

# Matemática para o exame da ANPEC <sup>1</sup>

Alexandre L. Madureira

LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA—LNCC, BRASIL  
*URL:* <http://www.lncc.br/~alm>

---

<sup>1</sup>04 de setembro de 2012

RESUMO. Estas notas de aula são relativas ao curso de preparação promovido pela FGV para a parte de matemática do exame da ANPEC. Estas notas devem servir de apoio, e certamente não eliminam a necessidade de se usar os já clássicos, aprimorados e vários livros didáticos. Mencionamos alguns deles na bibliografia.

Neste curso apresento alguns tópicos de álgebra linear, cálculo e análise que estão presentes no exame da ANPEC, e que são importantes para uma formação mais sólida de futuros pós-graduandos em economia. Espero apresentar algum rigor matemático aos alunos, e mostrar como este deve ser utilizado em conjunto com a intuição matemática, nunca esquecendo o objetivo que é aprimorar a arte de resolver questões.

Uma particularidade das notas é que, ao fim destas há soluções de questões das provas da ANPEC de matemática dos últimos anos. Isto não seria possível sem a ajuda de Gustavo Lopo Andrade, Gustavo Pereira, e Lucas Alves, que gentilmente concordaram em apresentar suas soluções em  $\text{\TeX}$ . Acho que estas soluções serão úteis para a grande comunidade de alunos que se prepara para os exames da ANPEC. Meus agradecimentos mais sinceros a estes três jovens!

Eu tomei a liberdade de modificar minimamente a notação usada em algumas das questões, a fim de torná-la homogênea e coincidir com as notações usadas nestas notas. Editei também minimamente as questões submetidas pelos alunos, a fim de tornar suas (deles) soluções mais próximas do estilo, linguagem e notações usadas no restante das notas.

São usadas estas notas várias ideias e notações de outros livros, como [4, 13, 20] em álgebra linear. A bibliografia básica sugerida pela ANPEC é dada por [4, 5, 23], e a complementar é [1, 9, 10, 13, 27].

## Sumário

Capítulo 1. Noções de Conjuntos	1
1.1. Conjuntos	1
1.2. Exercícios	4
Capítulo 2. Noções de geometria analítica	5
2.1. Coordenadas	5
2.2. Distância, norma, produtos escalar e vetorial	6
2.3. Produto vetorial	7
2.4. A reta no plano e espaço	8
2.5. Planos no espaço	9
2.6. Desigualdade lineares	9
2.7. Cônicas no plano	10
2.8. Exercícios	12
Capítulo 3. Álgebra Linear	13
3.1. Operações com matrizes	13
3.2. Matriz inversa, transposta e adjunta	13
3.3. Resolução de sistemas lineares	14
3.4. Determinantes e a regra de Cramer	15
3.5. Espaços vetoriais, subespaços, base e dimensão	17
3.6. Produto interno, ortogonalidade e projeções	20
3.7. Transformações lineares, núcleo, imagem e representações matriciais	22
3.8. Autovalores, polinômios característicos e operadores diagonalizáveis	24
3.9. Operadores auto-adjuntos, operadores ortogonais	26
3.10. Formas lineares e bilineares	27
3.11. Exercícios	28
Capítulo 4. Limites de funções	29
4.1. Definições básicas envolvendo funções	29
4.2. Intervalos na reta	30
4.3. Funções inversas	31
4.4. Limites de funções	32
4.5. Limites laterais, infinitos e no infinito	35
4.6. Exercícios	36
Capítulo 5. Continuidade e Funções Contínuas	37
5.1. Introdução e exemplos	37
5.2. Funções Contínuas em intervalos fechados e limitados	39

5.3. Exercícios	39
Capítulo 6. Diferenciação	41
6.1. Definições e Exemplos	41
6.2. Propriedades da Derivada	42
6.3. Aplicações	44
6.4. Teorema de Taylor e Aplicações	46
6.5. Regra de L'Hôpital	49
6.6. Exercícios	50
Capítulo 7. Funções trigonométricas, logarítmicas e exponenciais	51
7.1. Funções trigonométricas	51
7.2. Funções log e exponencial	52
Capítulo 8. Integração	55
8.1. Propriedade básicas de integrais de funções limitadas	55
8.2. Áreas planas	58
8.3. Integrais impróprias	59
Capítulo 9. Sequências e Séries	61
9.1. Definição e resultados preliminares	61
9.2. Sequências Monótonas	68
9.3. Séries	70
9.4. Exercícios	72
Capítulo 10. Funções de várias variáveis	73
10.1. Introdução	73
10.2. Derivadas parciais e planos tangentes	73
10.3. Diferenciabilidade	74
10.4. Matriz Hessiana, Fórmula de Taylor e pontos críticos	79
10.5. Teorema da Função Inversa e da Função Implícita	85
10.6. Minimização com restrições	87
10.7. Exercícios	91
Apêndice A. Uma introdução não tão formal aos fundamentos da matemática	93
A.1. Argumentação formal	93
A.2. Demonstração por indução e contradição	96
A.3. Exercícios	97
Apêndice B. Soluções das provas da ANPEC	99
Apêndice. Referências Bibliográficas	119

Os tópicos destas notas seguem a orientação da própria ANPEC. São eles:

- (1) Noção de Conjunto  
Relação de pertinência. Relação de inclusão, operações de interseção, união, diferença. Produto cartesiano. Relações.
- (2) Noções de Geometria Analítica  
Coordenadas no plano e no espaço. Fórmulas de distância. Vetores livres no plano e no espaço. Produto escalar, produto vetorial, perpendicularidade. Equações da reta no plano e no espaço, equações de planos. Inequações lineares. Parábola e hipérbole.
- (3) Funções  
Funções injetoras, sobrejetoras e bijetoras. Representação gráfica. Soma, diferença, produto, quociente e composição de funções.
- (4) Álgebra Linear  
Operações com matrizes. Matriz inversa, transposta e adjunta. Resolução de sistemas lineares. Determinantes. Regra de Cramer. Espaços vetoriais. Subespaços. Base e dimensão. Produto interno, ortogonalidade. Projeções. Transformações lineares. Núcleo e imagem. Matriz de uma transformação linear. Autovalores e autovetores. Polinômios característicos operadores diagonalizáveis. Operadores auto-adjuntos, operadores ortogonais. Formas bilineares.
- (5) Funções de uma variável real -  
Limites. Funções contínuas. Funções deriváveis. Reta tangente e reta normal. Regras de derivação: derivada da soma, do produto, do quociente, regra da cadeia, derivada da inversa. Elasticidade. Derivadas sucessivas. Funções trigonométricas. Função exponencial e logarítmica. Regra de L'Hôpital. Intervalos de concavidade e convexidade. Ponto de inflexão. Polinômio de Taylor.
- (6) Integrais  
Teorema fundamental do cálculo, primitivação por partes e por substituição. Áreas planas. Integrais impróprias.
- (7) Sequências e séries  
Convergência e divergência de seqüências e séries. Série geométrica, teste da comparação, da razão, da raiz, teste da integral. Séries alternadas.
- (8) Matemática financeira  
Juros simples. Juros compostos. Desconto e taxa de desconto. Séries de pagamento. Fluxo de caixa. Sistema de amortização.
- (9) Funções de várias variáveis reais  
Derivadas parciais. Diferencial total. Gradiente. Regra da cadeia. Funções implícitas. Teorema do envelope. Funções homogêneas. Teorema de Euler. Condições de 1ª e 2ª ordens para máximos e mínimos de funções de várias variáveis reais. Condições de 1ª e 2ª ordens para otimização condicionada com restrições de igualdade e desigualdade. Integrais duplas. Mudança de variáveis em integrais duplas.
- (10) Equações diferenciais e em diferenças  
Equações lineares de 1ª ordem e equações lineares de 2ª ordem com coeficientes constantes. Sistema de duas equações lineares de 1ª ordem homogêneo com coeficientes constantes.



