, o principal objetivo da Álgebra Linear, dando algumas definições e relacionando as propriedades fundamentais de tal estrutura. Nosso trabalho será desenvolvido no espaço \mathbb{R}^n .

Definição 2.1. *Sejam as n -uplas de números reais $\mathbf{x} = (x_1, \dots, x_n)$ e $\mathbf{y} = (y_1, \dots, y_n)$. Definimos sua soma por*

$$\mathbf{x} + \mathbf{y} = (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n). \quad (2-1)$$

O resultado da adição é novamente uma n -upla de números reais. Podemos definir de modo semelhante a multiplicação de uma n -upla $\mathbf{x} = (x_1, \dots, x_n)$ por um número real λ .

Definição 2.2. *Seja $\lambda \in \mathbb{R}$ e $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$. Definimos então*

$$\lambda \mathbf{x} = \lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n) \in \mathbb{R}^n. \quad (2-2)$$

Ao realizarmos estas operações com as n -uplas, nada mais estamos a fazer que realizando operações básicas com suas componentes, que são números reais. Daí, continuam válidas as regras usuais aplicáveis aos números. Referente a adição, temos:

$$(A_1) \text{ Para qualquer } \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n \text{ temos } (\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z}).$$

(A₂) Para qualquer $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ temos $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.

(A₃) Para qualquer $\mathbf{x} \in \mathbb{R}^n$ temos $\mathbf{x} + \mathbf{0} = \mathbf{x}$, onde $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^n$.

(A₄) Para qualquer $\mathbf{x} \in \mathbb{R}^n$ temos $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$, onde $-\mathbf{x} = -(x_1, \dots, x_n) \in \mathbb{R}^n$.

No que se refere a multiplicação por números reais, temos:

(M₁) Para qualquer $\lambda, \mu \in \mathbb{R}$ e $\mathbf{x} \in \mathbb{R}^n$ temos $\lambda(\mu\mathbf{x}) = (\lambda\mu)\mathbf{x}$.

(M₂) Para qualquer $\mathbf{x} \in \mathbb{R}^n$ temos $1\mathbf{x} = \mathbf{x}$.

Além disso, as duas propriedades distributivas satisfazem a compatibilidade da adição com a multiplicação:

(AM₁) Para qualquer $\lambda \in \mathbb{R}$ e $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ temos $\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$.

(AM₂) Para qualquer $\lambda, \mu \in \mathbb{R}$ e $\mathbf{x} \in \mathbb{R}^n$ temos $(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$.

Definição 2.3. Uma tripla $(V, +, \cdot)$ consistindo em um conjunto V , uma transformação (chamada adição)

$$\begin{aligned} + : \quad V \times V &\longrightarrow V \\ (\mathbf{x}, \mathbf{y}) &\longrightarrow \mathbf{x} + \mathbf{y} \end{aligned}$$

e outra transformação (denominada multiplicação por escalar)

$$\begin{aligned} \cdot : \quad \mathbb{R} \times V &\longrightarrow V \\ (\lambda, \mathbf{x}) &\longrightarrow \lambda\mathbf{x} \end{aligned}$$

é denominada um espaço vetorial real se os oito axiomas a seguir são válidos para as transformações $(+)$ e (\cdot) :

1. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ para quaisquer $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$.
2. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ para quaisquer $\mathbf{x}, \mathbf{y} \in V$.
3. Existe um elemento $\mathbf{0} \in V$ (chamado zero ou valor nulo), tal que $\mathbf{x} + \mathbf{0} = \mathbf{x}$ para qualquer $\mathbf{x} \in V$.
4. Para cada elemento $\mathbf{x} \in V$ existe um elemento $-\mathbf{x} \in V$ tal que $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.
5. $\lambda(\mu\mathbf{x}) = (\lambda\mu)\mathbf{x}$ para quaisquer $\lambda, \mu \in \mathbb{R}$ e $\mathbf{x} \in V$.
6. $1\mathbf{x} = \mathbf{x}$ para qualquer $\mathbf{x} \in V$.
7. $\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$ para quaisquer $\lambda \in \mathbb{R}$ e $\mathbf{x}, \mathbf{y} \in V$.

8. $(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$ para quaisquer $\lambda, \mu \in \mathbb{R}$ e $\mathbf{x} \in V$.

Proposição 2.4. *Em um espaço vetorial existe apenas um vetor zero (nulo).*

Demonstração.

Vamos supor, por absurdo, que exista $\mathbf{0}$ e $\mathbf{0}^*$. Daí, teremos após o uso dos Axiomas 2 e 3 da Definição 2.3 de espaços vetoriais reais as seguintes igualdades:

$$\mathbf{0} = \mathbf{0} + \mathbf{0}^* = \mathbf{0}^* + \mathbf{0} = \mathbf{0}^*.$$

Portanto, obrigatoriamente $\mathbf{0} = \mathbf{0}^*$. □

Proposição 2.5. *Em um espaço vetorial para cada \mathbf{x} existe somente um vetor $-\mathbf{x}$.*

Demonstração.

Consideremos $\mathbf{x} + \mathbf{a} = \mathbf{0}$ e $\mathbf{x} + \mathbf{b} = \mathbf{0}$. Daí, teremos após o uso dos Axiomas 1, 2 e 3 da Definição 2.3 de espaços vetoriais reais as seguintes igualdades:

$$\mathbf{a} = \mathbf{a} + \mathbf{0} = \mathbf{a} + (\mathbf{x} + \mathbf{b}) = (\mathbf{a} + \mathbf{x}) + \mathbf{b} = (\mathbf{x} + \mathbf{a}) + \mathbf{b} = \mathbf{0} + \mathbf{b} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Logo $\mathbf{a} = \mathbf{b}$. □

2.1.2 Exemplos de Espaços Vetoriais

- (i) $V = M_{2 \times 2}$ é o espaço vetorial das matrizes 2×2 .
- (ii) $V = P_n$ é o conjunto dos polinômios com coeficientes reais, de grau menor ou igual a n (incluindo o zero). As operações são soma de polinômios e a multiplicação destes por números reais.
- (iii) V é o conjunto das matrizes 2×2 , cujos elementos são números complexos. As operações são adição de matrizes e multiplicação destas por números complexos. Exemplificando:

$$\begin{bmatrix} 1+i & 0 \\ 3 & -5i \end{bmatrix} + \begin{bmatrix} 0 & 7 \\ 2+i+i & 5i \end{bmatrix} = \begin{bmatrix} 1+i & 7 \\ 5+i & 0 \end{bmatrix},$$

$$(1+i) \begin{bmatrix} -i & 1 \\ 0 & 1+i \end{bmatrix} = \begin{bmatrix} -1-i & 1-i \\ 0 & 2 \end{bmatrix}.$$

2.2 Subespaços Vetoriais

Sejam V um espaço vetorial e U um subconjunto não vazio de V . Dizemos que U é um subespaço vetorial de V , ou simplesmente subespaço de V , se U é um espaço vetorial com relação à adição e à multiplicação por escalar definidas em V .

Teorema 2.6. *Um subconjunto U não vazio, de um espaço vetorial V , é um espaço vetorial de V se, e somente se, estiverem satisfeitas as seguintes condições:*

- (1) $\forall \mathbf{u}, \mathbf{v} \in U$ tem-se $\mathbf{u} + \mathbf{v} \in U$.
- (2) $\forall \alpha \in \mathbb{R}, \forall \mathbf{u} \in U$ tem-se $\alpha \mathbf{u} \in U$.

Demonstração.

Seja $\mathbf{u} \in U$ e considerando a condição (2), isto é, $\alpha \mathbf{u} \in U, \forall \alpha \in \mathbb{R}$. Fazendo $\alpha = 0$, temos $0\mathbf{u} = \mathbf{0} \in U$. Além disso, quando $\alpha = -1$, temos $-1\mathbf{u} = -\mathbf{u} \in U$. Como as propriedades restantes são consequência de U ser um subconjunto não vazio de V , a demonstração é desnecessária. \square

2.2.1 Exemplos de Subespaços Vetoriais

- (i) O conjunto $\{\mathbf{0}\}$, constituído apenas do vetor nulo, e também todo o espaço V são subespaços de V . Estes dois subespaços de V são conhecidos como subespaços triviais e os demais subespaços, se existirem, são chamados subespaços próprios. O subespaço $\{\mathbf{0}\}$ é chamado de espaço vetorial nulo.
- (ii) No espaço vetorial das matrizes quadradas de ordem n , $M_{n \times n}$, os conjuntos das matrizes triangulares inferiores, das matrizes triangulares superiores e das matrizes diagonais, são subespaços vetoriais.
- (iii) $V = \mathbb{R}^2$, o plano, onde W é uma reta deste plano, que passa pela origem. W é um subespaço vetorial de V .

Teorema 2.7. *A intersecção de dois subespaços de um espaço vetorial V é um subespaço de V .*

Demonstração.

Sejam U_1 e U_2 subespaços de V . Inicialmente, podemos observar que $U_1 \cap U_2$ é não vazio, pois ambos contém o vetor nulo de V . Agora, consideremos $\mathbf{u}, \mathbf{v} \in U_1 \cap U_2$, isto é, \mathbf{u} e \mathbf{v} pertencem a U_1 , e também a U_2 . Daí, $\mathbf{u} + \mathbf{v} \in U_1$ e $\mathbf{u} + \mathbf{v} \in U_2$, pois U_1 e U_2 são subespaços de V , logo $\mathbf{u} + \mathbf{v} \in U_1 \cap U_2$. Além disso, consideremos $\alpha \in \mathbb{R}$,

tal que $\alpha \mathbf{u} \in U_1$ e $\alpha \mathbf{u} \in U_2$, logo $\alpha \mathbf{u} \in U_1 \cap U_2$, o que conclui a demonstração. \square

Observemos que o principal problema quando consideramos a união de subespaços é que se tomarmos um vetor em cada subespaço, a soma deles pode não pertencer à união. Seria, então, natural considerarmos o conjunto soma definido a seguir.

Definição 2.8. *Dados U e W subespaços de um espaço vetorial V , definimos a soma de U e W , denotada por $U + W$, como o conjunto*

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U \text{ e } \mathbf{w} \in W\}.$$

Com isso, quando somamos um elemento de um subespaço com um elemento do outro, automaticamente, a soma destes elementos está na soma dos subespaços.

Teorema 2.9. *Sejam U e W subespaços de um espaço vetorial V . Então, o conjunto $U + W = \{\mathbf{v} \in V \mid \mathbf{v} = \mathbf{u} + \mathbf{w}, \mathbf{u} \in U \text{ e } \mathbf{w} \in W\}$ é subespaço de V .*

Demonstração.

Consideremos $a \in \mathbb{R}$ e $\mathbf{v}_1, \mathbf{v}_2 \in U + W$. Como $\mathbf{v}_1, \mathbf{v}_2 \in U + W$, existem \mathbf{u}_1 e $\mathbf{u}_2 \in U$ e existem \mathbf{w}_1 e $\mathbf{w}_2 \in W$ tais que

$$\mathbf{v}_1 = \mathbf{u}_1 + \mathbf{w}_1 \text{ e } \mathbf{v}_2 = \mathbf{u}_2 + \mathbf{w}_2.$$

Então,

$$\mathbf{v}_1 + a\mathbf{v}_2 = (\mathbf{u}_1 + \mathbf{w}_1) + a(\mathbf{u}_2 + \mathbf{w}_2) = (\mathbf{u}_1 + a\mathbf{u}_2) + (\mathbf{w}_1 + a\mathbf{w}_2) \in U + W.$$

Assim, $U + W$ é subespaço de V .

2.3 Combinação Linear

Quando estamos em um espaço vetorial, é possível encontrarmos novos vetores a partir de vetores que já conhecemos.

Definição 2.10. *Sejam V um espaço vetorial real, $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$ e a_1, a_2, \dots, a_n números reais. Então, o vetor $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_n\mathbf{v}_n$ é um elemento de V ao que chamamos combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.*

Uma vez fixados os vetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ em V , o conjunto W de todos os vetores de V que são combinação linear destes, é um subespaço vetorial. W é chamado de subespaço gerado por $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ e usamos a notação $W = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$.

Note que, formalmente, podemos escrever

$$W = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] = \{\mathbf{v} \in V \mid \mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_n\mathbf{v}_n, a_i \in \mathbb{R}, 1 \leq i \leq n\}.$$

2.4 Dependência e Independência Linear

Os conjuntos $W_1 = [(1,0), (0,1), (2,1)]$ e $W_2 = [(1,0), (0,1)]$ representam o mesmo espaço gerado, isto se dá, pois o vetor $(2,1)$ é combinação linear dos vetores $(1,0)$ e $(0,1)$, ou seja, $(2,1) = 2(1,0) + 1(0,1)$. Portanto, o vetor $(2,1)$ é desnecessário para formarmos o subespaço W_1 , sendo suficiente apenas $(1,0)$ e $(0,1)$. Em álgebra linear, necessitaremos diversas vezes sabermos se existem vetores desnecessários, ou seja, se algum desses vetores é uma combinação linear dos outros. Para tal definiremos dependência e independência linear.

Definição 2.11. *Sejam V um espaço vetorial e $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$. Dizemos que os vetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ são linearmente independentes, se a equação $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_n\mathbf{v}_n = 0$ é satisfeita somente quando $a_1 = a_2 = \dots = a_n = 0$. Caso exista algum $a_i \neq 0$, dizemos que os vetores são linearmente dependentes.*

Teorema 2.12. *O conjunto $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ é linearmente dependente se, e somente se, um destes vetores for combinação linear dos outros.*

Demonstração.

Sejam $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ linearmente dependentes e $a_1\mathbf{v}_1 + \dots + a_j\mathbf{v}_j + \dots + a_n\mathbf{v}_n = 0$. De acordo com a definição, um dos coeficientes deve ser não nulo. Vamos supor que $a_j \neq 0$.

Logo

$$\mathbf{v}_j = -\frac{1}{a_j}(a_1\mathbf{v}_1 + \dots + a_{j-1}\mathbf{v}_{j-1} + a_{j+1}\mathbf{v}_{j+1} + \dots + a_n\mathbf{v}_n).$$

Portanto, \mathbf{v}_j é uma combinação linear dos outros vetores. Além disso, se tivermos $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ tal que para algum j , $\mathbf{v}_j = b_1\mathbf{v}_1 + \dots + b_{j-1}\mathbf{v}_{j-1} + b_{j+1}\mathbf{v}_{j+1} + \dots + b_n\mathbf{v}_n$, temos $b_1\mathbf{v}_1 + \dots + b_{j-1}\mathbf{v}_{j-1} + b_{j+1}\mathbf{v}_{j+1} + \dots + b_n\mathbf{v}_n - 1\mathbf{v}_j = 0$, e com isso, V é linearmente dependente. \square

2.5 Base de um Espaço Vetorial

Quando encontramos um conjunto finito de vetores de um espaço vetorial V , tal que todos os demais vetores de V sejam escritos como combinação linear dos vetores deste conjunto, estamos diante de uma base do espaço V .

Definição 2.13. Um conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ de vetores de V será uma base de V se:

(1) $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ é Linearmente Independente..

(2) $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] = V$.

Exemplos:

(a) $V = \mathbb{R}^2$; $V = [(1,0), (0,1)]$, $e_1 = (1,0)$ e $e_2 = (0,1)$ é conhecida como base canônica de \mathbb{R}^2 .

(b) $V = \mathbb{R}^3$; $V = [(1,0,0), (0,1,0), (0,0,1)]$, $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ e $e_3 = (0,0,1)$ é conhecida como base canônica de \mathbb{R}^3 .

(c) $V = M_{2 \times 2}$; $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ é uma base de V .

2.6 Dimensão de um Espaço Vetorial

Definição 2.14. O número de elementos de uma base de um espaço vetorial não nulo V de dimensão finita é chamado de dimensão de V e denotado por $\dim V$.

Teorema 2.15. Qualquer base de um espaço vetorial tem sempre o mesmo número de elementos.

Demonstração.

Consideremos $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ e $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ duas bases de V . Como $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ geram V e $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ são Linearmente Independentes, temos que $m \leq n$, pois qualquer conjunto Linearmente Independente de V tem no máximo n vetores. Além disso, analogamente, $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ geram V e $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ são Linearmente Independentes, então $n \leq m$, pois qualquer conjunto Linearmente Independente de V tem no máximo m vetores. Portanto $n = m$.

□

2.7 Produto Interno

Definição 2.16. Seja V um espaço vetorial. Um produto interno em V é uma função que a cada par de vetores \mathbf{u} e \mathbf{v} em V associa a um número real, denotado por $\langle \mathbf{u}, \mathbf{v} \rangle$, que satisfaz as seguintes condições:

Para quaisquer vetores \mathbf{u} , \mathbf{v} e \mathbf{w} de V e qualquer número real k ,

1. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$;
2. $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ se, e somente se, $\mathbf{v} = \mathbf{0}$;
3. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$;
4. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$;
5. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$.

Um espaço vetorial com um produto interno é chamado, abreviadamente, de espaço com produto interno.

Definição 2.17. (Norma de um vetor) *Seja V um espaço com produto interno. Definimos a norma do vetor \mathbf{v} de V , ou comprimento de \mathbf{v} , denotado por $\|\mathbf{v}\|$, como o número real*

$$\|\mathbf{v}\| = \langle \mathbf{v}, \mathbf{v} \rangle^{1/2}.$$

Se $\|\mathbf{v}\| = 1$, dizemos que \mathbf{v} é um vetor unitário.

A norma goza das seguintes propriedades:

1. $\|\mathbf{x}\| \geq 0$, valendo $\|\mathbf{x}\| = 0$ somente quando $\mathbf{x} = \mathbf{0}$;
2. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$;
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Definição 2.18. (Distância entre dois vetores) *A distância $d(\mathbf{u}, \mathbf{v})$ entre dois vetores \mathbf{u} e \mathbf{v} de V é definida como*

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}.$$

Definição 2.19. *Seja \mathbf{w} um vetor não-nulo de V . Se $\mathbf{v} \in V$, então a projeção de \mathbf{v} ao longo de \mathbf{w} é denotada por $\text{proj}_{\mathbf{w}}(\mathbf{v})$ e é definida por*

$$\text{proj}_{\mathbf{w}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w}.$$

2.8 Desigualdades

Teorema 2.20. (Desigualdade de Cauchy-Schwarz) *Se \mathbf{u} e \mathbf{v} são vetores de um espaço com produto interno V , então*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|,$$

com igualdade valendo se, e somente se, \mathbf{u} e \mathbf{v} são linearmente dependentes.