## CAPÍTULO 1

### Noções de Conjuntos

1

Neste capítulo falaremos sobre conjuntos, e em particular descreveremos relações de pertinência e inclusão, operações de interseção, união, diferença, produto cartesiano, e relações.

#### 1.1. Conjuntos

Esta parte do texto pretende apenas expor algumas dificuldades básicas, da parte talvez mais fundamental da matemática (excluindo-se a lógica). Duas referências também introdutórias, mas muito mais completas, são os livros do Terence Tao [25], e do Paul Halmos [17].

A primeira dificuldade encontrada é definir o que é um conjunto. Uma saída (questionável) é simplesmente dizer que um conjunto é uma “coleção” ou família de objetos (ou elementos ou membros). Se um objeto  $x$  faz parte de um conjunto  $A$ , dizemos que ele pertence à  $A$  e escrevemos  $x \in A$  (o símbolo  $\notin$  indica que quando um elemento não pertence a um conjunto).

Espera-se que o uso da palavra “coleção” acima não traga confusões. O termo coleção será a seguir utilizado para conjuntos cujos elementos são também conjuntos.

Considere agora dois conjuntos  $A$  e  $B$ .

- Dizemos que  $A$  está contido em  $B$  e escrevemos  $A \subseteq B$  se todo elemento de  $A$  é elemento de  $B$ . Pode-se também escrever  $B \supseteq A$  (lê-se  $B$  contém  $A$ ) para indicar  $A \subseteq B$ .
- Se  $A$  não está contido em  $B$  escrevemos  $A \not\subseteq B$ .
- Dizemos que dois conjuntos  $A$  e  $B$  são iguais, e escrevemos  $A = B$  se  $A \subseteq B$  e  $B \subseteq A$ .
- Se não forem iguais, dizemos que são diferentes e escrevemos  $A \neq B$ .
- Também escrevemos  $A \subsetneq B$  se  $A \subseteq B$  mas  $A \neq B$ . Dizemos neste caso que  $A$  está *propriamente* contido em  $B$ .

O seguinte axioma é importante, nos garante que a “forma usual” de definir conjuntos é “segura,” ou seja, quando definimos um conjunto obtemos um e apenas um conjunto (mesmo que seja vazio).

**AXIOMA 1.1.1** (da especificação). Seja  $A$  um conjunto, e para cada  $x \in A$ , seja  $P(x)$  uma afirmativa (verdadeira ou falsa). Então existe um único conjunto  $B$  composto de todos os elementos  $x$  de  $A$  tais que  $P(x)$  é verdade.

---

<sup>1</sup>Última Atualização: 17/06/2012

O conjunto acima é denotado por  $\{x \in A : P(x) \text{ é verdade}\}$ . Quando o conjunto  $A$  é claro pelo contexto, podemos escrever simplesmente  $\{x : P(x) \text{ é verdade}\}$ . Este conjunto é formado por *todos os elementos*  $x$  que estejam em  $A$  e tais que a propriedade  $P(x)$  seja verdadeira. Uma última forma de denotar os conjuntos é simplesmente descrever seus elementos entre as chaves. Por exemplo, o conjunto dos números pares pode ser denotado por

$$\{x \in \mathbb{Z} : x \text{ é divisível por } 2\}.$$

Sendo um pouco menos formal, pode-se escrever este mesmo conjunto como  $\{2x : x \in \mathbb{Z}\}$  ou ainda enumerar todos os elementos do conjunto:  $\{\dots, -4, -2, 0, 2, 4, 6, \dots\}$ .

Vale aqui descrever uma situação interessante dada pelo *Paradoxo de Russel*. É natural perguntar-se o quão grande podem ser os conjuntos. Por exemplo, existe um conjunto  $U$  tal que todos os conjuntos existentes sejam *elementos* de  $U$ ? Se  $U$  existe, então, pelo Axioma da especificação (Axioma 1.1.1) podemos formar

$$R = \{x \in U : x \text{ é conjunto e } x \notin x\}.$$

Então  $R \notin U$ . De fato, se  $R \in U$ , então  $R \in R$  ou  $R \notin R$ . Vamos dividir em dois casos:

- (1) Se  $R \in R$ , então  $R \notin R$  pois por definição,  $R$  é formado pelos conjuntos que *não* se autocontém.
- (2) Se  $R \notin R$ , então  $R$  não satisfaz as propriedades que definem  $R$ . No caso de *não* se autoconter. Logo  $R \in R$ .

Em ambas possibilidades (1) e (2) obtemos absurdos. Logo  $R \notin U$ . Mas  $U$  é exatamente o conjunto que contém *todos* os outros. . . . Somos levados a concluir que tal conjunto  $U$  não pode existir.

O próximo passo é definir as operações usuais. Por incrível que possa parecer, o mais difícil é definir a união entre dois conjuntos, e para isto é necessário um axioma.

AXIOMA 1.1.2 (da união). Para qualquer coleção de conjuntos, existe um conjunto que contém todos os elementos pertencentes a pelo menos um conjunto da coleção.

Podemos agora definir a união entre dois conjuntos  $A$  e  $B$ . Para tanto, note que pelo Axioma da união, existe um conjunto  $U$  que contém todos os elementos de  $A$  e de  $B$ . Definimos então  $A \cup B = \{x \in U : x \in A \text{ ou } x \in B\}$ .

Observe entretanto a seguinte armadilha. O Axioma da união não garante que o tal conjunto contendo  $A$  e de  $B$  é único, somente garante que existe. Podemos ter por exemplo um outro conjunto  $\hat{U}$  contendo  $A$  e de  $B$ . Seja agora  $C = \{x \in \hat{U} : x \in A \text{ ou } x \in B\}$ . Para a união ser definida de forma única, temos que garantir que  $C = A \cup B$ . Isto é verdade, e para provar basta argumentar que  $C \subseteq A \cup B$  e  $C \supseteq A \cup B$ .

Com o Axioma da especificação, podemos definir as seguintes operações.

- O conjunto interseção entre  $A$  e  $B$  é  $A \cap B = \{x \in A : x \in B\}$ .
- O conjunto diferença  $A$  menos  $B$  é  $A \setminus B = \{x \in A : x \notin B\}$ . O conjunto resultante também denotado por  $A - B$  e chamado de complemento de  $B$  em relação à  $A$ .
- Quando é claro quem é o conjunto  $A$ , denotamos  $A \setminus B$  por  $\mathcal{C}(B)$ , e o chamamos de complemento de  $B$ .



OBSERVAÇÃO. É fácil generalizar os conceitos acima para uniões e interseções arbitrárias de conjuntos.

Finalmente, é útil a regra de *De Morgam*, que diz que para conjuntos  $E_n$ , onde  $n \in \mathbb{N}$ , temos que

$$(1.1.1) \quad \mathcal{C}(\cup_{i \in \mathbb{N}} E_n) = \cap_{i \in \mathbb{N}} \mathcal{C}(E_n), \quad \mathcal{C}(\cap_{i \in \mathbb{N}} E_n) = \cup_{i \in \mathbb{N}} \mathcal{C}(E_n).$$

Outro conceito útil é o de *par ordenado*. Dados dois elementos, ou objetos  $a$  e  $b$ , formamos o par  $(a, b)$ , e chamamos  $a$  e  $b$  de (primeiro e segundo) componentes de  $(a, b)$ . Dizemos (definimos) que um par ordenado é *igual* a outro se os respectivos componentes forem iguais, i.e.,  $(a, b) = (a', b')$  se  $a = a'$  e  $b = b'$ .

Do ponto de vista axiomático, não é claro que dados dois elementos, exista o par ordenado formado por eles. Viveremos por enquanto com esta dúvida. O importante é como pares ordenados são formados (por elementos de dois conjuntos) e quando são iguais (quando os componentes são iguais).

Definimos agora *produtos cartesianos*. Dados dois conjuntos  $A$  e  $B$ , definimos o conjunto  $A \times B = \{(a, b) : a \in A, b \in B\}$  como sendo o composto pelos pares ordenados.

OBSERVAÇÃO. A extensão destes conceitos para  $n$ -úplas ordenadas e produtos cartesianos com  $n$  conjuntos é natural.

Chamamos  $R$  de *relação entre  $A$  e  $B$*  se  $R$  é subconjunto de  $A \times B$ . Similarmente, dizemos que  $a \in A$  e  $b \in B$  são *relacionados* se  $(a, b) \in R$ . Desta definição vem o importante conceito de função. Uma *função entre  $A$  e  $B$*  nada mais é que uma relação entre  $A$  e  $B$ , e sendo assim  $f \subseteq A \times B$ . Esta relação entretanto satisfaz a seguinte restrição: para todo  $a \in A$  existe um único  $b \in B$  tal que  $(a, b) \in f$ . Denotamos esta relação especial por  $f : A \rightarrow B$ . Dado  $a \in A$ ,  $b \in B$ , dizemos que  $f(a) = b$  se  $(a, b) \in f$ .

Na prática, comumente nos "esquecemos" desta definição e tratamos funções de forma mais informal e direta. Este peccadilho matemático não chega a atrapalhar nossos objetivos, mas é importante ter em mente a definição formal.

Uma relação  $R \subseteq A \times A$  é uma ordenação parcial se

- (1)  $(a, b) \in R$  e  $(b, c) \in R$  implica em  $(a, c) \in R$ ,
- (2)  $(a, b) \in R$  e  $(b, a) \in R$  implica em  $a = b$ ,

e é uma ordenação simples se, além disto,

- (1)  $(a, b) \in R$  ou  $(b, a) \in R$  para todo  $a, b \in A$ .

Um exemplo de ordenação parcial é dada pela relação de pertinência ( $\subseteq$ ) entre conjuntos. Um exemplo de ordenação simples é dada nos reais, com a relação de maior. Finalmente, temos que  $R \subseteq A \times A$  é uma relação de equivalência se para todos elementos  $a, b$  e  $c \in A$  temos

- (1)  $(a, a) \in R$ ,
- (2)  $(a, b) \in R$  implica em  $(b, a) \in R$ ,
- (3)  $(a, b) \in R$  e  $(b, c) \in R$  implica em  $(a, c) \in R$ .

**1.2. Exercícios**

EXERCÍCIO 1.1. Mostre que

- (1)  $\{x \in \mathbb{R} : x^2 \geq 0\} = \mathbb{R}$ .
- (2)  $\{x \in \mathbb{R} : x > 0\} \subsetneq \{x \in \mathbb{R} : x^2 \geq 0\}$ .
- (3)  $\mathbb{R} \not\subseteq \{x \in \mathbb{R} : x^2 \geq 0\}$ .

EXERCÍCIO 1.2. Mostre a regra de *De Morgan* dada em (1.1.1).

EXERCÍCIO 1.3. Mostre que  $\{a, a\} = \{a\}$ .

EXERCÍCIO 1.4. Sejam  $A$  e  $B$  dois conjuntos disjuntos, i.e.,  $A \cap B = \emptyset$ . Seja  $X = A \cup B$ . Mostre que  $A = X \setminus B$  e  $B = X \setminus A$ .

EXERCÍCIO 1.5. Sejam  $A$  e  $B$  dois conjuntos, e  $C = (A \setminus B) \cup (B \setminus A)$ . Mostre que  $C = (A \cup B) \setminus (A \cap B)$  e que  $C \cap A \cap B = \emptyset$ .

EXERCÍCIO 1.6. Mostre que a relação de pertinência ( $\subseteq$ ) entre conjuntos define uma ordenação parcial, e que a relação de maior nos reais define uma ordenação simples.