Demonstração.

A desigualdade é clara se \mathbf{u} é o vetor nulo de V . Suponhamos, então, \mathbf{u} diferente do vetor nulo. Para qualquer $t \in \mathbb{R}$, temos que $\langle t\mathbf{u} + \mathbf{v}, t\mathbf{u} + \mathbf{v} \rangle \geq 0$, ou seja, para qualquer $t \in \mathbb{R}$,

$$\langle \mathbf{u}, \mathbf{u} \rangle t^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle t + \langle \mathbf{v}, \mathbf{v} \rangle \geq 0. \quad (2-3)$$

Definamos $p(t) = \langle \mathbf{u}, \mathbf{u} \rangle t^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle t + \langle \mathbf{v}, \mathbf{v} \rangle$, $t \in \mathbb{R}$. Por (2-3), p é uma função polinomial não negativa. Além disso, como o coeficiente do termo quadrático é não negativo, segue que o discriminante Δ de $p(t)$ é um número real não positivo. Portanto,

$$\begin{aligned} \Delta &= 4\langle \mathbf{u}, \mathbf{v} \rangle^2 - 4\langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle \\ &= 4\langle \mathbf{u}, \mathbf{v} \rangle^2 - 4\|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \leq 0, \end{aligned}$$

o que equivale a

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

Extraindo a raiz quadrada em ambos os lados da desigualdade acima, obtemos o resultado desejado. \square

Proposição 2.21. (*Desigualdade Triangular*) Se \mathbf{u} e \mathbf{v} são vetores em um espaço V com produto interno e se $k \in \mathbb{R}$, então

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

Demonstração.

Temos

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v} \rangle| + \|\mathbf{v}\|^2, \end{aligned} \quad (2-4)$$

pois $x \leq |x|$ para todo $x \in \mathbb{R}$. Pelo Teorema 2.20, temos

$$\begin{aligned} \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v} \rangle| + \|\mathbf{v}\|^2 &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\| \|\mathbf{v}\| + \|\mathbf{v}\|^2 \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2. \end{aligned} \quad (2-5)$$

De (2-4) e (2-5), segue que

$$\|\mathbf{u} + \mathbf{v}\|^2 \leq (\|\mathbf{u}\| + \|\mathbf{v}\|)^2.$$

Extraindo as raízes quadradas de ambos os lados da desigualdade acima, obtemos a desigualdade desejada \square

2.9 Tópicos de Análise

2.9.1 Algumas Noções Topológicas: O espaço euclidiano n -dimensional

O espaço euclidiano n -dimensional \mathbb{R}^n é o produto cartesiano de n fatores iguais a \mathbb{R} : $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$. Seus elementos, portanto, são as sequências ou listas de n termos reais $x = (x_1, \dots, x_n)$. Para cada $i = 1, \dots, n$, o termo x_i chama-se a i -ésima coordenada de x . Se $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$, tem-se $x = y$ se, e somente se, $x_1 = y_1, \dots, x_n = y_n$. Assim, toda igualdade entre dois elementos de \mathbb{R}^n equivale a n igualdades entre números reais. Quando $n = 1$, \mathbb{R} é o conjunto dos números reais, \mathbb{R}^2 é o plano e \mathbb{R}^3 é o modelo do espaço euclidiano tridimensional. Os elementos de \mathbb{R}^n às vezes são chamados de *pontos* e às vezes *vetores*. Este segundo nome se aplica principalmente quando se considerarem entre eles as operações que aparecem nas Definições 2.1 e 2.2.

A revisão aqui apresentada está baseada nas referências [13] e [14].

2.9.2 Tipos Especiais de Conjuntos

Apresentaremos nesta subseção, definições que trazem algumas categorias especiais de conjuntos que serão úteis no desenvolver do trabalho.

Definição 2.22. (*Conjuntos Finitos*) Um conjunto X diz-se finito quando é vazio ou então existem $n \in \mathbb{N}$ e uma bijeção $f : I_n \rightarrow X$, onde $I_n = \{p \in \mathbb{N} \mid p \leq n\}$. Escrevendo $x_1 = f(1)$, $x_2 = f(2)$, \dots , $x_n = f(n)$ temos então $X = \{x_1, x_2, \dots, x_n\}$. A bijeção f chama-se contagem dos elementos de X e o número n chama-se o número de elementos do conjunto finito X .

Definição 2.23. (*Conjuntos Infinitos*) Diz-se que um conjunto é infinito quando não é finito. Assim, X é infinito quando não é vazio nem existe, seja qual for $n \in \mathbb{N}$, uma bijeção $F : I_n \rightarrow X$.

Definição 2.24. (*Conjuntos Enumeráveis*) Um conjunto X diz-se enumerável quando é finito ou quando existe uma bijeção $f : \mathbb{N} \rightarrow X$.

Definição 2.25. (Conjunto Complementar) Um conjunto $A^c \subseteq \mathbb{R}^n$ é dito complementar do conjunto A , quando é formado apenas por elementos de \mathbb{R}^n que não pertencem a A , isto é, $A^c = \mathbb{R}^n \setminus A$.

Definição 2.26. Uma bola aberta do \mathbb{R}^n é um conjunto de pontos da forma

$$B(a, r) = \{x \in \mathbb{R}^n \mid \|x - a\| < r\},$$

onde $a \in \mathbb{R}^n$ e $r > 0$.

Definição 2.27. Uma bola fechada do \mathbb{R}^n é um conjunto de pontos da forma

$$B[a, r] = \{x \in \mathbb{R}^n \mid \|x - a\| \leq r\},$$

onde $a \in \mathbb{R}^n$ e $r > 0$.

Por sua vez, a esfera de centro a e raio r é o conjunto

$$S[a; r] = \{x \in \mathbb{R}^n \mid \|x - a\| = r\}.$$

Definição 2.28. (Conjunto Limitado) Um conjunto $A \subset \mathbb{R}^n$ é um conjunto limitado se existir uma bola $B \subset \mathbb{R}^n$, aberta ou fechada, tal que $A \subseteq B$.

Definição 2.29. (Conjuntos Abertos) Diz-se que o ponto a é interior ao conjunto $X \subset \mathbb{R}^n$ quando, para algum $r > 0$, tem-se $B(a, r) \subset X$. O conjunto dos pontos interiores a X chama-se o interior do conjunto X e representa-se pela notação $\text{int}(X)$. Quando $a \in \text{int}(X)$ diz-se que o conjunto X é uma vizinhança de a . Um conjunto $A \subset \mathbb{R}^n$ chama-se aberto de A quando $A = \text{int}(A)$, isto é, quando todos os pontos de A são interiores a A . Ou ainda, se para cada $a \in A$ existir uma bola aberta $B(a, r)$ contida em A , com $r > 0$.

Observações:

- O conjunto vazio e o \mathbb{R}^n são subconjuntos abertos de \mathbb{R}^n ;
- A intersecção de dois conjuntos abertos também é um conjunto aberto;
- A união de uma família de conjuntos abertos é um conjunto aberto.

Teorema 2.30. (Intersecção e União de Conjuntos Abertos)

1. Se A_1 e A_2 são conjuntos abertos em \mathbb{R}^n então a intersecção $A_1 \cap A_2$ é um conjunto aberto.
2. Se (A_λ) é uma família qualquer de conjuntos abertos $A_\lambda \subset \mathbb{R}^n$, então a reunião $A = \bigcup_\lambda A_\lambda$ é um conjunto aberto.

Definição 2.31. A fronteira de um conjunto $X \subset \mathbb{R}^n$ é o conjunto $fr(X)$ formado pelos pontos de X que não são interiores a x , juntamente com os pontos de $\mathbb{R}^n - X$ que não são interiores a $\mathbb{R}^n - X$. De forma mais simples: tem-se $x \in fr(X)$ quando toda bola de centro x contém pontos de X e pontos de $\mathbb{R}^n - X$.

Exemplo 2.32. Seja $X = \{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$ o semi-plano superior fechado. Se $p = (a, b)$ com $b > 0$, então $p \in int(X)$. Com efeito, afirmamos que $B = B(p; b) \subset X$; veja a Figura 2.1.

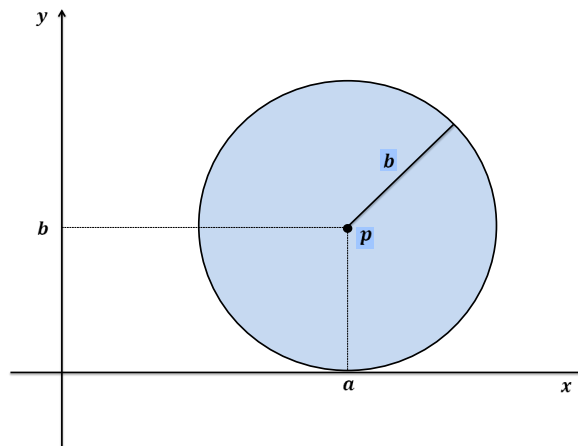


Figura 2.1: $X = \{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$, $p \in int(X)$.

Em termos mais precisos, argumentamos assim:

Se $(x, y) \in B$, então $\sqrt{(x-a)^2 + (y-b)^2} < b$. Daí, $(y-b)^2 < b^2$ e assim $y^2 - 2by + b^2 < b^2$. Logo, $y^2 < 2by$ e segue-se que $y > 0$ pois $b > 0$. Portanto, $(x, y) \in X$.

De forma análoga, todo ponto $\mathbb{R}^2 - X = \{(x, y) \in \mathbb{R}^2 \mid y < 0\}$ é um ponto interior, ou seja, $\mathbb{R}^2 - X$ é um conjunto aberto. Logo, nenhum ponto de $\mathbb{R}^2 - X$ pode estar na fronteira de X . Portanto, $fr(X) = \{(x, 0) \mid x \in \mathbb{R}\}$.

Definição 2.33. (Conjuntos Fechados) Diz-se que um ponto \bar{x} é aderente ao conjunto $X \subset \mathbb{R}^n$ quando \bar{x} é o limite de alguma sequência de pontos $x_k \in X$.

Chama-se fecho de um conjunto $X \subset \mathbb{R}^n$ ao conjunto $cl(X)$ formado por todos os pontos aderentes a X . Evidentemente, todo ponto $\bar{x} \in X$ é aderente a $cl(X)$. Tem-se que $X \subset cl(X)$. Se $X \subset Y$ então $cl(X) \subset cl(Y)$. Um conjunto X se diz fechado quando $X = cl(X)$, isto é, quando todo ponto aderente a X pertence a X . Em outras palavras, um conjunto $X \subseteq \mathbb{R}^n$ é dito **fechado** se o seu complementar for um conjunto aberto.

Teorema 2.34. (União e Intersecção de Conjuntos Fechados)

1. Se F_1 e F_2 são subconjuntos fechados de \mathbb{R}^n então $F_1 \cup F_2$ é um conjunto fechado.

2. Se (F_λ) é uma família qualquer de conjuntos fechados $(F_\lambda) \subset \mathbb{R}^n$, então a intersecção $F = \bigcap_\lambda F_\lambda$ é um conjunto fechado.

Teorema 2.35. Um ponto a é aderente ao conjunto X se, e somente se, toda bola de centro a contém algum ponto de X .

Corolário 2.36. O fecho de qualquer conjunto é um conjunto fechado.

Teorema 2.37.

1. Se F_1 e F_2 são conjuntos abertos então $F_1 \cup F_2$ é um conjunto aberto.
2. Se (F_λ) é uma família qualquer de conjuntos fechados, a intersecção $F = \bigcap_\lambda F_\lambda$ é um conjunto fechado.

Corolário 2.38. Os únicos conjuntos de \mathbb{R}^n que são simultaneamente fechados e abertos são o conjunto vazio e o \mathbb{R}^n .

Definição 2.39. (Ponto de Acumulação) Um ponto $\bar{x} \in \mathbb{R}^n$ é ponto de acumulação de um conjunto A , quando toda bola aberta de centro \bar{x} possui algum ponto de A , diferente de \bar{x} . Ou seja, para todo $\varepsilon > 0$, deve existir $x \in A$ tal que $0 < \|x - \bar{x}\| < \varepsilon$.

Teorema 2.40. (Teorema de Bolzano-Weierstrass) Toda sequência limitada em \mathbb{R}^n possui uma subsequência convergente.

Definição 2.41. (Conjunto Compacto) Um conjunto $X \subset \mathbb{R}^n$ chama-se compacto, quando X é limitado e fechado.

2.9.3 Princípio de Indução Matemática

Definição 2.42. (Axioma de Indução) Seja S um subconjunto de \mathbb{N} tal que

- (1) $0 \in S$.
- (2) S é fechado com respeito à operação "somar 1" a seus elementos, ou seja, para todo $n \in S$, tem-se $n+1 \in S$.

Então, $S = \mathbb{N}$.

Teorema 2.43. (Princípio de Indução Matemática) Seja $a \in \mathbb{N}$ e seja $p(n)$ uma sentença aberta em n ¹. Suponha que

¹Uma sentença aberta em n é uma frase de conteúdo matemático onde figura a letra n como palavra e que se torna uma sentença verdadeira ou falsa quando n é substituído por um número natural bem determinado.

(1) $p(a)$ é verdade, e que

(2) $\forall n \geq a, p(n) \Rightarrow p(n+1)$ é verdade, então, $p(n)$ é verdade para todo $n \geq a$.

Demonstração.

Seja $v = \{n \in \mathbb{N} \mid p(n)\}$; ou seja, v é o subconjunto dos elementos de \mathbb{N} para os quais $p(n)$ é verdade.

Considere o conjunto

$$S = \{m \in \mathbb{N} \mid a+m \in v\},$$

que verifica trivialmente $a+S \subset v$.

Como, pela condição (1), temos que $a+0 = a \in v$, segue-se que $0 \in S$.

Por outro lado, se $m \in S$, então $a+m \in v$ e, por (2), temos que $a+m+1 \in v$; logo $m+1 \in S$. Assim, pelo Axioma de Indução, temos que $S = \mathbb{N}$. Portanto,

$$\{m \in \mathbb{N} \mid m \geq a\} = a+\mathbb{N} \subset v,$$

o que prova o resultado. □

Para ilustrar o Primeiro Princípio de Indução Matemática, vamos determinar uma fórmula exata em função de $n \geq 1$ para a soma dos n primeiros números naturais ímpares. É importante mencionar que este exemplo constitui o primeiro registro da utilização deste princípio e foi desenvolvido por Francesco Maurolico em 1575 ². Assim, estamos em busca de uma fórmula para

$$S_n = 1 + 3 + 5 + \dots + (2n+1).$$

Vamos calcular S_n para alguns valores de n :

$$S_1 = 1, \quad S_2 = 4, \quad S_3 = 9, \quad S_4 = 16, \quad S_5 = 25.$$

Os casos particulares acima nos conduzem a conjecturar que $S_n = n^2$.

Então, definamos $p(n) = n^2$.

Temos que $p(1) = 1 = 1^2$, portanto verdade. Para provar que $p(n) \Rightarrow p(n+1)$ é verdade para todo $n \in \mathbb{N}$, basta mostrar que, se supusermos $p(n)$ verdade, então $p(n+1)$ é verdade, qualquer que seja $n \in \mathbb{N}$.

²Grande matemático e astrônomo grego que realizou contribuições nos campos da Geometria, Óptica, Cônica, Mecânica, Música e Astronomia. Em latim, Franciscus Maurolycus (1494-1575)

De fato, supondo $p(n)$ verdade, ou seja, $S_n = n^2$, e somando $2n + 1$ a ambos os lados desta última igualdade, obtemos:

$$S_{n+1} = S_n + 2n + 1 = n^2 + 2n + 1 = (n+1)^2,$$

o que nos diz que $p(n+1)$ é verdade.

Pelo Princípio da Indução Matemática, $p(n)$ é verdade para todo $n \in \mathbb{N}^*$.

Conjuntos Convexos

3.1 Um Pouco de História

A convexidade tem uma longa história. Na grande obra "Os Elementos" de Euclides (300 a.C.), aparecem várias contribuições ao assunto, relativas principalmente as propriedades dos polígonos e poliedros. Sem dúvida foi Arquimedes (287-212 a.C.), em seu livro "Sobre a Esfera e o Cilindro", o primeiro a definir o que se entendia por uma curva ou superfície convexa. Entre as diferentes propriedades obtidas por Arquimedes sobre convexidade, merecem maior destaque os postulados e resultados referentes ao centro de gravidade de conjuntos planos e sua descrição dos 13 poliedros semirregulares, também 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. Posteriormente, os sólidos arquimedianos foram redescobertos por Kepler (1571 - 1630), que em seu livro "Harmonices Mundi" (1619), demonstrou que, efetivamente, só podiam existir 13 tipos diferentes. Resultados referentes ao famoso problema isoperimétrico, cuja origem data de cerca de 810 a.C., o devido a Zenodorus (200 a. C.) parece ser o mais importante: ele mostra que, entre todos os n -ângulos convexos de mesmo perímetro, o regular é o que possui maior área. Além de outras contribuições esporádicas, no final do século XIX apareceram diversos resultados de grande importância em convexidade, graças a matemáticos como Brunn e Minkowski; contudo, o interesse real pela geometria convexa é relativamente recente, pois um primeiro estudo sistemático foi encontrado em 1934, no livro de Bonnesen e Fenchel "Theorie der Konvexen Körper". Ao longo dos anos 40 e 50 se descobriram numerosas aplicações importantes para conjuntos convexos, principalmente no campo da Otimização Geométrica, que despertou o interesse por esta teoria. A fundamentação teórica deste capítulo está baseada, principalmente, nas referências [10, 2, 4, 6].

3.2 Conjuntos Convexos

Definição 3.1. (*Conjunto convexo*) Um conjunto $D \subset \mathbb{R}^n$ é dito convexo quando dados dois pontos x_1 e $x_2 \in D$, então o ponto genérico $x = \lambda x_1 + (1 - \lambda)x_2 \in D$ para cada $\lambda \in [0, 1]$.