## CAPÍTULO 2

### Noções de geometria analítica

1

Neste capítulo falaremos sobre noções como coordenadas, distância, vetores, produtos escalar e vetorial, perpendicularidade, equações da reta no plano e espaço, equações de planos, inequações lineares, parábolas, hipérbolas.

Consideraremos o  $\mathbb{R}^n$  o como o conjunto das  $n$ -úplas ordenadas de números reais, como definido abaixo.

**DEFINIÇÃO 2.0.1.** *Seja  $\mathbb{R}^n$  o conjunto das  $n$ -úplas ordenadas de números reais, i.e.,*

$$\mathbb{R}^n = \{\mathbf{x} = (x_1, \dots, x_n) : x_i \in \mathbb{R} \text{ para } i = 1, \dots, n\}.$$

*Definimos então as operações produto por escalar e soma da seguinte forma:*

$$\alpha \mathbf{x} = (\alpha x_1, \dots, \alpha x_n), \quad \mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n),$$

*onde  $\mathbf{x} = (x_1, \dots, x_n)$  e  $\mathbf{y} = (y_1, \dots, y_n)$  estão em  $\mathbb{R}^n$ , e  $\alpha \in \mathbb{R}$ . Pode-se checar que  $\mathbb{R}^n$  é espaço vetorial com as operações acima descritas.*

#### 2.1. Coordenadas

Seja  $\mathcal{B}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  base do  $\mathbb{R}^n$ . Então, segundo o Teorema 3.5.11, todo vetor do  $\mathbb{R}^n$  pode ser escrito de forma única como combinação linear dos vetores de  $\mathcal{B}$ , i.e., dado um vetor  $\mathbf{w} \in \mathbb{R}^n$  qualquer, existem números reais  $\alpha_1, \dots, \alpha_n$  que são os únicos tais que

$$\mathbf{w} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n.$$

Dizemos então que  $\alpha_1, \dots, \alpha_n$  são as *coordenadas* de  $\mathbf{w}$  na base  $\mathcal{B}$ .

A base mais simples que existe é a *base canônica*, dada por  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ , onde, para  $i \in \{1, \dots, n\}$ , o vetor  $\mathbf{e}_i$  é definido tal que a  $i$ -ésima coordenada vale um e as demais coordenadas valem zero, i.e.,

$$\mathbf{e}_1 = (1, 0, 0, \dots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \dots, 0), \quad \dots, \quad \mathbf{e}_n = (0, 0, \dots, 0, 1).$$

Chamamos estes vetores de *vetores da base canônica*. Note que podemos escrever um ponto  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  como  $\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n$ . Neste caso,  $x_1, \dots, x_n$  são as *coordenadas de  $\mathbf{x}$  na base canônica*.

Existe uma identificação natural dos pontos em  $\mathbb{R}^n$  com suas coordenadas na base canônica. Usaremos neste texto a seguinte notação. Para cada  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ , indicaremos

---

<sup>1</sup>Última Atualização: 18/06/2012

por  $\vec{\mathbf{x}} \in \mathbb{R}^{n \times 1}$  a matriz coluna das coordenadas na base canônica dada por

$$(2.1.1) \quad \vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Na verdade não seremos tão preciosistas e escreveremos que  $\vec{\mathbf{x}} \in \mathbb{R}^n$  também.

EXEMPLO 2.1. Nem sempre as bases são tão simples. Por exemplo,  $\{(1, 1); (0, 1)\}$  determina uma base em  $\mathbb{R}^2$ . Para determinar as coordenadas de um vetor  $(a, b)$  qualquer em  $\mathbb{R}^2$ , temos que achar  $\alpha_1, \alpha_2 \in \mathbb{R}$  tais que

$$(a, b) = \alpha_1(1, 1) + \alpha_2(0, 1).$$

Nesta base, o vetor  $(1, 1)$  tem 1 e 0 como coordenadas, e o vetor  $(0, 1)$  tem 0 e 1 como coordenadas, pois

$$(1, 1) = 1 \cdot (1, 1) + 0 \cdot (0, 1),$$

$$(0, 1) = 0 \cdot (1, 1) + 1 \cdot (0, 1).$$

Já o vetor  $(-2, -1)$  tem  $-2$  e  $1$  como coordenadas, pois se

$$(-2, -1) = \alpha_1(1, 1) + \alpha_2(0, 1),$$

então  $\alpha_1 = -2$  e  $\alpha_1 + \alpha_2 = -1$ . Logo  $\alpha_2 = 1$ .

## 2.2. Distância, norma, produtos escalar e vetorial

Dados dois vetores  $\mathbf{x} = (x_1, x_2)$  e  $\mathbf{y} = (y_1, y_2)$  do  $\mathbb{R}^2$ , a distância entre eles é dada pelo “tamanho” do vetor  $\mathbf{x} - \mathbf{y} = (x_1 - y_1, x_2 - y_2)$ .

\*\*\*\*\* por figura aqui \*\*\*\*\*

Para medir tamanho de vetores, usamos a noção de *norma*. No  $\mathbb{R}^2$ , definimos a norma de um vetor  $\mathbf{x} = (x_1, x_2)$  por

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}.$$

Esta é a *norma euclidiana*, que no caso mais geral, em  $\mathbb{R}^n$ , é dada por

$$\|(x_1, \dots, x_n)\| = \sqrt{x_1^2 + \dots + x_n^2}.$$

Voltando ao conceito de distância, temos que a distância entre dois pontos do  $\mathbb{R}^2$  dados por  $\mathbf{x}$  e  $\mathbf{y}$ , nas várias normas, é dada por

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}, \quad \|\mathbf{x} - \mathbf{y}\|_\infty = \max\{|x_1 - y_1|, |x_2 - y_2|\}.$$

EXEMPLO 2.2. Considere o vetor  $(3, 4)$ . Então  $\|(3, 4)\| = \sqrt{9 + 16} = 5$  e  $\|(3, 4)\|_\infty = 4$ .

Os dois exemplos de norma acima são casos particulares da noção mais geral de norma, que se aplica em espaços vetoriais em geral, ver a Definição 3.6.2.

Uma outra importante *ferramenta* matemáticas quando se trabalha em espaços vetoriais é o conceito de produto interno.

Em  $\mathbb{R}^2$ , se  $\mathbf{x} = (x_1, x_2)$ , e  $\mathbf{y} = (y_1, y_2)$ , o produto interno canônico é dado por

$$\mathbf{x} \cdot \mathbf{y} = \vec{\mathbf{x}}^T \vec{\mathbf{y}} = x_1 y_1 + x_2 y_2.$$

Em  $\mathbb{R}^n$ , para  $\mathbf{x} = (x_1, \dots, x_n)$ , e  $\mathbf{y} = (y_1, \dots, y_n)$ , definimos

$$\mathbf{x} \cdot \mathbf{y} = \vec{\mathbf{x}}^T \vec{\mathbf{y}} = x_1 y_1 + \dots + x_n y_n.$$

Note que podemos definir as normas euclidianas usando o produto interno

$$(2.2.1) \quad \|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + \dots + x_n^2}$$

Assim como no caso de norma, um produto interno não precisa ser o canônico, basta obedecer algumas “regras”. Veja a Definição 3.6.1. O que é interessante é que a relação entre norma e produto interno vista em (2.2.1) é somente um exemplo do caso mais geral. Sempre que temos um produto interno, podemos definir uma norma. Isto será visto no Capítulo 3. Abaixo temos a desigualdade de *Cauchy–Schwartz* no  $\mathbb{R}^n$ . Deixaremos a demonstração para o caso geral visto no Teorema 3.6.3.

**TEOREMA 2.2.1.** *Considere a norma e o produto interno canônicos do  $\mathbb{R}^n$ . Então vale a desigualdade de Cauchy–Schwartz*

$$(2.2.2) \quad |\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\| \quad \text{para todo } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Além disto, a igualdade  $|\mathbf{x} \cdot \mathbf{y}| = \|\mathbf{x}\| \|\mathbf{y}\|$  vale se e somente se  $\mathbf{x} = \alpha \mathbf{y}$  para algum  $\alpha \in \mathbb{R}$ .

Finalmente, dados dois vetores  $\mathbf{x}, \mathbf{y}$  do  $\mathbb{R}^n$ , definimos o cosseno do ângulo formado por eles por

$$(2.2.3) \quad \cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

Dizemos então que dois vetores  $\mathbf{x}, \mathbf{y}$  são ortogonais, ou perpendiculares, se  $\mathbf{x} \cdot \mathbf{y} = 0$ . Note que devido à desigualdade de Cauchy–Schwartz (2.2.2) que o cosseno toma sempre valores entre  $-1$  e  $1$ .

### 2.3. Produto vetorial

Uma outra operação com vetores é o *produto vetorial*. Sejam  $\mathbf{x}, \mathbf{y}$  vetores em  $\mathbb{R}^3$ . Então definimos

$$\mathbf{x} \times \mathbf{y} = (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1).$$

Uma outra forma de escrever é

$$\mathbf{x} \times \mathbf{y} = \left( \det \begin{pmatrix} x_2 & x_3 \\ y_2 & y_3 \end{pmatrix}, -\det \begin{pmatrix} x_1 & x_3 \\ y_1 & y_3 \end{pmatrix}, \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \right),$$

onde  $\det(A)$  denota o determinante da matriz  $A$ . Algumas propriedades do produto vetorial são dadas abaixo:

- (1)  $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$
- (2)  $\mathbf{x} \times \mathbf{y}$  é ortogonal a  $\mathbf{x}$  e  $\mathbf{y}$
- (3)  $(\alpha \mathbf{x}) \times \mathbf{y} = \alpha(\mathbf{x} \times \mathbf{y}) = \mathbf{x} \times (\alpha \mathbf{y})$ , para todo  $\alpha \in \mathbb{R}$
- (4)  $(\mathbf{x} + \mathbf{y}) \times \mathbf{z} = \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z}$
- (5)  $\|\mathbf{x} \times \mathbf{y}\|^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2$
- (6)  $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta$

$$(7) \mathbf{x} \times \mathbf{x} = 0$$

$$(8) \mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) = \det \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$$

$$(9) \mathbf{x} \times \mathbf{y} = 0 \text{ se e somente se } \mathbf{x} = \alpha \mathbf{y} \text{ para algum } \alpha \in \mathbb{R}$$

$$(10) \mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3, \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1, \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$$

## 2.4. A reta no plano e espaço

Uma reta é um conjunto de pontos do  $\mathbb{R}^n$  que pode ser definida por um ponto a ela pertencente, e a uma direção dada. Se chamamos de  $r$  uma reta, seja  $\mathbf{x} \in r$  e um vetor  $\mathbf{v}$  na direção de  $r$ . Definimos então

$$r = \{\mathbf{x} + t\mathbf{v} : t \in \mathbb{R}\}.$$

Analogamente, se  $\mathbf{x}, \mathbf{y}$  são dois pontos de  $r$ , então  $\mathbf{v} = \mathbf{x} - \mathbf{y}$  determina a direção da reta.

EXEMPLO 2.3. Seja  $r_1$  reta dada por

$$r_1 = \{(1, 2) + t(3, -1) : t \in \mathbb{R}\}.$$

Ache  $r_2$  passando por  $(1, 1)$  e paralela à  $r_1$ .

**Solução.** A solução é simples pois como  $r_2$  é paralela à  $r_1$ , ambas tem a mesma direção, que no caso é  $(3, -1)$ . Como  $(1, 1) \in r_2$ , então  $r_2 = \{(1, 1) + t(3, -1) : t \in \mathbb{R}\}$ .

O exemplo abaixo lida com interseção de retas.

EXEMPLO 2.4. Determine se as retas  $r_1 = \{(0, 0, 1) + t(1, -1, 3) : t \in \mathbb{R}\}$  e  $r_2 = \{(1, 2, 0) + t(0, 3, -1) : t \in \mathbb{R}\}$  se interseptam, e em qual ponto.

**Solução.** Note que as retas se interseptam se e somente se elas tiverem um ponto em comum, ou seja se existirem  $t, s$  tais que  $(1, -2, 1) + t(1, -1, 3) = (1, 2, 0) + s(0, 3, -1)$ . Isto equivale a resolver o sistema

$$1 + t = 1; \quad -2 - t = 2 + 3s; \quad 1 + 3t = -s.$$

Este sistema de equações pode ter uma, zero ou infinitas soluções.

No exemplo a seguir, consideramos como, dada a equação reta, podemos determinar sua direção.

EXEMPLO 2.5. Seja agora um reta  $r$  no plano, i.e., no  $\mathbb{R}^2$ , dada por  $ax + by + d = 0$ , onde  $a$  e  $b$  não são simultaneamente nulos. Caso  $b = 0$ , temos que a reta é simplesmente dada por  $x - d/a$ , ou seja, é a reta vertical dada por  $x$  constante. Suponha agora  $b \neq 0$ . Para determinarmos sua direção, vamos achar dois pontos pertencentes à  $r$ . Para tal, basta determinar o valor de  $y$  quando  $x$  for igual a zero e a um, por exemplo. No caso temos que  $(0, -d/b)$  e  $(1, (-d-a)/b)$  pertencem à  $r$ . Logo,  $\mathbf{v} = (1, (-d-a)/b) - (0, -d/b) = (1, -a/b)$  é paralelo à  $r$ . Note que  $-a/b$  é exatamente o número que indica a *inclinação* de  $r$ . É interessante e útil notar que  $(a, b)$  é vetor perpendicular à  $r$ .