Em outras palavras, podemos dizer que um conjunto é convexo se cada ponto no conjunto puder ser visto por todos os outros pontos, ao longo de uma trajetória retilínea desobstruída entre eles, onde o meio desobstruído está no conjunto. Todo conjunto afim é também convexo, pois ele contém as retas inteiras entre quaisquer dois pontos distintos nele, e, portanto, o segmento de reta entre os pontos.

Observe o conjunto $X \subset \mathbb{R}^2$ da forma $X = \{(x, y) \in \mathbb{R}^2 \mid y \leq x^2\}$. Veja que X não é convexo, pois se tomarmos, por exemplo, os pontos $a = (-1, 1)$ e $b = (1, 1) \in X$, temos que,

$$\frac{1}{2}a + \frac{1}{2}b = (0, 1) \notin X.$$

Geometricamente, podemos visualizar a Definição 3.1 na Figura 3.1.

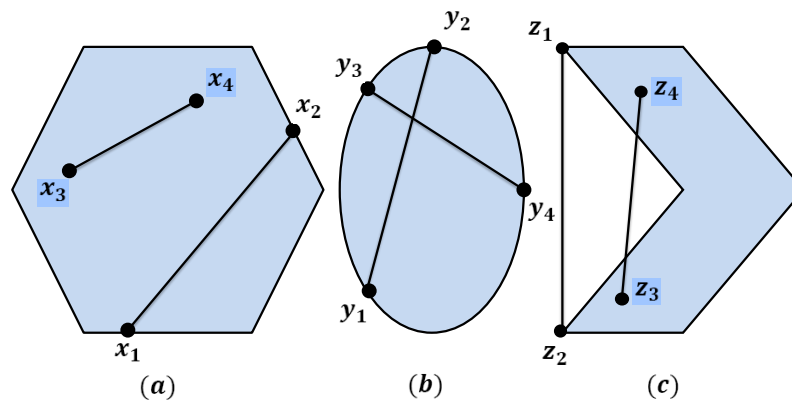


Figura 3.1: (a) e (b) conjuntos convexos; (c) Conjunto não-convexo.

Definição 3.2. (*Combinação Linear*) Dizemos que um vetor $x \in \mathbb{R}^n$ é uma combinação linear dos vetores x_1, \dots, x_k se existem escalares $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ adequados, tais que $x = \lambda_1 x_1 + \dots + \lambda_k x_k$. Além disso,

- Se $\sum_{i=1}^k \lambda_i = 1$, então dizemos que x é uma combinação afim dos vetores x_i .

- Se $\lambda_i \geq 0$ para todo $1 \leq i \leq k$, então dizemos que x é uma combinação positiva dos vetores x_1, \dots, x_k .

Definição 3.3. (Combinação Convexa) Dado $x_i \in \mathbb{R}^n$, $\alpha_i \in [0, 1]$, $i = 1, \dots, p$, tais que $\sum_{i=1}^p \alpha_i = 1$, o ponto $\sum_{i=1}^p \alpha_i x_i$ denomina-se combinação convexa dos pontos $x_i \in \mathbb{R}^n$ com parâmetros α_i , $i = 1, \dots, p$.

Em outras palavras, uma combinação é convexa quando ela é concomitantemente combinação afim e positiva.

Exemplo 3.4. Podemos representar uma combinação convexa de vetores do \mathbb{R}^2 da seguinte forma. Tomemos dois vetores $v_1 = (2, 0)$ e $v_2 = (5, 0)$ e em seguida escrevemos $v = \lambda v_1 + (1 - \lambda)v_2$, com $0 \leq \lambda \leq 1$. Repare que o vetor $v \in \mathbb{R}^2$, escrito desta maneira, obedece à regra de convexidade, pois se fizermos λ variar entre 0 e 1, v será representado por pontos situados no segmento que une as extremidades de v_1 e v_2 . A Figura 3.2 abaixo ilustra a combinação convexa deste exemplo.

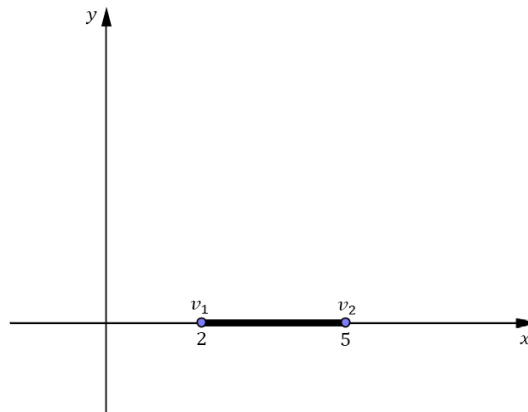


Figura 3.2: Combinação convexa

Definição 3.5. (Conjunto Afim) Um conjunto $A \subseteq \mathbb{R}^n$ é afim se a reta que une quaisquer dois pontos distintos em A estiver em A , isto é, se para algum $x_1, x_2 \in A$ onde $\lambda \in \mathbb{R}$, tem-se que $\lambda x_1 + (1 - \lambda)x_2 \in A$. Em outras palavras, A contém a combinação linear de quaisquer dois pontos em A , contanto que a soma dos coeficientes na combinação linear seja igual a 1.

São exemplos de conjuntos afins: retas, subespaços e hiperplano, sendo que este último definiremos ainda neste capítulo.

Segue abaixo alguns resultados importantes sobre conjunto convexo.

Lema 3.6. Qualquer subespaço de \mathbb{R}^n é um conjunto convexo.

Demonstração.

Seja D um subespaço do \mathbb{R}^n . Se $D = \emptyset$ ou $D = \mathbb{R}^n$ ou ainda se D é unitário, a conclusão é imediata, pois são conjuntos convexos triviais.

Caso contrário, se $x \in D$, $y \in D$ e $\alpha \in [0, 1]$, então

$$x + y \in D, \text{ e } \alpha x \in D,$$

pois D é um subespaço do \mathbb{R}^n .

Com isso, $\alpha x + (1 - \alpha)y = \alpha(x - y) + y \in D$, já que $\alpha(x - y) \in D$ e $y \in D$. \square

Conjuntos convexos aparecem na matemática com muita frequência. Um importante exemplo aparece na continuação.

Lema 3.7. *A solução de um sistema de equações lineares $S = \{x \in \mathbb{R}^n \mid Ax = b\}$, onde $A \in \mathbb{R}^{m \times n}$ e $b \in \mathbb{R}^m$, é um conjunto convexo.*

Demonstração.

De fato, sejam $x_1, x_2 \in S$. Então, para todo $\lambda \in [0, 1]$, temos:

$$\begin{aligned} A[\lambda x_1 + (1 - \lambda)x_2] &= A\lambda x_1 + A(1 - \lambda)x_2 \\ &= \lambda Ax_1 + (1 - \lambda)Ax_2 \\ &= \lambda b + (1 - \lambda)b \\ &= \lambda b - \lambda b + b = b. \end{aligned}$$

Portanto, $Ax = b$ com $x = \lambda x_1 + (1 - \lambda)x_2 \in S$. O que conclui a demonstração. \square

3.3 Propriedades dos Conjuntos Convexos

O Teorema abaixo afirma que a interseção de conjuntos convexos, ainda é um conjunto convexo.

Teorema 3.8. *Suponha que D_1, D_2, \dots, D_q são conjuntos convexos do \mathbb{R}^n . Então,*

$$D = D_1 \cap D_2 \cap \dots \cap D_q$$

é um conjunto convexo.

Demonstração.

Se a interseção é vazia, então D é convexo por definição. Caso contrário, fixemos arbitrariamente $x, y \in D$.

Tomemos $\lambda_1, \lambda_2 \in [0, 1]$, com $\lambda_1 + \lambda_2 = 1$. Considere $i = 1, 2, \dots, q$. Uma vez que $x, y \in D$, segue-se pela definição de interseção de conjuntos que $x, y \in D_i$, para cada i . Logo, para cada i , $z = \lambda_1 x + \lambda_2 y \in D_i$, pela convexidade de D_i . Segue-se, novamente pela definição de interseção de conjuntos, que $z \in D$. Pela arbitrariedade de $x, y \in D$, D é um conjunto convexo. \square

Apesar da interseção de conjuntos convexos preservar a convexidade, o mesmo não ocorre com a união, o que é bastante óbvio. O exemplo seguinte possibilita uma melhor compreensão.

Exemplo 3.9. *Exemplo: A representação dos conjuntos*

$$\varepsilon_1 = \left\{ (x, y) \in \mathbb{R}^2 \mid \frac{x^2}{4} + y^2 = 1 \right\} \text{ e}$$

$$\varepsilon_2 = \left\{ (x, y) \in \mathbb{R}^2 \mid (x-5)^2 + \frac{y^2}{4} = 1 \right\}$$

no \mathbb{R}^2 ajuda a visualizar graficamente que a união de conjuntos convexos nem sempre é um conjunto convexo. Conforme podemos observar na Figura 3.3 abaixo.

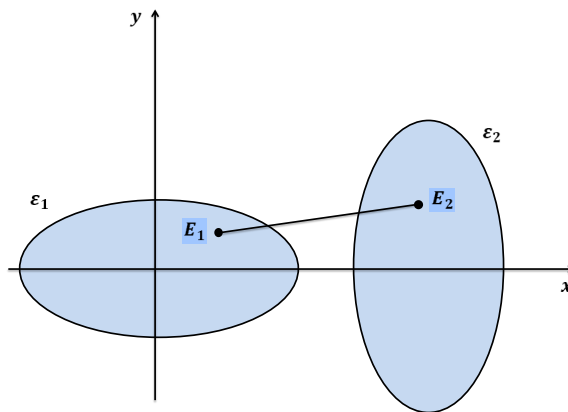


Figura 3.3: *União de conjuntos convexos*

Definição 3.10. *(Soma/Subtração de Conjuntos Convexos)* Sejam D_1 e D_2 dois conjuntos convexos. O conjunto soma/subtração destes conjuntos é um conjunto de pontos da forma

$$D_1 \pm D_2 = \{x \in \mathbb{R}^n \mid x = d_1 \pm d_2, d_1 \in D_1, d_2 \in D_2\}.$$

Lema 3.11. *Sejam D_1 e D_2 conjuntos convexos de \mathbb{R}^n , então $D_1 \pm D_2$ é um conjunto convexo.*

Demonstração.

Se $D_1 = \emptyset$ ou $D_2 = \emptyset$, o resultado é imediato, pois $D_1 + \emptyset = D_1, \forall D_1 \in \mathbb{R}^n$. Por outro lado, se tivermos $D_1 \neq \emptyset$ e $D_2 \neq \emptyset$ e, além disso, $x, y \in D_1 + D_2$ e $\lambda \in [0, 1]$. Teremos, com base na Definição 3.10, $x = x_1 + x_2$, com $x_1 \in D_1$ e $x_2 \in D_2$, e $y = y_1 + y_2$, com $y_1 \in D_1$ e $y_2 \in D_2$. Consideremos $z = (1 - \lambda)x + \lambda y$, e com isso escrevemos

$$\begin{aligned} z &= (1 - \lambda)x + \lambda y \\ &= (1 - \lambda)(x_1 + x_2) + \lambda(y_1 + y_2) \\ &= (1 - \lambda)x_1 + \lambda y_1 + (1 - \lambda)x_2 + \lambda y_2 \end{aligned}$$

Definindo, $z_1 = (1 - \lambda)x_1 + \lambda y_1 \in D_1$ e $z_2 = (1 - \lambda)x_2 + \lambda y_2 \in D_2$, obtemos

$$z = z_1 + z_2.$$

Portanto, $D_1 \pm D_2$ é um conjunto convexo. □

Definição 3.12. *Seja D um conjunto convexo e $\alpha \in \mathbb{R}$ um escalar arbitrário. A multiplicação de um escalar por um conjunto convexo é definida por*

$$\alpha D = \{x \in \mathbb{R}^n \mid x = \alpha y; y \in D, \alpha \in \mathbb{R}\}.$$

Lema 3.13. *A multiplicação de um conjunto convexo por um escalar também é um conjunto convexo.*

Demonstração.

Quando $D = \emptyset$, o resultado é óbvio. Vejamos então, o caso de $D \neq \emptyset$. Consideremos $x, y \in D$ e $\lambda \in [0, 1]$. Escrevendo $x = \alpha d_1$ e $y = \alpha d_2$, sendo que, $d_1, d_2 \in D$. Com isso, temos $z = (1 - \lambda)x + \lambda y$ e daí,

$$\begin{aligned} z &= (1 - \lambda)(\alpha d_1) + \lambda(\alpha d_2) \\ &= \alpha[(1 - \lambda)d_1 + \lambda d_2] \\ &= \alpha d, \end{aligned}$$

onde $d \in D$. Portanto, αD é um conjunto convexo. □

Lema 3.14. *Considere D um conjunto convexo e $A \in \mathbb{R}^{m \times n}$ uma matriz arbitrária. O conjunto da forma $A(D) = \{y \in \mathbb{R}^m \mid y = Ax, x \in D\}$ é convexo.*

Demonstração.

De fato, sejam y_1 e $y_2 \in A(D)$, isto é, $y_1 = Ax_1$ e $y_2 = Ax_2$, com x_1 e $x_2 \in D$. Com isso, considerando $\lambda \in [0, 1]$, temos:

$$\begin{aligned}\lambda y_1 + (1 - \lambda)y_2 &= \lambda Ax_1 + (1 - \lambda)Ax_2 \\ &= A[\lambda x_1 + (1 - \lambda)x_2] \\ &= Ax.\end{aligned}$$

□

Portanto, pela convexidade do conjunto D , temos que $A(D)$ é convexo.

Lema 3.15. *Considere D um conjunto convexo e $A \in \mathbb{R}^{n \times m}$ uma matriz arbitrária. O conjunto da forma $A^{-1}(D) = \{x \in \mathbb{R}^m \mid Ax \in D\}$ é convexo.*

Demonstração.

Consideremos, arbitrariamente, dois elementos x_1 e $x_2 \in A^{-1}(D)$, ou seja, Ax_1 e $Ax_2 \in D$. Pela convexidade de D , temos que $\lambda Ax_1 + (1 - \lambda)Ax_2 \in D$. Logo,

$$A\lambda x_1 + A(1 - \lambda)x_2 = A[\lambda x_1 + (1 - \lambda)x_2].$$

□

Com isso, podemos concluir que $A^{-1}(D)$ é um conjunto convexo.

Lema 3.16. *Toda bola aberta, dada pela Definição 2.26, ou fechada, dada pela Definição 2.27, em \mathbb{R}^n é um conjunto convexo.*

Demonstração.

De fato, sejam x_1 e $x_2 \in B[a, r]$ e $\lambda \in [0, 1]$, ou seja, $\|x_1 - a\| \leq r$ e $\|x_2 - a\| \leq r$. Assim,

$$\begin{aligned}\|\lambda x_1 + (1 - \lambda)x_2 - a\| &= \|\lambda(x_1 - a) + (1 - \lambda)(x_2 - a)\| \\ &\leq \lambda\|x_1 - a\| + (1 - \lambda)\|x_2 - a\| \\ &\leq \lambda r + (1 - \lambda)r = r.\end{aligned}$$

Mostrando, com isso, que $B[a, r]$ é um conjunto convexo. A demonstração para o caso $B(a, r)$, isto é, de uma bola aberta é análogo. □

Lema 3.17. *O conjunto $E = \{x \in \mathbb{R}^n \mid \|x\| = c\}$, onde $c \in \mathbb{R}_+$, é convexo se, e somente se, $c = 0$.*

Em particular, uma esfera não-trivial, isto é, com $c > 0$ é não-convexa.

Demonstração.

De fato, sejam x_1 e $x_2 \in E$ e $\lambda \in [0, 1]$, ou seja, $\|x_1\| = c$ e $\|x_2\| = c$. Assim, pela Definição 3.16 temos que: $\|\lambda x_1 + (1 - \lambda)x_2\| \leq c$, com a igualdade acontecendo somente se $c = 0$. Caso $c > 0$, a não convexidade é imediata. \square

Definição 3.18. *Seja $A \subseteq \mathbb{R}^n$. Diz-se que $a \in A$ é ponto interior de A se existir algum $r > 0$ tal que $B(a, r) \subseteq A$. Ao conjunto de todos os pontos interiores do conjunto A chama-se interior de A e denota-se por $\text{int}(A)$.*

A seguir mostraremos que o interior de um conjunto convexo, dado na Definição 3.18 e o fecho deste conjunto, dado na Definição 2.33 são conjuntos convexos.

Proposição 3.19. *Seja $D \subset \mathbb{R}^n$ um conjunto convexo. Então $\text{cl}(D)$ e $\text{int}(D)$ são conjuntos convexos.*

Demonstração.

Sejam $x \in \text{cl}(D)$, $y \in \text{cl}(D)$, $\alpha \in [0, 1]$. Escolhemos sequências $\{x^k\} \subset D$ e $\{y^k\} \subset D$ tais que $\{x^k\} \rightarrow x$ e $\{y^k\} \rightarrow y$ quando $k \rightarrow \infty$. Pela convexidade de D , $\alpha x^k + (1 - \alpha)y^k \in D$ para todo k . Portanto,

$$\alpha x + (1 - \alpha)y = \lim_{k \rightarrow \infty} (\alpha x^k + (1 - \alpha)y^k) \in \text{cl}(D),$$

isto é, $\text{cl}(D)$ é convexo.

Sejam $x \in \text{int}(D)$ e $y \in \text{int}(D)$. Fixamos $\varepsilon > 0$ tal que $B(x, \varepsilon) \subset D$ e $B(y, \varepsilon) \subset D$. Seja $\alpha \in [0, 1]$. Para mostrar que $\alpha x + (1 - \alpha)y \in \text{int}(D)$, verificamos que $B(\alpha x + (1 - \alpha)y, \varepsilon) \subset D$.

Seja $z \in B(\alpha x + (1 - \alpha)y, \varepsilon)$, isto é, existe $q \in \mathbb{R}^n$ tal que

$$z = \alpha x + (1 - \alpha)y + q,$$

onde $\|q\| \leq \varepsilon$. Observamos que

$$z = \alpha(x + q) + (1 - \alpha)(y + q).$$

Obviamente, $x + q \in B(x, \varepsilon) \subset D$ e $y + q \in B(y, \varepsilon) \subset D$. Pela convexidade do conjunto D , concluímos que $z \in D$. acabamos de mostrar que $B(\alpha x + (1 - \alpha)y, \varepsilon) \subset D$. Mas isto significa que $\alpha x + (1 - \alpha)y \in \text{int}(D)$. \square

Proposição 3.20. *Sejam $D_i \subset \mathbb{R}^n$, $i = 1, 2$, conjuntos convexos fechados. Se um deles também é limitado, então o conjunto $D_1 + D_2$ é convexo e fechado.*

Demonstração.