### 2.5. Planos no espaço

Um plano no espaço é definido por um ponto a ele pertencente e a um vetor ortogonal ao plano. Seja  $\Sigma$  um plano,  $\hat{\mathbf{x}} \in \Sigma$  e  $\mathbf{n}$  vetor perpendicular a  $\Sigma$ . Então

$$\Sigma = \{\mathbf{x} \in \mathbb{R}^3 : (\mathbf{x} - \hat{\mathbf{x}}) \cdot \mathbf{n} = 0\}.$$

Expandindo nas coordenadas, temos que para  $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3)$ , e  $\mathbf{n} = (n_1, n_2, n_3)$ , que um ponto qualquer de  $\Sigma$  satisfaz  $(x_1 - \hat{x}_1)n_1 + (x_2 - \hat{x}_2)n_2 + (x_3 - \hat{x}_3)n_3 = 0$ . Reescrevemos esta equação como  $n_1x + n_2x + n_3x = \hat{x}_1n_1 + \hat{x}_2n_2 + \hat{x}_3n_3 = \hat{\mathbf{x}} \cdot \mathbf{n}$ , que é da forma geral

$$ax_1 + bx_2 + cx_3 = d,$$

com  $a = n_1$ ,  $b = n_2$ ,  $c = n_3$ ,  $d = \hat{\mathbf{x}} \cdot \mathbf{n}$ .

EXEMPLO 2.6. Ache a menor distância do ponto  $\hat{\mathbf{p}} = (1, 0, 1)$  ao plano dado por  $x + 2y - z = 2$ .

**Solução.**

Passo i: precisamos primeiro achar algum ponto  $\mathbf{p}_0$  pertencente ao plano. Por exemplo  $(1, 1, 1)$ .

Passo ii: Seja  $\mathbf{v} = (1, 1, 1) - (1, 0, 1) = (0, 1, 0)$ . Então a projeção de  $\mathbf{v}$  em  $\mathbf{n} = (1, 2, -1)$  é dada por

$$\mathbf{w} = \frac{\mathbf{n} \cdot \mathbf{v}}{\|\mathbf{n}\|^2} \mathbf{n} = \frac{1}{3} \mathbf{n}.$$

Então a distância de  $\hat{\mathbf{p}}$  ao plano é dada simplesmente pela norma de  $\mathbf{w}$ , o seja, a distância é de  $\|\mathbf{w}\| = \|\mathbf{n}\|/3 = \sqrt{6}/3$ .

Passo iii: para achar o ponto  $\mathbf{p}_M$  do plano que tem distância mínima até  $\hat{\mathbf{p}}$ , basta notar que  $\mathbf{p}_M + \mathbf{w} = \hat{\mathbf{p}}$ , e portanto

$$\mathbf{p}_M = \hat{\mathbf{p}} - \mathbf{w}.$$

Outra forma de se definir um plano é, dados três pontos não colineares a ele pertencentes, definir um vetor normal ao plano via produto vetorial. De fato, se  $\mathbf{x}$ ,  $\mathbf{y}$ ,  $\mathbf{z}$  pertencem a um plano, então  $\mathbf{n} = (\mathbf{y} - \mathbf{x}) \times (\mathbf{z} - \mathbf{x})$  é perpendicular a este mesmo plano.

EXEMPLO 2.7. Dadas duas retas  $r_1 = \{\mathbf{p}^1 + t\mathbf{d}^1 : t \in \mathbb{R}\}$  e  $r_2 = \{\mathbf{p}^2 + t\mathbf{d}^2 : t \in \mathbb{R}\}$ , ache pontos  $\mathbf{x}^1 \in r_1$  e  $\mathbf{x}^2 \in r_2$  que têm distância mínima.

**Solução.** Seja  $\mathbf{n} = \mathbf{d}^1 \times \mathbf{d}^2$  e defina o plano  $\Sigma$  passando por  $\mathbf{p}$  perpendicular à  $\mathbf{n}$ . Então este plano  $\Sigma$  é paralelo à  $r_2$ , e basta achar a distância de  $\mathbf{p}$  até  $\Sigma$ .

### 2.6. Desigualdade lineares

Às vezes precisamos otimizar uma certa função definida no  $\mathbb{R}^n$  em domínios que satisfazem alguma restrição, por exemplo que as coordenadas sejam todas não negativas, i.e.  $x_i \geq 0$  para  $i = 1, \dots, n$ . Estes tipos de restrição são dadas por *desigualdade lineares*. Por simplicidade, ficaremos apenas no caso do plano, quando  $n = 2$ , mas o caso geral é análogo.

Em geral as desigualdades lineares são dadas na forma  $ax_1 + bx_2 + c \leq 0$ , onde  $a$ ,  $b$ ,  $c$  são números reais (para evitar trivialidades, suporemos sempre que  $a$  ou  $b$  são não nulos). Estas desigualdades determinam a região do plano  $\{(x, y) \in \mathbb{R}^2 : ax + by + c \leq 0\}$ . Tendo várias desigualdades, podemos considerar a interseção entre os domínios por elas determinados, a chamada *região admissível*. Esta interseção pode ser nula, não limitado, ou limitada. Neste último caso, a região será dada por um polígono.

EXEMPLO 2.8. Ache os pontos de  $\mathbb{R}^2$  tais que

$$\begin{aligned} 3x + 4y - 5 &\leq 0, \\ y &\leq 1, \\ x &\geq 0, \\ y &\geq 0. \end{aligned}$$

É um problema típico tentar agora minimizar uma função linear nalguma região como a dada no exemplo 2.8.

EXEMPLO 2.9. Ache o mínimo de  $p(x, y) = 2x + 3 - 5$  na região determinada no exemplo 2.8.

**Solução.** Note que as curvas de nível da função  $p$  são dadas por retas no plano. Neste caso, para achar os pontos de máximo e mínimo de  $p$ , basta procurar entre os vértices. Este é apenas um exemplo do caso geral, como enunciado no resultado a seguir.