Já mostramos no Lema 3.11 que $D = D_1 + D_2$ é convexo. Mostraremos que D é fechado, supondo que D_2 seja limitado.

Seja $\{x^k\} \rightarrow x$ quando $k \rightarrow \infty$, $x^k \in D$ para todo k . Precisamos verificar que $x \in D$. Como $x^k \in D$, para todo k existem então $x^{k,i} \in D_i$, $i = 1, 2$, tais que $x^k = x^{k,1} + x^{k,2}$. É claro que a sequência $\{x^{k,2}\}$ é limitada, pois D_2 é limitado. Como a sequência $\{x^{k,1} + x^{k,2}\}$ converge, segue que $\{x^{k,1}\}$ também é limitada. Portanto, $\{x^{k,1}\}$ tem um ponto de acumulação, digamos \bar{x}^1 . Seja $\{x^{k_i,1}\} \rightarrow \bar{x}^1$ quando $i \rightarrow \infty$. Podemos admitir que $\{x^{k_i,2}\} \rightarrow \bar{x}^2$ quando $i \rightarrow \infty$, escolhendo uma subsequência de $\{k_i\}$, se for necessário. Como D_1 e D_2 são fechados, $\bar{x}^1 \in D_1$ e $\bar{x}^2 \in D_2$. Portanto,

$$\begin{aligned} x = \lim_{k \rightarrow \infty} x^k &= \lim_{i \rightarrow \infty} x^{k_i} = \lim_{i \rightarrow \infty} (x^{k_i,1} + x^{k_i,2}) \\ &= \bar{x}^1 + \bar{x}^2 \in D_1 + D_2, \end{aligned}$$

o que mostra que D é fechado. □

Observamos que para D_1 e D_2 convexos e fechados é possível que $D_1 + D_2$ não seja fechado; conforme podemos perceber no exemplo a seguir.

Exemplo 3.21. *Consideremos os dois conjuntos*

$$D_1 = \{x \in \mathbb{R}^2 \mid x_2 = 0\} \text{ e } D_2 = \{x \in \mathbb{R}^2 \mid x_1 > 0, x_2 \geq 1/x_1\}$$

ilustrados na Figura 3.4. Veja que os dois conjuntos D_1 e D_2 assim definidos são convexos e fechados, mas nenhum é limitado.

Para $x = x^1 + x^2$, onde $x^i \in D^i$, $i = 1, 2$, tem-se que $x_2 > 0$ sempre, enquanto x_1 pode assumir um valor qualquer. Temos então que $D_1 + D_2 = \{x \in \mathbb{R}^2 \mid x_2 > 0\}$, um semiespaço aberto.

Teorema 3.22. *Um conjunto $D \subset \mathbb{R}^n$ é convexo se, e somente se, para quaisquer $p \in \mathbb{N}$, $x^i \in D$ e $\alpha_i \in [0, 1]$, $i = 1, \dots, p$, tais que $\sum_{i=1}^p \alpha_i = 1$, a combinação convexa $\sum_{i=1}^p \alpha_i x^i$ pertence a D .*

Demonstração.

Se D contém todas as combinações convexas de seus pontos, ele contém em particular as combinações convexas de quaisquer dois pontos, isto é, para $p = 2$. Logo, pela Definição 3.1, D é convexo.

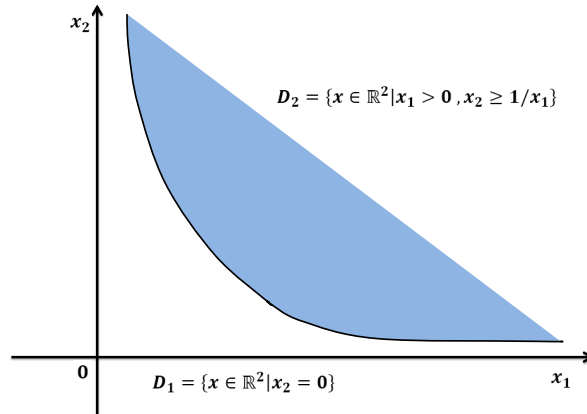


Figura 3.4: Um exemplo que mostra que a soma de conjuntos convexos fechados pode não ser fechada: $D_1 + D_2 = \{x \in \mathbb{R}^2 \mid x_2 > 0\}$.

Suponhamos agora que D seja convexo. A prova é feita por indução em relação ao número de pontos $p \in \mathbb{N}$ e $x^i \in D$, $\alpha_i \in [0, 1]$, $i = 1, \dots, p$ tais que $\sum_{i=1}^p \alpha_i = 1$, definimos $x = \sum_{i=1}^p \alpha_i x^i$.

Se $p = 1$, tem-se $\alpha_1 = 1$ e, portanto, $x = x^1 \in D$.

Suponhamos que qualquer combinação convexa de quaisquer $j \geq 1$ pontos de D pertença a D , e consideremos o caso de $p = j + 1$ pontos.

Se $\alpha_{j+1} = 1$, então $\alpha_i = 0$ para todo $i = 1, \dots, j$. Neste caso, $x = x^{j+1} \in D$.

Seja $\alpha_{j+1} \in [0, 1)$. Como $1 - \alpha_{j+1} > 0$, podemos escrever

$$\begin{aligned} x &= \sum_{i=1}^{j+1} \alpha_i x^i \\ &= (1 - \alpha_{j+1}) \sum_{i=1}^j \frac{\alpha_i}{1 - \alpha_{j+1}} x^i + \alpha_{j+1} x^{j+1} \\ &= (1 - \alpha_{j+1}) y + \alpha_{j+1} x^{j+1}, \end{aligned} \tag{3-1}$$

onde

$$y = \sum_{i=1}^j \beta_i x^i,$$

$\beta_i = \frac{\alpha_i}{1 - \alpha_{j+1}} \geq 0$, $i = 1, \dots, j$. Como $1 = \sum_{i=1}^{j+1} \alpha_i = \sum_{i=1}^j \alpha_i + \alpha_{j+1}$, segue-se que

$$\begin{aligned} \sum_{i=1}^j \beta_i &= (1 - \alpha_{j+1})^{-1} \sum_{i=1}^j \alpha_i \\ &= (1 - \alpha_{j+1})^{-1} (1 - \alpha_{j+1}) = 1. \end{aligned}$$

Portanto, y é uma combinação convexa de j pontos de D . Pela hipótese de indução, $y \in D$. Agora, a Equação (3-1) mostra que x é uma combinação convexa de dois pontos de D , em particular, y e x^{j+1} . Como D é convexo, obtemos que $x \in D$, o que completa a prova. \square

Teorema 3.23. (Teorema de Carathéodory) *Seja $x \in \mathbb{R}^n$ uma combinação convexa de pontos do conjunto $D \subset \mathbb{R}^n$. Então existem $x^i \in D$ e $\alpha_i \in \mathbb{R}_+$, $i = 1, \dots, n+1$, tais que $x = \sum_{i=1}^{n+1} \alpha_i x^i$, $\sum_{i=1}^{n+1} \alpha_i = 1$.*

3.4 Fecho Convexo

Definição 3.24. (Fecho Convexo) *Seja $D \subset \mathbb{R}^n$. O fecho convexo ou envoltória convexa de D , representada como $\text{conv}(D)$, é o menor conjunto convexo que contém D . Ou, equivalentemente, a interseção de todos os conjuntos convexos em \mathbb{R}^n que contém D .*

Lema 3.25. *Sejam A e B dois conjuntos não-vazios de \mathbb{R}^n , temos que*

$$\text{conv}(A+B) = \text{conv}(A) + \text{conv}(B).$$

Demonstração.

Vejamos primeiramente que $\text{conv}(A+B) \subset \text{conv}(A) + \text{conv}(B)$:

Dado $x \in \text{conv}(A+B)$, existem $a_1, a_2, \dots, a_k \in A$, $b_1, b_2, \dots, b_k \in B$ e $\lambda_k \geq 0$ com $\sum_{i=1}^k \lambda_i = 1$ e $\sum_{i=1}^k \lambda_i (a_i + b_i) = x$.

Assim,

$$x = \sum_{i=1}^k \lambda_i (a_i + b_i) = \sum_{i=1}^k \lambda_i a_i + \sum_{i=1}^k \lambda_i b_i \in \text{conv}(A) + \text{conv}(B).$$

Reciprocamente, vejamos que pela Definição 3.10, $x + y \in \text{conv}(A+B) \forall x \in \text{conv}(A), y \in \text{conv}(B)$:

Existem $a_1, \dots, a_k \in A$, $b_1, \dots, b_k \in B$, $\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_m \geq 0$ com

$$\sum_{i=1}^k \lambda_i = \sum_{j=1}^m \mu_j = 1, \quad x = \sum_{i=1}^k \lambda_i a_i \quad \text{e} \quad y = \sum_{j=1}^m \mu_j b_j.$$

Assim, considerando a Definição 3.24, temos que:

$$\begin{aligned} x+y &= \left(\sum_{j=1}^m \mu_j \right) \left(\sum_{i=1}^k \lambda_i a_i \right) + \left(\sum_{i=1}^k \lambda_i \right) \left(\sum_{j=1}^m \mu_j b_j \right) \\ &= \sum_{i=1, \dots, k} \sum_{j=1, \dots, m} \lambda_i \mu_j (a_i + b_j). \end{aligned}$$

Repare que o membro da direita é uma combinação linear convexa de $a_i + b_i \in A + B$, já que os coeficientes $\lambda_i \mu_j$ são não negativos e somam 1. Portanto, $x + y \in \text{conv}(A + B)$ o que conclui a demonstração. \square

A Figura 3.5 é um exemplo simplório de um fecho convexo.

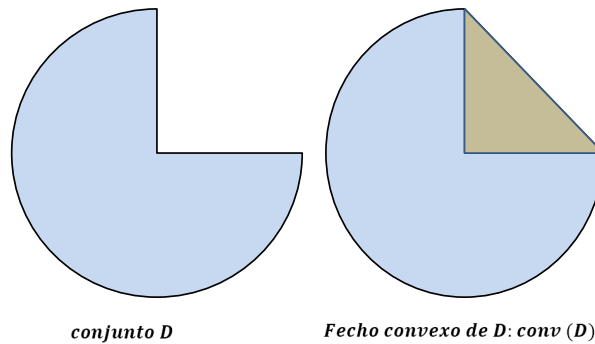


Figura 3.5: fecho convexo

Como um outro exemplo, podemos afirmar que o fecho da esfera é a bola, isto é, para

$$D = \{x \in \mathbb{R}^n \mid \|x\| = c\},$$

onde $c \in \mathbb{R}$, tem-se

$$\text{conv}(D) = \{x \in \mathbb{R}^n \mid \|x\| \leq c\}.$$

E ainda, como trata-se de uma interseção de conjuntos, é claro que $\text{conv}(D)$ é um conjunto convexo para qualquer $D \subset \mathbb{R}^n$. Também é óbvio que se D é convexo, então $\text{conv}(D) = D$.

Para ilustrarmos ainda mais a definição de fecho convexo, considere o conjunto $\Omega \subset \mathbb{R}^2$, no plano Euclidiano, definido por:

$$\Omega := \{(-1, 0)\} \cup \{(x_1, x_2) \mid x_1^2 + x_2^2 = 1, x_1 \geq 0\}.$$

O gráfico que representa este conjunto é mostrado na Figura 3.6.

3.5 Ponto Extremo

Definição 3.26. (*Ponto Extremo*) Diz-se que $z \in \mathbb{R}^n$ é um ponto extremo de D se não existem $x, y \in D$, $x \neq y$ tais que $z = \alpha x + (1 - \alpha)y$ para algum $\alpha \in (0, 1)$, isto é, z

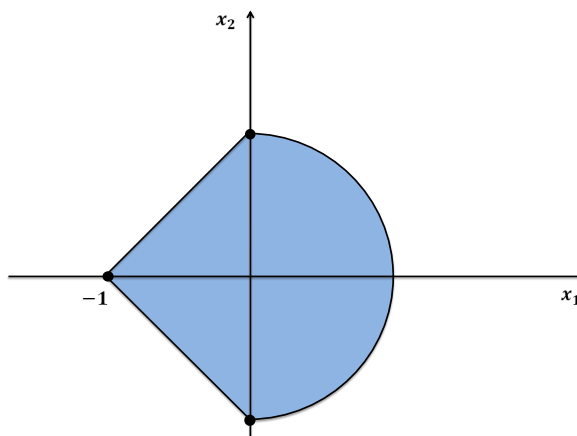


Figura 3.6: $\Omega := \{(-1, 0)\} \cup \{(x_1, x_2) \mid x_1^2 + x_2^2 = 1, x_1 \geq 0\}$

não pode ser escrito como combinação linear convexa de quaisquer dois pontos distintos de D .

Na figura 3.7, apenas x_1 é ponto extremo do conjunto representado.

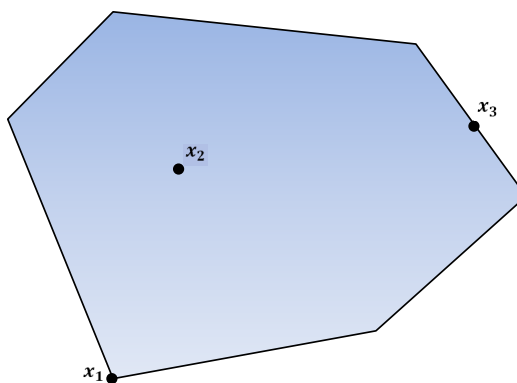


Figura 3.7: Exemplo de ponto extremo de um conjunto

3.6 Hiperplano

Definição 3.27. Um hiperplano do \mathbb{R}^n é um conjunto de pontos da forma

$$\mathcal{H} = \{x \in \mathbb{R}^n \mid \langle a, x \rangle = c\},$$

onde $a \in \mathbb{R}^n$ e $c \in \mathbb{R}$.

Lema 3.28. Todo hiperplano em \mathbb{R}^n é um conjunto convexo.

Demonstração.

Considere $x_1, x_2 \in \mathcal{H}$ e $\lambda \in [0, 1]$. Queremos mostrar que $z = \lambda x_1 + (1 - \lambda)x_2 \in \mathcal{H}$.
Ora,

$$\begin{aligned} \langle a, z \rangle &= \langle a, \lambda x_1 + (1 - \lambda)x_2 \rangle \\ &= \langle a, \lambda x_1 \rangle + \langle a, (1 - \lambda)x_2 \rangle \\ &= \lambda \langle a, x_1 \rangle + (1 - \lambda) \langle a, x_2 \rangle \\ &= \lambda c + (1 - \lambda)c = c. \end{aligned}$$

Portanto, $z \in \mathcal{H}$, ou seja, \mathcal{H} é um conjunto convexo. □

Outra forma de representação de um hiperplano seria:

$$\mathcal{H} = \{x \in \mathbb{R}^n \mid a^T x = c\},$$

onde $a \in \mathbb{R}^n$, $a \neq 0$ é o vetor normal de \mathcal{H} e $c \in \mathbb{R}$.

Observações:

- \mathcal{H} é uma variedade linear: se $x, y \in \mathcal{H}$ então definindo-se $z := \alpha x + (1 - \alpha)y$, $\alpha \in \mathbb{R}$,

$$\begin{aligned} a^T z &= \alpha a^T x + (1 - \alpha) a^T y \\ &= \alpha c + (1 - \alpha)c \\ &= c, \forall \alpha \in \mathbb{R}. \end{aligned}$$

Ou seja $\alpha x + (1 - \alpha)y \in \mathcal{H}$ para todo $\alpha \in \mathbb{R}$.

- \mathcal{H} é a translação do subespaço

$$\mathcal{S} = \{x \in \mathbb{R}^n \mid c^T x = 0\}.$$

\mathcal{S} é o conjunto dos pontos $x \in \mathbb{R}^n$ ortogonais a c .

Portanto, $\dim(\mathcal{S}) = \dim(\mathcal{H}) = n - 1$; \mathcal{H} é a maior variedade própria contida no \mathbb{R}^n .

3.6.1 Hiperplano Suporte e Hiperplano Separador

Definição 3.29. \mathcal{H} é um *hiperplano suporte* de um conjunto convexo Ω , se $\Omega \subset \mathcal{H}_{\leq}$ ou $\Omega \subset \mathcal{H}_{\geq}$ e \mathcal{H} contém um ponto de $\overline{\Omega}$.

Podemos visualizar um hiperplano suporte mediante a Figura 3.8.

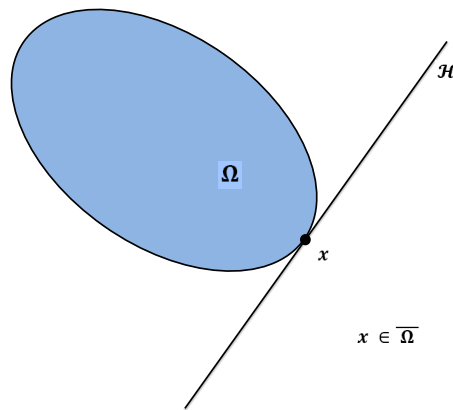


Figura 3.8: hiperplano suporte

Definição 3.30. Sejam $\Omega, \Gamma \subset \mathbb{R}^n$ conjuntos quaisquer. Diz-se que

$$\mathcal{H} = \{x \in \mathbb{R}^n \mid c^T x = k\}$$

é um **hiperplano separador** de Ω e Γ se $\Omega \subset \mathcal{H}_{\geq}$ e $\Gamma \subset \mathcal{H}_{\leq}$.

Observemos exemplos de hiperplanos separadores na Figura 3.9.

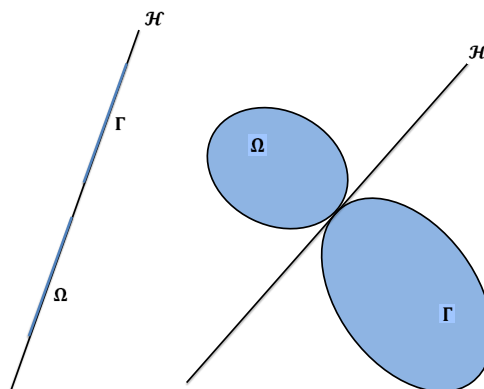


Figura 3.9: hiperplano separador

Dependendo da dimensão do espaço onde estamos trabalhando, um hiperplano pode ter várias representações. O plano Euclidiano, formado pelo conjunto dos pares ordenados de números reais, escritos na forma $\mathbb{R}^2 = \{(x_1, x_2) \mid x_1, x_2 \in \mathbb{R}\}$ é representado geometricamente como no gráfico apresentado na Figura 3.10.

No \mathbb{R}^2 a equação $a_1x_1 + a_2x_2 = c$, onde $a_1, a_2, c \in \mathbb{R}$, representa uma reta. Por exemplo, a equação $2x_1 + 5x_2 = 10$ é representada graficamente pela reta r que aparece na Figura 3.11.

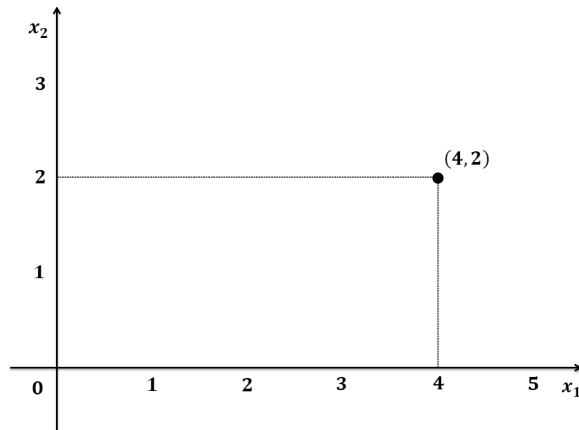


Figura 3.10: Representação Gráfica do \mathbb{R}^2 .

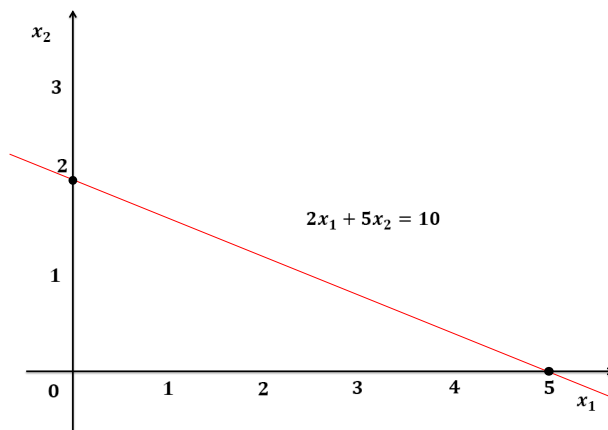


Figura 3.11: Hiperplano $r: 2x_1 + 5x_2 = 10$ no \mathbb{R}^2 .

Uma inequação do tipo $a_1x_1 + a_2x_2 \leq c$ é o conjunto dos pontos da reta $a_1x_1 + a_2x_2 = c$, juntamente com os pontos que estão em um dos lados da reta. Por exemplo, $2x_1 + 5x_2 \leq 10$ é o conjunto dos pontos que aparecem sombreados no gráfico da Figura 3.12.

Portanto, no espaço de dimensão 2, o hiperplano é representado por uma reta. E o conjunto dos pontos que satisfazem a desigualdade da forma $a_1x_1 + a_2x_2 \leq c$, ou da forma $a_1x_1 + a_2x_2 \geq c$, onde pelo menos uma das constantes a_1 ou a_2 é diferente de zero, é o que definimos anteriormente como semi-espaço fechado.

Analogamente, o espaço Euclidiano de dimensão 3, pode ser representado por triplas ordenadas e escrito como

$$\mathbb{R}^3 = \{(x_1, x_2, x_3) \mid x_1, x_2, x_3 \in \mathbb{R}\}.$$

No \mathbb{R}^3 a equação $a_1x_1 + a_2x_2 + a_3x_3 = c$, onde a_1, a_2, a_3 e c são constantes,

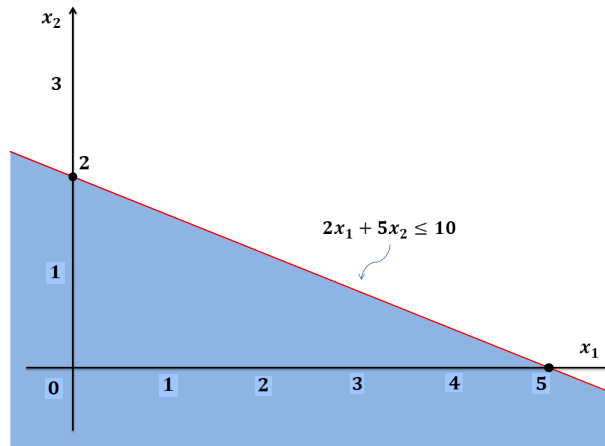


Figura 3.12: Semi-espço $2x_1 + 5x_2 \leq 10$ no \mathbb{R}^2 .

representa um plano. Por exemplo, a equação $3x_1 - x_2 + 4x_3 = 6$ é um plano. O conjunto dos pontos que satisfazem as desigualdades da forma

$$a_1x_1 + a_2x_2 + a_3x_3 \leq c,$$

e da forma

$$a_1x_1 + a_2x_2 + a_3x_3 \geq c.$$

são os semi-espços fechados do \mathbb{R}^3 .

Podemos generalizar tal ideia a um espaço Euclidiano de dimensão n , escrevendo da seguinte forma:

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in \mathbb{R}\}.$$

No \mathbb{R}^n a equação $a_1x_1 + a_2x_2 + \dots + a_nx_n = c$, onde a_1, a_2, \dots, a_n e c são constantes, é a representação de um Hiperplano. E, da mesma maneira, um semi-espço fechado no \mathbb{R}^n é o conjunto dos pontos que satisfazem uma desigualdade da forma

$$a_1x_1 + a_2x_2 + \dots + a_nx_n \leq c,$$

ou da forma

$$a_1x_1 + a_2x_2 + \dots + a_nx_n \geq c.$$

3.7 Semi-espços

Definição 3.31. Sejam $a \in \mathbb{R}^n$ e $c \in \mathbb{R}$. O subconjunto do \mathbb{R}^n definido por $\mathcal{H}_< = \{x \in \mathbb{R}^n \mid \langle a, x \rangle < c\}$ ou $\mathcal{H}_> = \{x \in \mathbb{R}^n \mid \langle a, x \rangle > c\}$ é denominado semi-espço aberto, enquanto que o subconjunto $\mathcal{H}_\leq = \{x \in \mathbb{R}^n \mid \langle a, x \rangle \leq c\}$ ou

$\mathcal{H}_{\geq} = \{x \in \mathbb{R}^n \mid \langle a, x \rangle \geq c\}$ por semi-espço fechado do \mathbb{R}^n .

Note que a união destes semi-espços representa o espço completo \mathbb{R}^n . Assim, um hiperplano separa o espço \mathbb{R}^n em duas regiões ou semi-espços. Também é usual, em determinados momentos, utilizarmos o termo semi-plano para substituir semi-espço.

Lema 3.32. *Qualquer semi-espço em \mathbb{R}^n é um conjunto convexo.*

Demonstração.

Consideremos, nesta demonstração, apenas o caso em que o semi-espço é o conjunto da forma $\mathcal{H}_{\leq} = \{x \in \mathbb{R}^n \mid \langle a, x \rangle \leq c\}$, os demais casos são análogos. Ora, sejam x_1 e $x_2 \in \mathcal{H}_{\leq}$ e $\lambda \in [0, 1]$, isto é, $\langle a, x_1 \rangle \leq c$ e $\langle a, x_2 \rangle \leq c$. Então:

$$\begin{aligned} \langle a, \lambda x_1 + (1 - \lambda)x_2 \rangle &= \langle a, \lambda x_1 \rangle + \langle a, (1 - \lambda)x_2 \rangle \\ &= \lambda \langle a, x_1 \rangle + (1 - \lambda) \langle a, x_2 \rangle \\ &\leq \lambda c + (1 - \lambda)c = c. \end{aligned}$$

Assim, concluímos o que queríamos demonstrar. \square

Definição 3.33. (Conjunto Poliedral) *Um conjunto poliedral, ou simplesmente poliedro, $\Omega \subset \mathbb{R}^n$ é a interseção de um número finito de semi-espços fechados. Em outras palavras, poliedro seria um conjunto solução de um número finito de igualdades e desigualdades lineares.*

Lema 3.34. *Um conjunto poliedral em \mathbb{R}^n é convexo.*

Demonstração.

Um conjunto poliedral é o conjunto solução de um sistema finito de equações e inequações lineares, isto é, uma interseção de semi-espços e hiperplanos que são conjuntos convexos. Como a interseção de conjuntos convexos é um conjunto convexo, temos que um conjunto poliedral é convexo. \square

Os conjuntos afins, como por exemplo os subespços, hiperplanos e retas, assim como os raios, segmentos de reta e semi-espços são todos exemplos de poliedros. A Figura 3.13 apresenta exemplos de conjuntos poliedrais.

Como um exemplo de conjunto poliedral, podemos citar o conjunto

$$\Omega := \{x \in \mathbb{R}^n \mid (c^i)^T x \leq k_i, i = 1, 2, \dots, m\}$$

onde $c^i \neq 0$, $i = 1, 2, \dots, m$.

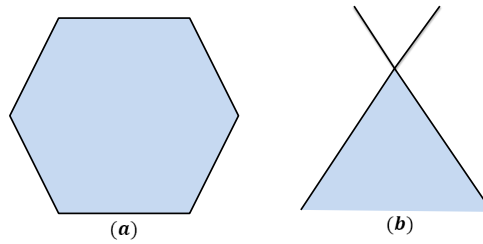


Figura 3.13: (a) Poliedro limitado; (b) Poliedro ilimitado

Conjuntos poliedrais são fechados e convexos, mas podem ser ilimitados, conforme podemos verificar no exemplo a seguir. Para $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$,

$$\Omega := \{x \in \mathbb{R}^n \mid Ax \leq b\}.$$

3.7.1 Politopo

Definição 3.35. *Seja $S \subset \mathbb{R}^n$. Dizemos que um conjunto S é um **politopo**, quando S é um poliedro limitado.*

Alguns autores utilizam a convenção contrária, isto é, politopo de forma geral e poliedro para um politopo limitado. Dessa forma, é comum encontrarmos a definição que politopos são generalizações dos conceitos de polígonos e poliedros, os quais são bem conhecidos e trabalhados na matemática básica. Podemos dizer que politopos são basicamente conhecidos como o termo geral da sequência {ponto, segmento, polígono, poliedro, ...}. Portanto, quando estamos trabalhando no \mathbb{R}^2 , um politopo seria simplesmente um polígono, enquanto que no \mathbb{R}^3 seria a representação de um poliedro. E ainda, podemos facilmente, notar outra propriedade: com dois pontos distintos podemos formar um segmento de reta, com três ou mais segmentos de reta é possível formar um polígono, com quatro ou mais polígonos podemos formar um poliedro, e assim sucessivamente.

3.8 O Operador de Projeção

Definição 3.36. *Uma projeção ortogonal do ponto $x \in \mathbb{R}^n$ sobre um conjunto $D \subset \mathbb{R}^n$ é um ponto de D que está mais próximo de x , onde a distância é medida pela norma euclidiana. Em outras palavras, uma projeção de x sobre D é uma solução global do*

problema

$$\min \|y - x\|, \quad (3-2)$$

sujeito a $y \in D$.

Conforme veremos no próximo corolário, quando D é fechado uma projeção existe para qualquer $x \in \mathbb{R}^n$.

Corolário 3.37. *Seja $D \subset \mathbb{R}^n$ um conjunto fechado não-vazio. Então a projeção de x sobre D existe para todo ponto $x \in \mathbb{R}^n$.*

Mostraremos agora que se D é convexo e fechado, então a projeção é única.

A Figura 3.14 ilustra o teorema da projeção a seguir.

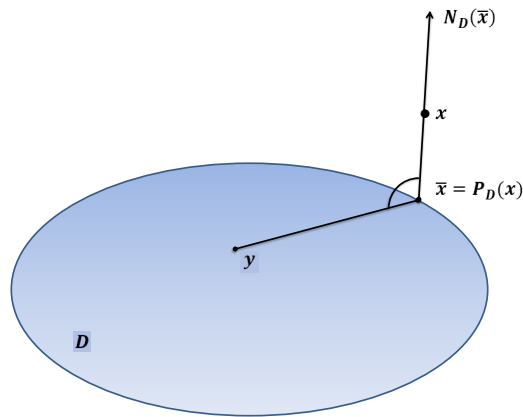


Figura 3.14: \bar{x} é a projeção de x sobre D ; tem-se que para todo $y \in D$, $\langle x - \bar{x}, y - \bar{x} \rangle \leq 0$ ou, equivalentemente, $x - \bar{x} \in N_D(\bar{x})$.

Teorema 3.38. *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,

$$\bar{x} \in D, \quad \langle x - \bar{x}, y - \bar{x} \rangle \leq 0 \quad \forall y \in D, \quad (3-3)$$

ou, equivalentemente,

$$\bar{x} \in D, \quad x - \bar{x} \in N_D(\bar{x}). \quad (3-4)$$

Demonstração.

Sejam \bar{x} uma solução do problema (3-2) e $y \in D$ qualquer. Como $\bar{x} \in D$ e D é convexo, $(1 - \alpha)\bar{x} + \alpha y = x(\alpha) \in D$ para todo $\alpha \in (0, 1]$. Temos então que

$\|x - \bar{x}\| \leq \|x - x(\alpha)\|$ e portanto,

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

Dividindo os dois lados da desigualdade acima por $2\alpha > 0$ e passando o limite quando $\alpha \rightarrow 0_+$, obtemos

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

sendo que $y \in D$ era arbitrário.

Suponhamos agora que um certo \bar{x} satisfaça (3-3). Então, para todo $y \in D$,

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

Temos então que $\|x - \bar{x}\| \leq \|x - y\|$ para todo $y \in D$, isto é, \bar{x} é uma projeção de x sobre D .

Finalmente, mostramos que a projeção é única. Seja \hat{x} alguma outra solução de (3-2).

Usando (3-3) para \bar{x} com $y = \hat{x} \in D$ e para \hat{x} com $y = \bar{x} \in D$, temos

$$\begin{aligned} \langle x - \bar{x}, \hat{x} - \bar{x} \rangle &\leq 0, \\ \langle x - \hat{x}, \bar{x} - \hat{x} \rangle &\leq 0. \end{aligned}$$

Somando, obtemos

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

□

3.9 Teoremas da Separação

Sejam D_1 e D_2 conjuntos não-vazios em \mathbb{R}^n e

$$H(a, c) = \{x \in \mathbb{R}^n \mid \langle a, x \rangle = c\}$$

um hiperplano, onde $a \in \mathbb{R}^n \setminus \{0\}$, $c \in \mathbb{R}$.

Definição 3.39. Dizemos que o hiperplano $H(a, c)$ separa os conjuntos D_1 e D_2 se

$$\langle a, x^1 \rangle \leq c \leq \langle a, x^2 \rangle \quad \forall x^1 \in D_1 \quad \forall x^2 \in D_2.$$

Dizemos que $H(a, c)$ separa estritamente D_1 e D_2 quando as desigualdades acima são estritas.

A noção de separação de conjuntos é muito importante. No sentido geométrico, separação significa que um dos conjuntos fica de um lado do hiperplano $H(a, c)$, enquanto o outro conjunto fica do outro lado, veja a Figura 3.15.

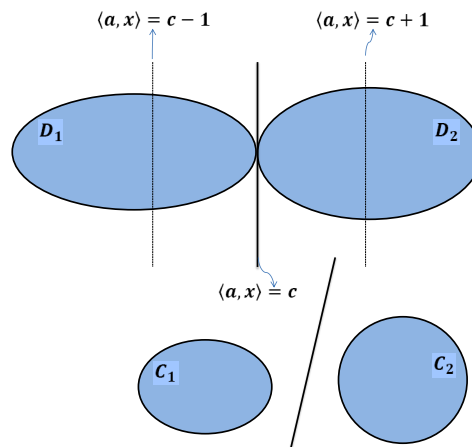


Figura 3.15: Os conjuntos D_1 e D_2 são separáveis; os conjuntos C_1 e C_2 são estritamente separáveis.

Como pode ser visto com facilidade fazendo-se desenhos geométricos, a possibilidade de separar dois conjuntos dados está ligada ao fato da interseção deles ser vazia ou não, e à convexidade deles, veja as Figuras 3.15 e 3.16.

Mostramos primeiro que um ponto que não pertence ao fecho de um conjunto convexo pode ser separado dele estritamente, veja a Figura 3.17.

Lema 3.40. (Lema de Minkowski) Seja $D \subset \mathbb{R}^n$ um conjunto convexo não-vazio. Se $x \notin cl(D)$, então existem $a \in \mathbb{R}^n \setminus \{0\}$ e $c \in \mathbb{R}$ tais que

$$\langle a, x \rangle = c, \quad \langle a, y \rangle > c \quad \forall y \in D.$$

Demonstração.

O conjunto $cl(D)$ é fechado e convexo. Seja \bar{x} a projeção de x sobre $cl(D)$, cuja existência e unicidade são asseguradas pelo Teorema 3.38.

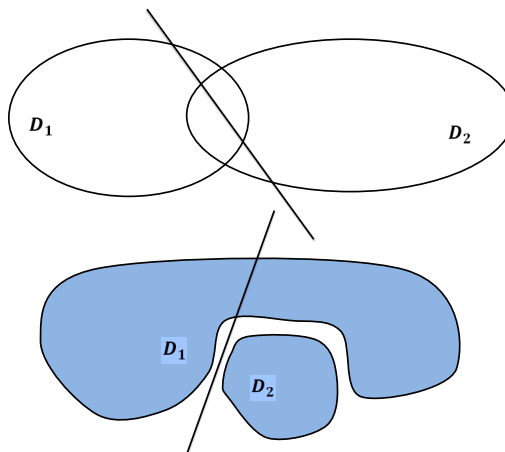


Figura 3.16: Os conjuntos D_1 e D_2 não são separáveis.