**TEOREMA 2.6.1.** *Se uma região admissível  $D$  definida por desigualdades lineares é limitada, então máximos e mínimos de  $p(x, y) = ax + by + c$  em  $D$  ocorrem nos vértices de  $D$ .*

## 2.7. Cônicas no plano

Uma cônica no plano é o conjunto de pontos  $\{(x, y) \in \mathbb{R}^2 : ax^2 + bxy + cy^2 + dx + ey + f = 0\}$ , onde  $a, \dots, f \in \mathbb{R}$ . Pedimos ainda que  $a, b$  ou  $c$  seja diferente de zero. Uma outra forma de exigir isto é impor que  $|a| + |b| + |c| \neq 0$ . Se definirmos a forma quadrática  $Q(x, y) = ax^2 + bxy + cy^2$ , e a forma linear  $F(x, y) = dx + ey$ , temos que  $Q(x, y) + F(x, y) + f = 0$ . A seguir mostramos exemplos de cônicas em sua *forma reduzida*.

EXEMPLO 2.10.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \text{ (elipse)}, \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \text{ (hipérbole)}, \quad x^2 - dx = 0 \text{ (parábola)}.$$

No caso da elipse e da hipérbole, impomos que  $a$  e  $b$  sejam não nulos. No caso da parábola, a imposição é que  $d$  seja não zero.

Podemos ainda ter casos degenerados, como o exemplo a seguir nos mostra.

EXEMPLO 2.11. O caso da hipérbole degenerada é dado, para  $a$  e  $b$  não nulos, por  $(x^2/a^2) - (y^2/b^2) = 0$ , o que implica em  $y = \pm bx/a$ .

No caso da parábola degenerada, para  $a \neq 0$ , temos  $ax^2 - f = 0$ , e portanto  $x = \pm \sqrt{f/a}$ .

Elipses degeneradas são dadas por  $ax^2 + by^2 = 0$ , com  $a$  e  $b$  positivos. Logo  $x = y = 0$  é o único ponto da cônica.

Finalmente, temos cônicas dadas por conjuntos vazios (elipses e parábolas degeneradas), se  $ax^2 + by^2 + r^2 = 0$  e  $r \neq 0$ ,  $a \geq 0$  e  $b \geq 0$ , e  $a$  ou  $b$  são não nulos.

Como dissemos, todos os exemplo acima estão em sua forma reduzida, mas este não é a forma mais geral possível. Porém, todas as cônicas podem ser reescritas em forma reduzida após mudanças de coordenadas. Veja o exemplo abaixo.



EXEMPLO 2.12. (Boldrini) Seja a cônica dada por  $2x^2 + 2y^2 + 4xy + 4\sqrt{2}x + 12\sqrt{2}y - 8 = 0$ . Para reescrevê-la na forma reduzida, seguimos os passos abaixo.

Passo i: reescrever a cônica em forma matricial:

$$\begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} + [4\sqrt{2} \quad 12\sqrt{2}] \begin{bmatrix} x \\ y \end{bmatrix} - 8 = 0, \text{ onde } A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

Passo ii: diagonalizar a matriz  $A$ . Primeiro vemos que  $A$  tem como autovalores  $\lambda_1 = 0$  e  $\lambda_2 = 4$  e os correspondentes autovetores

$$\vec{v}^1 = \frac{\sqrt{2}}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \vec{v}^2 = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Note que se definirmos a matriz  $M = [\vec{v}^1 \quad \vec{v}^2]$ , então  $M^{-1} = M$  e

$$\begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} = M^{-1}AM = MAM.$$

Se  $(x_1, y_1)$  são as coordenadas de  $(x, y)$  na base  $\{\mathbf{v}^1, \mathbf{v}^2\}$ , i.e., se

$$\vec{x} = x_1\vec{v}^1 + y_1\vec{v}^2 = M \begin{bmatrix} x_1 \\ y_1 \end{bmatrix},$$

então

$$Q(x, y) = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \end{bmatrix} M^T A M \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

Passo iii: reescrever a parte linear em termos de  $(x_1, x_2)$ . Note que temos

$$[4\sqrt{2} \quad 12\sqrt{2}] \begin{bmatrix} x \\ y \end{bmatrix} = [4\sqrt{2} \quad 12\sqrt{2}] M \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

Passo iv: eliminar as constantes. Note que em termos de  $(x_1, x_2)$  a cônica é dada por

$$\begin{bmatrix} x_1 & y_1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + [4\sqrt{2} \quad 12\sqrt{2}] M \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} - 8 = 0.$$

Reescrevendo a expressão acima em sua forma não matricial, temos que

$$y_1^2 + 2x_1 + 4y_1 - 2 = 0.$$

Completando quadrados temos que  $(y_1^2 + 4y_1 + 4) + 2x_1 - 6 = 0$ , e portanto

$$(y_1 + 2)^2 + 2(x_1 - 3) = 0.$$

Introduzindo novas coordenadas  $y_2 = y_1 + 2$  e  $x_2 = x_1 - 3$  obtemos que

$$y_2^2 + 2x_2 = 0,$$

e a cônica é uma parábola.

As contas acima podem ser feitas em geral, como mostra o resultado abaixo.

TEOREMA 2.7.1. *Seja a cônica definida por  $ax^2 + bxy + cy^2 + dx + ey + f = 0$ , i.e.,*

$$\begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} + [d \quad e] \begin{bmatrix} x \\ y \end{bmatrix} + f = 0, \text{ onde } A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix},$$

e sejam  $\lambda_1$  e  $\lambda_2$  os autovalores de  $A$ . Então

- (1) se  $\lambda_1\lambda_2 > 0$ , então a cônica é uma elipse  
 (2) se  $\lambda_1\lambda_2 < 0$ , então a cônica é uma hipérbole  
 (3) se  $\lambda_1\lambda_2 = 0$ , então a cônica é uma parábola

COROLÁRIO 2.7.2. Como o sinal de  $\lambda_1\lambda_2$  é o mesmo de  $-(b^2 - 4ac)$ , podemos concluir que

- (1) se  $b^2 - 4ac < 0$ , então a cônica é uma elipse  
 (2) se  $b^2 - 4ac > 0$ , então a cônica é uma hipérbole  
 (3) se  $b^2 - 4ac = 0$ , então a cônica é uma parábola

## 2.8. Exercícios

EXERCÍCIO 2.1. Ache as coordenadas de um vetor  $(u_1, u_2)$  qualquer na base exemplo 2.1. Mostre que são unicamente determinados por  $u_1$  e  $u_2$ .

EXERCÍCIO 2.2. No  $\mathbb{R}^n$ , seja  $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ . Mostre que os vetores de  $\mathcal{B}$  são *linearmente independentes* se e somente se as coordenadas de todo vetor do  $\mathbb{R}^n$  são unicamente determinadas.