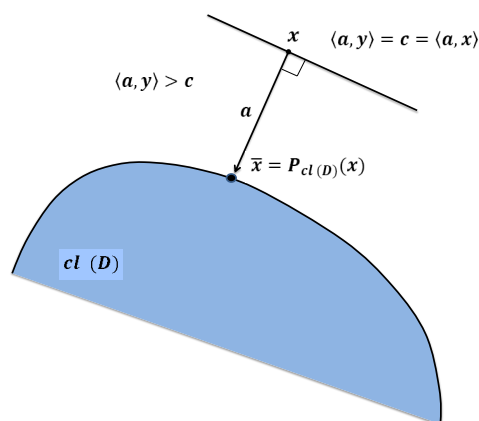


Figura 3.17: Ilustração do Lema de Minkowski.

Quando $x \notin cl(D)$, o ponto x pode ser separado de D estritamente.

Definimos $a = \bar{x} - x$ e $c = \langle \bar{x} - x, x \rangle$. Pela definição, $\langle a, x \rangle = c$. Seja $y \in D$. Temos que

$$\langle a, y \rangle = \langle \bar{x} - x, y \rangle \geq \langle \bar{x} - x, \bar{x} \rangle, \quad (3-5)$$

onde a desigualdade vale para todo $y \in cl(D)$ (e portanto para todo $y \in D$), pelo Teorema 3.38. Obtemos que

$$\begin{aligned} \langle \bar{x} - x, x \rangle &= \|x - \bar{x}\|^2 + \langle \bar{x} - x, x \rangle \\ &= \|x - \bar{x}\|^2 + c > c, \end{aligned}$$

onde usamos que $\|x - \bar{x}\| > 0$ (pela hipótese de que $x \notin cl(D)$). Combinando a última relação com (3-5), obtemos o resultado desejado. \square

Quando um ponto pertence ao fecho de um conjunto, a separação estrita não pode ser

garantida, como se vê no Exemplo 3.41 e na Figura 3.18.

Exemplo 3.41. *Consideremos o conjunto*

$$D = \{x \in \mathbb{R}^2 \mid 0 \neq x \geq 0\} \cup \{x \in \mathbb{R}^2 \mid x_1 < 0 < x_2\}. \quad (3-6)$$

Como pode ser visto fazendo um desenho geométrico, D é convexo. Veja a Figura 3.18.

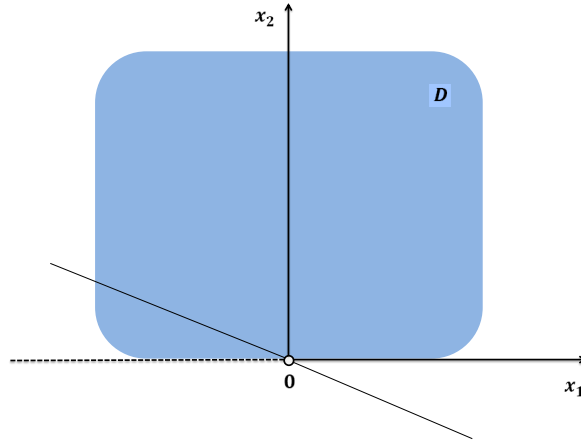


Figura 3.18: *Tem-se que $0 \in cl(D)$, onde D é dado por (3-6); mas o zero não é estritamente separável de D .*

O ponto $x = 0 \notin D$, mas $x \in cl(D)$. Temos $\langle a, x \rangle = 0$ para qualquer $a \in \mathbb{R}^2$, mas não existe nenhum $a \in \mathbb{R}^2$ tal que $\langle a, x \rangle = a_1x_1 + a_2x_2 > 0$ para todo $x \in D$ (isto fica claro na Figura 3.18, ou pode ser verificado fazendo-se o cálculo).

Portanto, 0 não pode ser separado de D estritamente. No entanto, observamos que a separação (não estrita) é possível: escolhendo $a = (0, 1) \in \mathbb{R}^2$ garantimos que $\langle a, x \rangle = x_2 \geq 0$ para todo $x \in D$.

Agora mostramos que um ponto na fronteira de um conjunto convexo pode ser separado dele. Lembramos que $x \in fr(D)$ quando toda vizinhança de x contém tanto pontos que pertencem a D quanto pontos que não pertencem a este conjunto.

Lema 3.42. *Seja $D \subset \mathbb{R}^n$ um conjunto convexo não-vazio. Se $x \in fr(D)$, então existem $a \in \mathbb{R}^n \setminus \{0\}$ e $c \in \mathbb{R}$ tais que*

$$\langle a, x \rangle = c, \quad \langle a, y \rangle \geq c \quad \forall y \in D.$$

Demonstração.

Como $x \in fr(D)$, podemos escolher uma sequência $\{x^k\} \rightarrow x$ quando $k \rightarrow \infty$ tal que $x^k \notin cl(D) \forall k$.

Pelo Lema 3.40, todo ponto x^k pode ser separado de D estritamente, isto é, para todo k existem $a^k \in \mathbb{R}^n \setminus \{0\}$ e $c_k \in \mathbb{R}$ tais que

$$\langle a^k, x^k \rangle = c_k, \quad \langle a^k, y \rangle > c_k \quad \forall y \in D. \quad (3-7)$$

Podemos admitir que $\{a^k/\|a^k\|\} \rightarrow a \neq 0$ quando k tende a zero (tomando uma subsequência, se for necessário). Obtemos que

$$\lim_{k \rightarrow \infty} c_k/\|a^k\| = \lim_{k \rightarrow \infty} \langle a^k, x^k \rangle/\|a^k\| = \langle a, x \rangle.$$

Definindo $c = \langle a, x \rangle$, dividindo a segunda relação em (3-7) por $\|a^k\| > 0$, e passando ao limite $k \rightarrow \infty$, obtemos que para todo $y \in D$ tem-se

$$\langle a, y \rangle \geq c,$$

o que conclui a prova. □

O hiperplano $H(a, c) = \{x \in \mathbb{R}^n \mid \langle a, x \rangle = c\}$, que separa $x \in fr(D)$ de D , se denomina o hiperplano de apoio do conjunto D no ponto $x \in fr(D)$. A Figura 3.19 ilustra o Lema 3.42. Observamos que o conjunto pode admitir mais de um hiperplano de apoio no mesmo ponto.

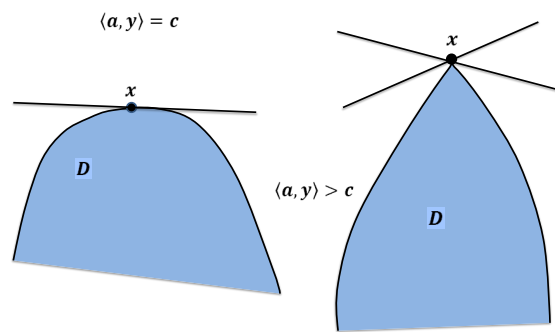


Figura 3.19: Separação de ponto x do conjunto D quando x pertence a fronteira de D ; hiperplanos de apoio de D em $x \in fr(D)$.

Mostramos a seguir que dois conjuntos convexos sem pontos em comum são separáveis. Veja a Figura 3.20.

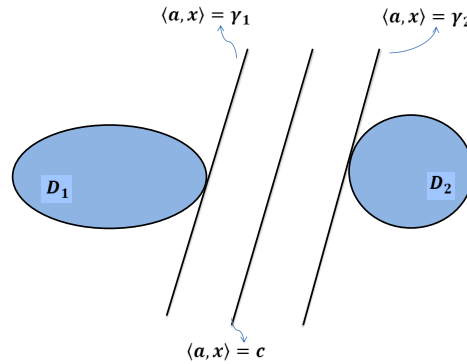


Figura 3.20: Separação de D_1 e D_2 quando $D_1 \cap D_2 = \emptyset$.

Teorema 3.43. Sejam $D_1 \subset \mathbb{R}^n$ e $D_2 \subset \mathbb{R}^n$ conjuntos convexos não-vazios tais que $D_1 \cap D_2 = \emptyset$. Então existem $a \in \mathbb{R}^n \setminus \{0\}$ e $c \in \mathbb{R}$ tais que

$$\langle a, x^1 \rangle \leq c \leq \langle a, x^2 \rangle \quad \forall x^1 \in D_1, x^2 \in D_2.$$

Demonstração.

O conjunto $D = D_2 - D_1$ é convexo, veja o Lema 3.11, e $0 \notin D$ porque $D_1 \cap D_2 = \emptyset$. Nesta situação há duas possibilidades: ou $0 \notin cl(D)$, ou $0 \in fr(D)$. Para separar 0 do conjunto D , no primeiro caso usamos o Lema 3.40 e no segundo caso usamos o Lema 3.42. Concluímos que existe $a \in \mathbb{R}^n \setminus \{0\}$ tal que

$$\langle a, x \rangle \geq 0 \quad \forall x \in D,$$

isto é, $\langle a, x^2 - x^1 \rangle \geq 0$ para todos $x^2 \in D_2, x^1 \in D_1$. Ou ainda,

$$\langle a, x^2 \rangle \geq \langle a, x^1 \rangle \quad \forall x^2 \in D_2, \forall x^1 \in D_1.$$

Em particular, a função $\langle a, \cdot \rangle$ é limitada inferiormente em D_2 e limitada superiormente em D_1 . Ainda, da relação acima obtemos que

$$\gamma_2 = \inf_{x^2 \in D_2} \langle a, x^2 \rangle \geq \sup_{x^1 \in D_1} \langle a, x^1 \rangle = \gamma_1.$$

Definindo $c = (\gamma_1 + \gamma_2)/2$, temos que $\gamma_2 \geq c \geq \gamma_1$ e

$$\forall x^1 \in D_1, \quad \langle a, x^1 \rangle \leq \sup_{x^1 \in D_1} \langle a, x^1 \rangle = \gamma_1 \leq c,$$

$$\forall x^2 \in D_2, \quad \langle a, x^2 \rangle \geq \inf_{x^2 \in D_2} \langle a, x^2 \rangle = \gamma_2 \geq c,$$

o que completa a prova do resultado anunciado. □

Funções Convexas

4.1 Caracterização da Função Convexa

As Funções convexas são de grande importância em diversas áreas como a Termodinâmica, a Mecânica Estatística, a Teoria das Probabilidades, Equações Diferenciais, no Cálculo Variacional e em diversos outros contextos. Este tipo de função é também utilizado para se demonstrar algumas desigualdades, como por exemplo a desigualdade de Young. Neste capítulo apresentaremos suas definições no espaço \mathbb{R}^n e enunciaremos algumas de suas propriedades básicas. A revisão apresentada neste capítulo baseia-se principalmente nas referências [4, 6]

Definição 4.1. Uma função $f : D \rightarrow \mathbb{R}$ definida sobre um conjunto convexo $D \subset \mathbb{R}^n$ se diz convexa se

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2),$$

$$\forall x_1, x_2 \in D, \forall \lambda \in [0, 1].$$

Quando a desigualdade se verifica no sentido estrito, então dizemos que a função é estritamente convexa, isto é, f é estritamente convexa em D , se

$$f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2),$$

$$\forall x_1, x_2 \in D, \forall \lambda \in (0, 1).$$

Definição 4.2. Se $D \subset \mathbb{R}^n$ é um conjunto convexo, dizemos que $f : D \rightarrow \mathbb{R}$ é uma função côncava em D quando a função $(-f)$ é convexa em D .

Lema 4.3. A função quadrática $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = ax^2 + bx + c$, com $b, c \in \mathbb{R}$ e $a \in \mathbb{R}_+^*$ é uma função estritamente convexa.

Demonstração.

Provemos inicialmente que, se $x_1 \neq x_2$ e $0 < \lambda < 1$, então

$$[\lambda x_1 + (1 - \lambda)x_2]^2 < \lambda x_1^2 + (1 - \lambda)x_2^2. \quad (4-1)$$

Neste caso, temos que

$$\begin{aligned} 0 &< \lambda x_1^2 + (1-\lambda)x_2^2 - [\lambda x_1 + (1-\lambda)x_2]^2 \\ 0 &< \lambda x_1^2 + (1-\lambda)x_2^2 - \lambda^2 x_1^2 - 2\lambda(1-\lambda)x_1x_2 - (1-\lambda)^2 x_2^2 \\ 0 &< (\lambda - \lambda^2)x_1^2 - 2\lambda(1-\lambda)x_1x_2 + (\lambda - \lambda^2)x_2^2 \\ 0 &< \lambda(1-\lambda)(x_1 - x_2)^2. \end{aligned}$$

E agora, façamos

$$\begin{aligned} f[\lambda x_1 + (1-\lambda)x_2] &= a[\lambda x_1 + (1-\lambda)x_2]^2 + b[\lambda x_1 + (1-\lambda)x_2] + c \\ &< a[\lambda x_1^2 + (1-\lambda)x_2^2] + b[\lambda x_1 + (1-\lambda)x_2] + c \\ &= \lambda a x_1^2 + \lambda b x_1 + \lambda c + (1-\lambda)x_2^2 + (1-\lambda)b x_2 + (1-\lambda)c \\ &= \lambda f(x_1) + (1-\lambda)f(x_2). \end{aligned}$$

Portanto, f é estritamente convexa. □

Além disso, quando f é uma função convexa no conjunto convexo e não vazio D , temos que $(-f)$ é uma função côncava em D .

A função f diz-se fortemente convexa com módulo $\gamma > 0$, quando para quaisquer $x_1 \in A$, $x_2 \in B$ e $\lambda \in [0, 1]$, tem-se

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) - \gamma \lambda(1-\lambda) \|x_1 - x_2\|^2.$$

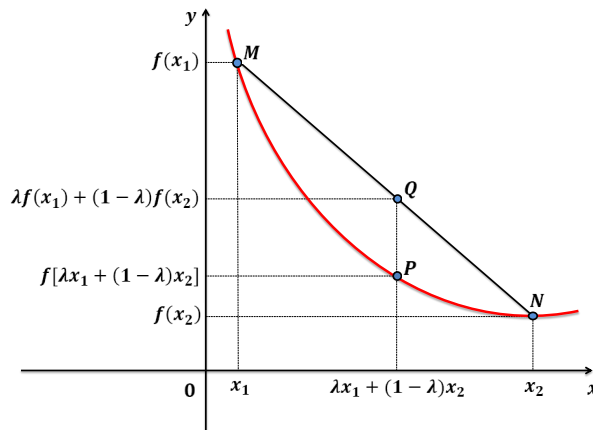


Figura 4.1: Gráfico de uma Função Convexa

Utilizando as informações da Figura 4.1, temos que $x_1 < x < x_2$ e fazendo $\mu = \frac{x-x_1}{x_2-x_1}$, constatamos que $0 \leq \mu \leq 1$, $x = x_1 + (x-x_1)$,

Daí,

$$x_1 + \mu(x_2 - x_1) = (1-\mu)x_1 + \mu x_2,$$

e ainda, chamando $\lambda = 1 - \mu$, temos,

$$x = \lambda x_1 + \mu x_2, \quad \lambda \geq 0, \quad \mu \geq 0 \quad e \quad \lambda + \mu = 1.$$

Repare que os pontos P e Q da Figura 4.1 possuem coordenadas

$$P = (\lambda x_1 + \mu x_2, f(\lambda x_1 + \mu x_2)) \quad e \quad Q = (\lambda x_1 + \mu x_2, \lambda f(x_1) + \mu f(x_2)).$$

Com isso, dizemos que a função f é convexa quando o ponto Q está sempre acima do ponto P . E assim,

$$\forall x_1, x_2 \in A, \forall \lambda, \mu \geq 0 \text{ com } \lambda + \mu = 1,$$

chegamos a desigualdade esperada

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Podemos também, interpretar geometricamente, a definição de função convexa como uma função tal que, se $z_1, z_2 \in \mathbb{R}^n$ são quaisquer dois pontos que estão sobre o gráfico de f , então os pontos do segmento de reta que une z_1 e z_2 estão sobre ou acima do gráfico de f , e ainda, a curva deve sempre estar sobre ou acima de qualquer de suas tangentes. As funções convexas e côncavas são unimodais, isto é, funções com um único ponto de máximo ou mínimo, tais que o extremo local é o extremo global.

A função $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$, é um exemplo de função fortemente convexa. Repare que $f''(x) = 2 > 0$ e toda reta tangente a $f(x)$ é um suporte para a função, pois $f(x)$ fica sempre acima da tangente. Particularmente, podemos verificar isso no ponto $(0, 0)$ e na reta $y = 0$. Este fato ocorre com toda função estritamente convexa. A função $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = e^x$, é estritamente convexa, mas não é fortemente convexa. Já a função $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x$, é convexa, contudo não é estritamente convexa. É de fácil entendimento que uma função fortemente convexa é estritamente convexa, e uma função estritamente convexa é convexa.

Nota Histórica: A noção de função convexa foi introduzida por Johan Ludwig William Valdemar Jensen (1859-1925), utilizando a desigualdade (4-2) conhecida como Desigualdade do Ponto Médio. Vejamos:

Seja $f: D \rightarrow \mathbb{R}$ uma função convexa em um conjunto D , tal que $f(x) = ax^2 + bx + c$

com $a \in \mathbb{R}_+^*$ e $b, c \in \mathbb{R}$, então vale a desigualdade:

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2}. \quad (4-2)$$

Demonstração.

Provemos inicialmente que, se $x_1 \neq x_2$ então $\left(\frac{x_1 + x_2}{2}\right)^2 < \frac{x_1^2 + x_2^2}{2}$. Assim, segue que

$$\begin{aligned} \frac{x_1^2 + x_2^2}{2} - \left(\frac{x_1 + x_2}{2}\right)^2 &= \frac{x_1^2 - 2x_1x_2 + x_2^2}{4} \\ &= \left(\frac{x_1 - x_2}{2}\right)^2 > 0, \end{aligned}$$

onde $x_1 \neq x_2$.

Agora, se $x_1 \neq x_2$ e $a > 0$, temos:

$$\begin{aligned} f\left(\frac{x_1 + x_2}{2}\right) &= a\left(\frac{x_1 + x_2}{2}\right)^2 + b\left(\frac{x_1 + x_2}{2}\right) + c \\ &< a\left(\frac{x_1^2 + x_2^2}{2}\right) + b\left(\frac{x_1 + x_2}{2}\right) + c \\ &= \frac{ax_1^2 + bx_1 + c + ax_2^2 + bx_2 + c}{2} \\ &= \frac{f(x_1) + f(x_2)}{2}. \end{aligned}$$

□

O Lema 4.4 seguinte é uma caracterização de funções convexas que deu origem à esta teoria e utiliza como hipótese a desigualdade (4-2).

Lema 4.4. *Seja $f : D \rightarrow \mathbb{R}$ tal que*

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2},$$

para todo $x_1, x_2 \in D$. Então, vale

$$f\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) \leq \frac{f(x_1) + f(x_2) + \dots + f(x_n)}{n},$$

para todo $n \in \mathbb{N}$ e para todo $x_1, x_2, \dots, x_n \in D$.

A prova é feita por indução sobre n , ver Definição 2.42, e fica a cargo do leitor.

Agora apresentaremos alguns critérios que podem ser usados para reconhecermos a que classe uma função dada pertence.

Definição 4.5. Define-se o epígrafo de uma função $f : D \rightarrow \mathbb{R}$ como

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

A relação entre convexidade de conjuntos e de funções é dada mais adiante pelo Teorema 4.7. E pode ser melhor compreendida na Figura 4.2.

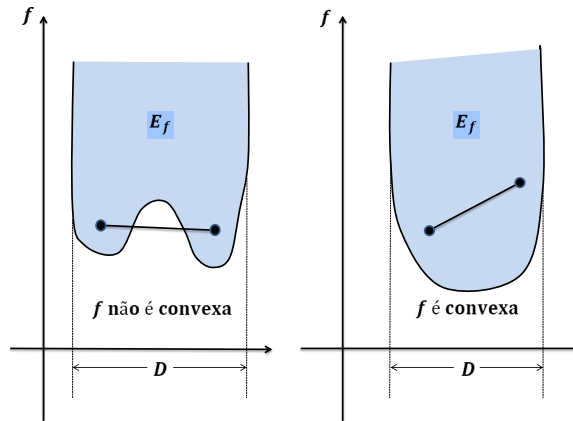


Figura 4.2: Convexidade da função $f \iff$ convexidade do epígrafo de f .

O seguinte teorema fornece uma primeira condição necessária e suficiente de convexidade:

Teorema 4.6. A condição necessária e suficiente para que a função $f(x)$ seja convexa no conjunto $D \subseteq \mathbb{R}^n$ é que, quaisquer que sejam $a, b \in D$, a função real de variável real $g(\lambda) = f[\lambda a + (1 - \lambda)b]$ seja convexa no intervalo $[0, 1]$.

Demonstração.

A condição é necessária. Admitindo $f(x)$ convexa no conjunto convexo D , considere-se a função $g(\lambda) = f[\lambda a + (1 - \lambda)b]$. Dados os números reais $\lambda, \mu \in [0, 1]$, e sejam $\alpha \geq 0$ e $\beta \geq 0$ tais que $\alpha + \beta = 1$. Tem-se então,

$$\begin{aligned} g(\alpha\lambda + \beta\mu) &= f[(\alpha\lambda + \beta\mu) \cdot a + (1 - \alpha\lambda - \beta\mu) \cdot b] = \\ &= f[(\alpha\lambda + \beta\mu) \cdot a + (\alpha + \beta - \alpha\lambda - \beta\mu) \cdot b] = \\ &= f[\alpha \cdot (\lambda a + b - \lambda b) + \beta \cdot (\mu a + b - \mu b)] = \\ &= f\{\alpha \cdot [\lambda a + (1 - \lambda)b] + \beta \cdot [\mu a + (1 - \mu)b]\}, \end{aligned}$$

e como o conjunto D é convexo temos,

$a, b \in D$, então $\lambda a + (1 - \lambda)b \in D$ e $\mu a + (1 - \mu)b \in D$,

Conclui-se, pela convexidade de $f(x)$ que:

$$\begin{aligned} g(\alpha\lambda + \beta\mu) &\leq \alpha \cdot f[\lambda a + (1 - \lambda)b] + \beta \cdot f[\mu a + (1 - \mu)b] \\ &= \alpha \cdot g(\lambda) + \beta \cdot g(\mu), \end{aligned}$$

provando assim, a convexidade da função $g(\lambda)$ no intervalo $[0, 1]$. Agora, verifiquemos que a condição é suficiente. Para todo $a, b \in D$, a função $g(\lambda) = f[\lambda a + (1 - \lambda)b]$ é convexa no intervalo $[0, 1]$, tem-se, com $\lambda \geq 0$, $\mu \geq 0$ e $\lambda + \mu = 1$,

$$\begin{aligned} f(\lambda a + \mu b) &= f[\lambda a + (1 - \lambda)b] = g(\lambda) = g(\lambda \cdot 1 + \mu \cdot 0) \leq \\ &\leq \lambda \cdot g(1) + \mu \cdot g(0) = \lambda f(a) + \mu f(b). \end{aligned}$$

Ficando, com isso, provado que $f(x)$ é convexa no conjunto D .

Teorema 4.7. *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.

Primeiramente vamos supor que E_f seja um conjunto convexo. Sejam $x \in D$ e $y \in D$ quaisquer. Obviamente, $(x, f(x)) \in E_f$ e $(y, f(y)) \in E_f$ pela própria definição. Como E_f é convexo $\forall \alpha \in [0, 1]$, 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, utilizando a definição de epígrafo, podemos dizer que

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

ou seja, f é uma função convexa.

Agora, vamos supor que f seja convexa. Sejam $(x, c_1) \in E_f$ e $(y, c_2) \in E_f$. Sabemos que $f(x) \leq c_1$ e $f(y) \leq c_2$, logo pela convexidade de f e para todo $\alpha \in [0, 1]$, temos que

$$\begin{aligned} f(\alpha x + (1 - \alpha)y) &\leq \alpha f(x) + (1 - \alpha)f(y) \\ &\leq \alpha c_1 + (1 - \alpha)c_2, \end{aligned}$$

o que é equivalente a

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

De acordo com o Teorema 4.7, concluímos que funções convexas são funções cujos epígrafos são conjuntos convexos. \square

Se invertermos o sentido das desigualdades na argumentação precedente, o teorema fica provado para o caso da concavidade.

4.2 Propriedades das Funções Convexas

Pelo Teorema que se segue, mostraremos que a soma de múltiplos não-negativos de um número finito de funções convexas é uma função convexa.

Teorema 4.8. (*Convexidade da soma de funções convexas*) *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f_i : D \rightarrow \mathbb{R}$, $i = 1, \dots, p$, funções convexas em D . Então para quaisquer $\mu_i \in \mathbb{R}_+$, $i = 1, \dots, p$, a função*

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

é convexa em D .

Demonstração.

Para $x \in D$, $y \in D$ e $\lambda \in [0, 1]$ quaisquer, temos que:

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

onde a desigualdade segue da convexidade de f_i e do fato de que $\mu_i \geq 0$, $i = 1, \dots, p$. \square

A seguir mostramos que o supremo de funções convexas também é uma função convexa. A demonstração é baseada no fato de que o epígrafo do supremo é a intersecção dos epígrafos das funções que definem o supremo. Veja a Figura 4.3.

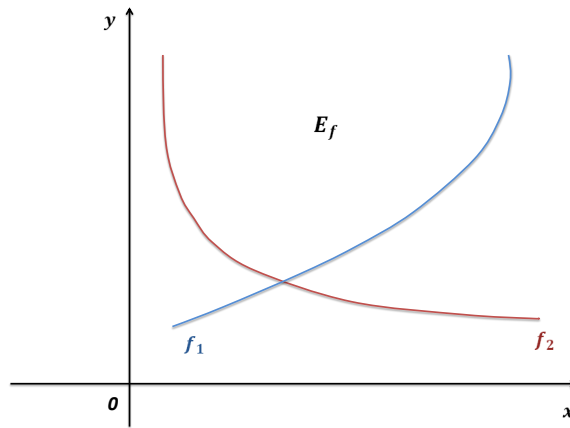


Figura 4.3: O epígrafo de supremo de funções convexas é um conjunto convexo (no desenho, $f(x) = \max\{f_1(x), f_2(x)\}$). Portanto, o supremo de funções convexas é uma função convexa.

Proposição 4.9. (Convexidade do Supremo de Funções Convexas) Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f_i : D \rightarrow \mathbb{R}$, $1 \leq i \leq p$, funções convexas em D . Suponhamos que exista $\beta \in \mathbb{R}$ tal que $f_i(x) \leq \beta$ para todo $x \in D$ e $1 \leq i \leq p$. Então a função

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

é convexa em D .

Demonstração.

Seja $c \in \mathbb{R}$ arbitrário. Tem-se que

$$\begin{aligned} E_f &= \{(x, c) \in D \times \mathbb{R} \mid f(x) \leq c\} \\ &= \{(x, c) \in D \times \mathbb{R} \mid f_i(x) \leq c, 1 \leq i \leq p\} \\ &= \bigcap_{1 \leq i \leq p} \{(x, c) \in D \times \mathbb{R} \mid f_i(x) \leq c\} \\ &= \bigcap_{1 \leq i \leq p} E_{f_i}. \end{aligned}$$

Pela convexidade de f_i , os epígrafos de E_{f_i} , $1 \leq i \leq p$, são convexas, ver o Teorema 4.7. Logo, a interseção deles é um conjunto convexo, ver o Teorema 3.8. Usando de novo o Teorema 4.7, a convexidade de E_f implica a convexidade de f . \square

Observemos que no resultado acima a condição de que as funções que definem o supremo sejam uniformemente limitadas (superiormente) é necessária somente para

garantir que f tenha valores finitos no conjunto D .

Outra forma de obter uma função convexa é a seguinte.

Proposição 4.10. *Sejam $g : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa, $h : \mathbb{R} \rightarrow \mathbb{R}$ uma função convexa e não-decrescente. Então, a função*

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

é convexa.

Demonstração.

Considerando a convexidade de g e $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$, $\alpha \in [0, 1]$, temos que

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

Considerando também a definição de f , podemos escrever

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

A primeira desigualdade se justifica pelo fato de h ser não-decrescente, e a segunda por h ser uma função convexa.

Portanto, a função $f = h \circ g$ é convexa. \square

Para esclarecer o Teorema 4.10, podemos considerar a função $g : \mathbb{R} \rightarrow \mathbb{R}$, $g(x) = e^x$ convexa e a função $h : \mathbb{R} \rightarrow \mathbb{R}$, $h(x) = -x$ também convexa, contudo decrescente. Repare que a função composta $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = h(g(x))$, definida por $f(x) = -e^x$, é côncava.

A seguir daremos uma definição importante que nos permite eliminar a hipótese de que o conjunto viável seja compacto, pois contribui para resultados de existência de solução ao custo de fortalecimento das hipóteses sobre função objetivo.

Definição 4.11. (Conjunto de Nível de Funções Convexas) *Seja $D \subset \mathbb{R}^n$ um conjunto convexo e seja $f : D \rightarrow \mathbb{R}$ uma função convexa em D . O conjunto de pontos da forma*

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

é denominado conjunto de nível da função convexa f . Veja a Figura 4.4.

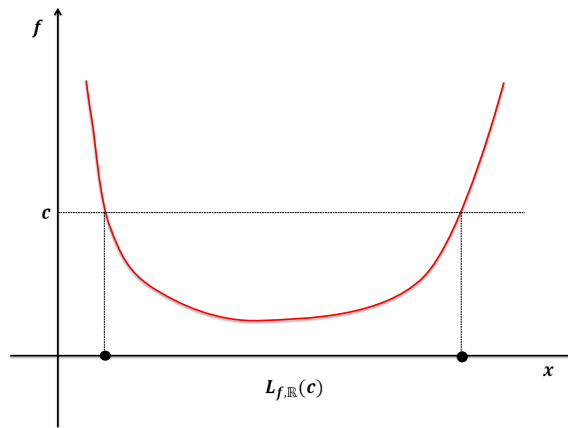


Figura 4.4: Conjuntos de nível de uma função convexa são convexas

Teorema 4.12. *O conjunto de nível de uma função convexa é convexo.*

Demonstração.

Considere uma função f convexa no conjunto D também convexo. Seja $c \in \mathbb{R}$.

Se $L_{f, D} = \emptyset$, não há o que provar, pois \emptyset é um conjunto convexo trivial.

Caso contrário, considere $x \in L_{f, D}(c)$ e $y \in L_{f, D}(c)$, e daí, $x \in D$, $f(x) \leq c$ e $y \in D$, $f(y) \leq c$. Como D é convexo, temos que $\alpha x + (1 - \alpha)y \in D$.

Pela convexidade de f em D , temos

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

Portanto, $(\alpha x + (1 - \alpha)y) \in L_{f, D}(c)$ e assim fica demonstrado que $L_{f, D}(c)$ é um conjunto convexo. \square

Dada a definição de conjunto de nível de uma função, podemos concluir que a convexidade de todos esses conjuntos não garante a convexidade da função. Um bom exemplo seria a função $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^3$, que obviamente não é convexa, mas possui o conjunto de nível convexo.

A constatação acima abre espaço para uma outra definição. Um tipo de função que aparece muito em problemas de otimização. Poderíamos dizer que são mais comuns que as funções convexas, são as quase-convexas.

Definição 4.13. *Seja $D \subset \mathbb{R}$ um conjunto convexo. Dizemos que $f : D \rightarrow \mathbb{R}$ é quase-convexa em D quando os conjuntos de nível $L_{f, D}(c)$ são convexas para todo $c \in \mathbb{R}$.*

A função $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^3$ e também a função cujo gráfico aparece representado na Figura 4.5 são bons exemplos de funções quase-convexas. Observe que nenhuma destas duas funções é convexa.

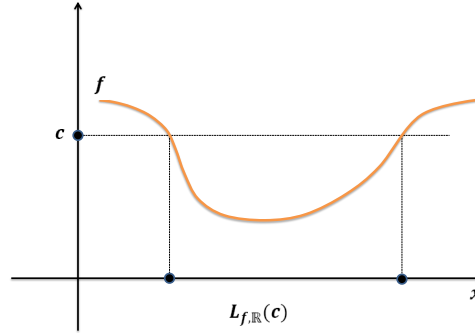


Figura 4.5: Um exemplo de função quase-convexa.

Corolário 4.14. Seja $\Omega \subset \mathbb{R}^n$ um conjunto convexo. Sejam $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma função convexa e $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ uma função afim. Então, o conjunto

$$D = \{x \in \Omega \mid h(x) = 0, g(x) \leq 0\}$$

é convexo.

Demonstração.

Como é fácil de observar, tem-se que

$$D = D_1 \cap D_2 \cap D_3,$$

onde

$$D_1 = \bigcap_{i=1}^l \{x \in \Omega \mid h_i(x) \leq 0\} = \bigcap_{i=1}^l L_{h_i, \Omega}(0),$$

$$D_2 = \bigcap_{i=1}^l \{x \in \Omega \mid -h_i(x) \leq 0\} = \bigcap_{i=1}^l L_{-h_i, \Omega}(0),$$

$$D_3 = \bigcap_{j=1}^m \{x \in \Omega \mid g_j(x) \leq 0\} = \bigcap_{j=1}^m L_{g_j, \Omega}(0).$$

Lembremos que quando h é afim, h e $-h$ são convexas. Portanto, pelo Teorema 4.12, todos os conjuntos de nível acima são convexas e, pelo Lema 3.8, a interseção deles também é um conjunto convexo. \square

Na sequência serão enunciados alguns outros teoremas importantes, sem a preocupação de demonstrarmos. O leitor interessado pode recorrer às referências [1, 4, 5, 10, 13, 14].

Definição 4.15. Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é dita *localmente Lipschitz-contínua* com constante $L > 0$ em $x \in \mathbb{R}^n$ se existe um número positivo r tal que

$$|f(y) - f(z)| \leq L \|y - z\|,$$

para todo $y, z \in B(x, r)$.

Comumente utiliza-se a denominação de localmente Lipschitz.

Se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função convexa em $x \in \mathbb{R}^n$, então f é localmente Lipschitz em x .

Teorema 4.16. (Continuidade de funções convexas) Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e aberto e $f : D \rightarrow \mathbb{R}$ uma função convexa em D .

Então f é localmente Lipschitz-contínua em D . Em particular, f é contínua em D .

Demonstração.

Seja $\bar{x} \in D$ qualquer. Como D é aberto, existe $\delta > 0$ (que depende de \bar{x}) tal que $U \subset D$, onde

$$U = \{x \in \mathbb{R}^n \mid -\delta \leq x_i - \bar{x}_i \leq \delta, i = 1, \dots, n\}.$$

Seja V o conjunto de vértices de U : V é composto por 2^n pontos e $U = \text{conv}(V)$.

Definamos $\beta = \max_{x \in V} f(x)$, que existe porque V tem um número finito de pontos.

Pela convexidade de f , o conjunto de nível $L_{f,D}(\beta)$ é convexo (Teorema 4.8). Como $V \subset L_{f,D}(\beta)$, segue-se que

$$U = \text{conv}(V) \subset \text{conv}(L_{f,D}(\beta)) = L_{f,D}(\beta).$$

Seja $x \in U$ tal que $0 < \|x - \bar{x}\| < \delta$. Definindo

$$d = \frac{\delta(x - \bar{x})}{\|x - \bar{x}\|} = \frac{(x - \bar{x})}{\alpha},$$

$$\alpha = \frac{\|x - \bar{x}\|}{\delta} \in (0, 1),$$

é fácil ver que $\bar{x} \pm d \in U$ (porque $\|d\| = \delta$). E ainda, podemos visualizar melhor com o auxílio da Figura 4.6,

$$x = \bar{x} + \alpha d = \alpha(\bar{x} + d) + (1 - \alpha)\bar{x},$$

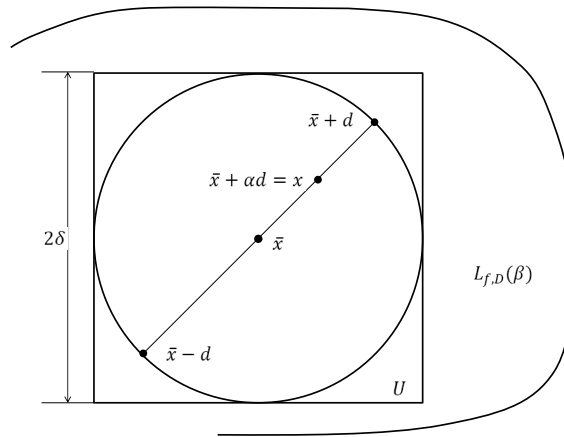


Figura 4.6: Ilustração de prova do Teorema 4.16.

e pela convexidade de f ,

$$\begin{aligned} f(x) &\leq \alpha f(\bar{x} + d) + (1 - \alpha)f(\bar{x}) \\ &\leq \alpha\beta + (1 - \alpha)f(\bar{x}), \end{aligned}$$

porque $\bar{x} + d \in U \subset L_{f,D}(\beta)$. De modo similar,

$$\bar{x} = x - \alpha d = x + \alpha(\bar{x} - d) - \alpha\bar{x},$$

o que implica em

$$\bar{x} = \frac{1}{1 + \alpha}x + \frac{\alpha}{1 + \alpha}(\bar{x} - d),$$

e pela convexidade de f ,

$$\begin{aligned} f(\bar{x}) &\leq \frac{1}{1 + \alpha}f(x) + \frac{\alpha}{1 + \alpha}f(\bar{x} - d) \\ &\leq \frac{f(x) + \alpha\beta}{1 + \alpha}, \end{aligned}$$

porque $\bar{x} - d \in U \subset L_{f,D}(\beta)$. Usando a última desigualdade, obtemos que

$$\begin{aligned} -\alpha(\beta - f(\bar{x})) &\leq f(x) - f(\bar{x}) \\ &\leq \alpha(\beta - f(\bar{x})), \end{aligned}$$

onde a segunda desigualdade se justifica por $f(x) \leq \alpha\beta + (1 - \alpha)f(\bar{x})$. Portanto,

$$\begin{aligned} |f(x) - f(\bar{x})| &\leq \alpha(\beta - f(\bar{x})) \\ &= \frac{\beta - f(\bar{x})}{\delta} \|x - \bar{x}\| \\ &= L\|x - \bar{x}\|, \end{aligned}$$

onde $L = (\beta - f(\bar{x}))/\delta$ depende de \bar{x} e da escolha de δ , mas não depende de $x \in \{x \in \mathbb{R}^n \mid 0 < \|x - \bar{x}\| < \delta\}$. \square

Em particular, se $f : D \rightarrow \mathbb{R}$ é convexa no conjunto convexo $D \subset \mathbb{R}^n$, então f é contínua no interior de D . É fácil ver que uma função convexa pode não ser contínua na fronteira de um conjunto fechado. Neste caso, ela não está definida em todo o espaço.

Para ilustrar o que foi enunciado no Teorema 4.16. Consideremos como exemplo, a função $f : D \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} 8, & x = -2 \\ x^2, & x > -2 \end{cases}$$

onde $D = \{x \in \mathbb{R} \mid x \geq -2\}$. Repare que f é convexa em D , e também que seu epígrafo é convexo, contudo f não é contínua no ponto $x = -2$. Veja a Figura 4.7.

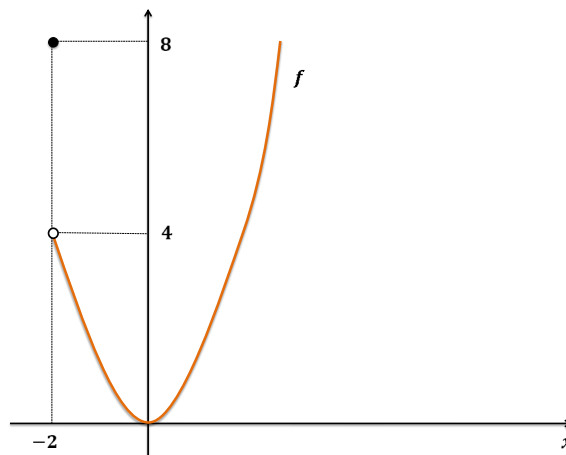


Figura 4.7: A função f é convexa em $D = \{x \in \mathbb{R} \mid x \geq -2\}$, mas não é contínua no ponto $x = -2$ da fronteira de D .

Se $f : D \rightarrow \mathbb{R}$ é convexa no conjunto convexo $D \subset \mathbb{R}^n$, então f é contínua no interior de D . A função do exemplo acima não está definida em todo o espaço.

O resultado do Teorema 4.17 a seguir garante a eficácia e a eficiência na forma de se resolver um problema de minimização convexo, isto é, onde $D \subset \mathbb{R}^n$ é um conjunto

convexo e $f : D \rightarrow \mathbb{R}$ é uma função convexa no conjunto D .

Teorema 4.17. (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 é minimizador global. Além disso, o conjunto de minimizadores é convexo. Se f é estritamente convexa, não pode haver mais de um minimizador.*

Demonstração.

Suponhamos que $\bar{x} \in D$ seja um minimizador local que não é global. Então existe $y \in D$, tal que $f(y) < f(\bar{x})$.

Como D é convexo, temos que $z = \alpha y + (1 - \alpha)\bar{x}$, $z \in D$ e $\alpha \in [0, 1]$. E ainda, como f é convexa, temos que

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

Se considerarmos $\alpha > 0$ suficientemente pequeno, podemos afirmar que z é arbitrariamente próximo a \bar{x} , e como $f(z) < f(\bar{x})$ e $z \in D$, temos uma contradição na hipótese de \bar{x} ser minimizador local que não seja global. Portanto, toda solução local deve ser global. \square

Agora, vamos comprovar a unicidade de minimizadores globais. Consideremos $S \subset D$ o conjunto dos minimizadores globais e $\bar{v} \in \mathbb{R}$ o valor ótimo (mínimo) do problema, isto é, $f(x) = \bar{v}$ para qualquer $x \in S$.

Para todo $x, \bar{x} \in S$ e $\alpha \in [0, 1]$ e pela convexidade de f , temos

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

o que mostra que S é convexo, pois $f(\alpha x + (1 - \alpha)\bar{x}) = \bar{v}$.

Suponhamos agora que f seja estritamente convexa e que existam $x, \bar{x} \in S$ com $x \neq \bar{x}$ e $x \in (0, 1)$. Como x e \bar{x} são minimizadores globais e $\alpha x + (1 - \alpha)\bar{x} \in D$, por D ser convexo, temos

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

Mas, pela convexidade estrita,

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

o que é uma contradição. Portanto, o minimizador é único. O Teorema 4.18 a seguir, afirma que a composição de funções convexas, pode ser convexa, desde que uma delas seja não-decrescente.

Quando a função objetivo de um problema de otimização for uma função fortemente convexa num conjunto fechado e não-vazio, então ele terá solução, que será única quando o conjunto for convexo. O resultado a seguir é de grande importância em computação.

Teorema 4.18. *(Compactidade de conjunto de nível de uma função fortemente convexa)* Suponhamos que a função $f: \mathbb{R}^n \rightarrow \mathbb{R}$ seja fortemente convexa no \mathbb{R}^n . Então o conjunto de nível

$$L_{f, \mathbb{R}^n}(c) = \{x \in \mathbb{R}^n \mid f(x) \leq c\}$$

é compacto para todo $c \in \mathbb{R}$.

Demonstração.

Pelos Teoremas 4.12 e 4.16, o conjunto $L_{f, \mathbb{R}^n}(c)$ é convexo e fechado. Suponhamos que $L_{f, \mathbb{R}^n}(c)$ seja ilimitado. Neste caso, existe uma direção de recessão [10] $d \in \mathfrak{R}_{L_{f, \mathbb{R}^n}(c)} \setminus \{0\}$, isto é, a semirreta $\{x + td \mid t \in \mathbb{R}_+\}$ pertence a $L_{f, \mathbb{R}^n}(c)$.

Seja $\gamma > 0$ o módulo de convexidade forte de f . Fixemos $t > 0$. Para todo $q > t$, obtemos que

$$\begin{aligned} f(x + td) &= f\left(\left(\frac{t}{q}\right)(x + qd) + \left(1 - \frac{t}{q}\right)x\right) \\ &\leq \left(\frac{t}{q}\right)f(x + qd) + \left(1 - \frac{t}{q}\right)f(x) - \gamma\left(\frac{t}{q}\right)\left(1 - \frac{t}{q}\right)\|x + qd - x\|^2 \\ &\leq \left(\frac{t}{q}\right)c + \left(1 - \frac{t}{q}\right)c - \gamma\left(\frac{t}{q}\right)\left(1 - \frac{t}{q}\right)q^2\|d\|^2 \\ &= c - \gamma t(q - t)\|d\|^2. \end{aligned}$$

Como $f(x + td)$ é um número fixo e

$$c - \gamma t(q - t)\|d\|^2 \rightarrow -\infty \quad \text{quando } q \rightarrow +\infty,$$

temos uma contradição. Concluímos que $L_{f, \mathbb{R}^n}(c)$ é ilimitado. □

Lema 4.19. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função fortemente convexa e $D \subset \mathbb{R}^n$, $D \neq \emptyset$, um conjunto fechado qualquer. Então f tem um minimizador em D . Se D é convexo, o minimizador é único.*

Demonstração.

Suponhamos que existe $c \in \mathbb{R}$ tal que o conjunto de nível $L_{f,D}(c)$ seja não vazio e compacto. Pelo Teorema de Weierstrass, o problema

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

tem solução global, digamos \bar{x} . Logo, para todo $x \in D \setminus L_{f,D}(c)$, temos que $f(x) > c \geq f(\bar{x})$, o que mostra que \bar{x} é um minimizador global de f não só em $L_{f,D}(c)$, mas também em D . Considerando também o Teorema 4.18, concluímos que para D não vazio fechado qualquer, f tem um minimizador em D e pelo Teorema 4.17, ele deve ser único. \square

4.3 Funções Convexas Diferenciáveis

Estudaremos nesta seção a caracterização de funções convexas diferenciáveis. Temos que quando uma função é diferenciável, a convexidade admite várias caracterizações que são muito úteis para determinar se uma função é convexa ou não.

Teorema 4.20. *(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 propriedades seguintes são equivalentes:*

- (1) A função f é convexa em D .
- (2) Para todo $x \in D$ e todo $y \in D$,

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

- (3) 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

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

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

Demonstração.

Mostraremos primeiro que (1) \Rightarrow (2) \Rightarrow (3).

Seja f convexa. Para $x \in D$, $y \in D$ e $\alpha \in (0, 1]$ quaisquer. Definindo $d = y - x$, temos que

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

donde

$$\alpha(f(y) - f(x)) \geq f(x + \alpha d) - f(x).$$

Agora, dividindo ambos os lados da desigualdade acima por $\alpha > 0$, e passando o limite quando $\alpha \rightarrow 0_+$, obtemos

$$\begin{aligned} f(x) - f(y) &\geq \lim_{\alpha \rightarrow 0_+} \frac{f(x + \alpha d) - f(x)}{\alpha} \\ &= \langle f'(x), d \rangle \\ &= \langle f'(x), y - x \rangle. \end{aligned}$$

Trocando agora o papel de x e y no item (2), temos

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

Somando esta desigualdade com a de (2), imediatamente obtemos (3).

Mostraremos agora que (3) \Rightarrow (2) \Rightarrow (1). Sejam $x \in D$ e $y \in D$. Pelo Teorema do Valor Médio, existe $\alpha \in (0, 1)$ tal que

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

Usando (3) para os pontos $(x + \alpha(y - x))$ e x , obtemos

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

Com isso temos (2).

Definindo novamente $d = y - x$, temos

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

$$f(y) \geq f(x + \alpha d) - (1 - \alpha)\langle f'(x + \alpha d), d \rangle,$$

onde usamos o item (2) para os pontos x e $x + \alpha d$; y e $x + \alpha d$, respectivamente. Multiplicando a primeira desigualdade por $1 - \alpha \geq 0$ e a segunda por $\alpha \geq 0$, e somando, obtemos

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

o que mostra que f é convexa.

Suponhamos agora que f seja duas vezes diferenciável em D . É suficiente mostrar que (2) \Leftrightarrow (4).

Fixemos $x \in D$ e $d \in \mathbb{R}^n$ quaisquer. Como D é aberto, $x + \alpha d \in D$ para todo $\alpha > 0$ suficiente pequeno. Por (2),

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

Usando ainda a diferenciabilidade de f ,

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

Dividindo por $\alpha^2 > 0$ e tomando o limite quando $\alpha \rightarrow 0+$, obtemos (4).

Sejam $x \in D$ e $y \in D$ quaisquer. Pelo Teorema do Valor Médio, existe $\alpha \in (0, 1)$ tal que

$$f(y) - f(x) - \langle f'(x), y - x \rangle = \frac{1}{2}\langle f''(x + \alpha(y - x))(y - x), y - x \rangle \geq 0,$$

onde a desigualdade segue de (4). Portanto, (4) \Rightarrow (2). □

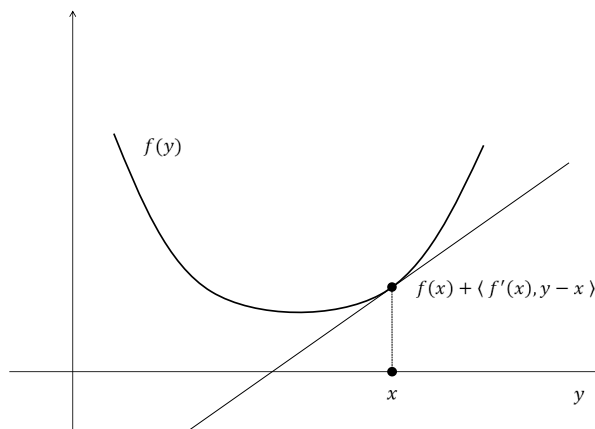


Figura 4.8: Ilustração do Teorema 4.20(2). Para todo y , tem-se que $f(y) \geq f(x) + \langle f'(x), y - x \rangle$, onde x é fixo, porém arbitrário.

O item (2) do Teorema 4.20 diz que, em todo ponto, a aproximação de primeira ordem de uma função convexa sempre está abaixo do gráfico da função. Veja a Figura 4.8. Enquanto que o item (3) diz que o gradiente de uma função convexa é monótono.

Considerações Finais e Perspectivas

A análise convexa constitui um pré-requisito básico ao estudante que deseja entender questões relacionadas à pesquisa operacional e otimização. O arcabouço teórico indispensável na resolução de problemas relacionados à Otimização encontra-se no bojo dos conjuntos e funções convexas. Mediante o estudo de algumas definições e a demonstração de alguns teoremas cria-se um ambiente seguro, onde o pesquisador responsável por modelar os problemas pode caminhar sem temor em estar utilizando um procedimento incorreto. É importante mencionar que desde o advento dos computadores, o apelo à pesquisa operacional vem crescendo de forma sistemática e, sincronicamente, percebe-se a necessidade de se obter cada vez mais conhecimento sobre Análise Convexa.

Nesse sentido, embora haja muitas aplicações relacionando os Conjuntos e as Funções Convexas, percebe-se que os professores de matemática da educação básica não possuem um conhecimento aprofundado sobre esses temas. Assim sendo, esperamos que este material cumpra uma função didática, no sentido que todos aqueles que desejam iniciar seus estudos em Análise Convexa, o utilizem com eficácia. Como o leitor pôde perceber, os principais teoremas, lemas e corolários sobre conjuntos e funções convexas foram demonstrados e as principais definições foram estabelecidas. E ainda, afim de tornar o texto autocontido, colocamos no Capítulo 2 os pré-requisitos fundamentais que o leitor necessita. Dessa forma, acreditamos que o texto tenha se tornado inteligível, principalmente aos iniciantes nestes temas.

Por fim, há alguns aspectos interessantes relacionados à Análise Convexa que ainda podem ser estudados, tais como: funções convexas diferenciáveis e não-diferenciáveis; problemas de otimização. Além disso, materiais didáticos voltados a um público mais amplo que envolva tais temas também merecem ser desenvolvidos.

Referências Bibliográficas

- [1] ANDRADE, E. L. **Introdução à Pesquisa Operacional: Métodos e Modelos para Análise de Decisão**. 3. ed. São Paulo: LTC, 2004.
- [2] BARBOSA, R. M. **Programação Linear**. São Paulo: Nobel, 1973.
- [3] BOLDRINI, J. L. **Álgebra Linear**. São Paulo: Harper & Row do Brasil, 1980.
- [4] BORWEIN, J. M.; VANDERWERFF, D. **Convex Functions**. Cambridge: Encyclopedia of mathematics, 2009.
- [5] BRONSON, R. **Pesquisa Operacional**. São Paulo: Schaum McGraw-Hill, 1985.
- [6] EKELAND, I.; TÉMAN, R. **Complex Analysis and Variational Problems**. Classics in applied mathematics, SIAM, 1999.
- [7] ELSEIT, H. A.; SANDBLOM, C.L. **Linear Programming and its applications**. Springer-Verlag Berlin, 2007.
- [8] GOLDBARG, M. C.; LUNA, H. P. **Otimização, Combinatória e Programação Linear**. Rio de Janeiro: Campus, 2000.
- [9] HEFEZ, A.; FERNANDEZ, C. S. **Introdução à Álgebra Linear**. Rio de Janeiro: SBM, 2012.
- [10] IZMAILOV, A.; SOLODOV, M. **Otimização Vol. 1**. Rio de Janeiro: IMPA, 2009.
- [11] JÄNICH, K. **Álgebra Linear**. São Paulo: LTC, 2002.
- [12] KOLMAN, B. **Introdução à Álgebra Linear com Aplicações**. 6 ed. Rio de Janeiro: LTC, 1999.
- [13] LIMA, E. L. **Análise Real**. vol 1. Rio de Janeiro: IMPA, 1999.
- [14] LIMA, E. L. **Análise Real**. vol 2. Rio de Janeiro: IMPA, 2004.
- [15] MIRSHAWKA, V. **Programação Linear**. 1 ed. São Paulo: Nobel, 1971.

-
- [16] PUCCINI, A. L. **Introdução a Programação Linear**. 1 ed. Rio de Janeiro: LTC, 1972.
- [17] SHAMBLIM, J. E.; STEVENS JR, G. T. **Programação Linear: uma abordagem básica**. São Paulo: Atlas, 1989.
- [18] SHOKRANIAN, S. **Introdução à Álgebra Linear**. Brasília: Editora UnB, 2002.
- [19] SILVA, E. M. et al. **Pesquisa Operacional**. São Paulo: Atlas, 1998.