EXERCÍCIO 2.3. Mostre que  $|\|\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x} - \mathbf{y}\|$  (isto vale para qualquer norma) e que  $\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\mathbf{x} \cdot \mathbf{y} + \|\mathbf{y}\|^2$  (isto vale para qualquer norma que venha de produto interno) para todo  $\mathbf{x}, \mathbf{y}$  do  $\mathbb{R}^n$ .

EXERCÍCIO 2.4. Considere uma norma vinda de produto interno. Prove o Teorema de Pitágoras.

EXERCÍCIO 2.5. Mostre que a norma euclidiana, e a norma do exemplo 3.13 são de fato normas, segundo a definição 3.6.2.

EXERCÍCIO 2.6. Mostre que existe uma constante  $c \in \mathbb{R}$  tal que

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\| \leq c\|\mathbf{x}\|_\infty$$

para todo  $\mathbf{x} \in \mathbb{R}^2$ . Mostre que o mesmo vale para vetores do  $\mathbb{R}^n$ . Como é que esta constante depende de  $n$ ?

EXERCÍCIO 2.7. Mostre que o ângulo  $\theta$  entre a diagonal de um cubo e as suas arestas é tal que  $\cos \theta = 1/\sqrt{3}$ .

EXERCÍCIO 2.8. Verifique que o triângulo com vértices em  $(0, -1, )$ ,  $(-1, 1)$ ,  $(2, 0)$  é retângulo.

EXERCÍCIO 2.9. Mostre a lei dos cossenos, que diz que um triângulo com lados de tamanho  $a$ ,  $b$  e  $c$ , e com os lados de tamanho  $a$  e  $b$  determinando um ângulo  $\theta$ , obedecem à relação:

$$c^2 = a^2 + b^2 - 2ab \cos \theta$$

EXERCÍCIO 2.10. Seja  $\mathbf{y}$  vetor não nulo. Mostre que se  $\mathbf{z}$  é a projeção de  $\mathbf{x}$  em  $\mathbf{y}$ , i.e.,  $\mathbf{z} = \alpha\mathbf{y}$  e  $(\mathbf{x} - \mathbf{z}) \cdot \mathbf{y} = 0$ , então  $\alpha = \mathbf{x} \cdot \mathbf{y} / \|\mathbf{y}\|^2$  e  $\|\mathbf{z}\| = \|\mathbf{x}\| \cos \theta$ .

EXERCÍCIO 2.11. Mostre que a área do paralelograma determinado pelos vetores  $\mathbf{x}$  e  $\mathbf{y}$  é dada por  $\|\mathbf{x} \times \mathbf{y}\|$ .

EXERCÍCIO 2.12. Usando a notação do Exemplo 2.7, ache a distância entre as duas retas, com  $\mathbf{p}^1 = (0, 0, 0)$ ,  $\mathbf{d}^1 = (1, 0, 1)$ ,  $\mathbf{p}^2 = (-1, 1, 2)$  e  $\mathbf{d}^2 = (-1, -1, 0)$ .

## CAPÍTULO 3

# Álgebra Linear

1

Neste capítulo trataremos resumidamente de várias noções de álgebra linear, como operações com matrizes, matriz inversa, transposta e adjunta, resolução de sistemas lineares, determinantes, regra de Cramer, espaços vetoriais e subespaços, base e dimensão, produto interno, ortogonalidade, projeções, transformações lineares, núcleo e imagem, matriz de uma transformação linear. Autovalores e autovetores, polinômios característicos, operadores diagonalizáveis, operadores auto-adjuntos, operadores ortogonais, e formas bilineares.

### 3.1. Operações com matrizes

Denotaremos por  $\mathbb{R}^{m \times n}$  o espaço das matrizes reais com  $m$  linhas e  $n$  colunas. Se  $A \in \mathbb{R}^{m \times n}$ , Denotaremos por  $A_{i,j}$  o elemento da  $i$ -ésima linha e  $j$ -ésima coluna de  $A$ . A soma e multiplicação de matrizes é definida da forma usual, isto é, se  $A, B \in \mathbb{R}^{m \times n}$ , então  $C = A + B \in \mathbb{R}^{m \times n}$  é dada por  $C_{i,j} = A_{i,j} + B_{i,j}$ . A multiplicação para matrizes  $D \in \mathbb{R}^{m \times n}$ ,  $E \in \mathbb{R}^{n \times o}$ , é definida tal que  $C = AB \in \mathbb{R}^{m \times o}$  é dada por  $C_{i,j} = \sum_{k=1}^n A_{i,k} B_{k,j}$ .

Chamaremos de matriz identidade, e denotaremos por  $I$ , à matriz tal que  $I_{i,i} = 1$  e  $I_{i,j} = 0$  se  $i \neq j$ , para  $i, j = 1, \dots, n$ .

### 3.2. Matriz inversa, transposta e adjunta

Dada  $A \in \mathbb{R}^{n \times n}$ , se existir  $B \in \mathbb{R}^{n \times n}$  tal que  $AB = I$  e  $BA = I$ , então dizemos que  $A$  é invertível e que  $B$  é a inversa de  $A$ . Escrevemos ainda  $B = A^{-1}$ .

Uma forma de se computar a matriz inversa de  $A$ , quando esta existir, é via *matriz dos cofatores*. Seja  $\hat{A}^{i,j} \in \mathbb{R}^{(n-1) \times (n-1)}$  obtida de  $A$  “retirando” de  $A$  sua  $i$ -ésima linha e  $j$ -ésima coluna. Por exemplo, dada

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

temos que

$$\hat{A}^{1,1} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}, \quad \hat{A}^{1,2} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}, \quad \hat{A}^{1,3} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}, \quad \hat{A}^{2,1} = \begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix}, \quad \hat{A}^{2,2} = \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix}, \\ \hat{A}^{2,3} = \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix}, \quad \hat{A}^{3,1} = \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}, \quad \hat{A}^{3,2} = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}, \quad \hat{A}^{3,3} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}.$$

---

<sup>1</sup>Última Atualização: 18/06/2012

Definimos  $\Delta \in \mathbb{R}^{n \times n}$  como sendo a *matriz cofator* de  $A$ , onde  $\Delta_{i,j} = (-1)^{i+j} \det \hat{A}^{i,j}$ . No caso do exemplo acima, temos

$$\Delta = \begin{bmatrix} \det \hat{A}^{1,1} & -\det \hat{A}^{1,2} & \det \hat{A}^{1,3} \\ -\det \hat{A}^{2,1} & \det \hat{A}^{2,2} & -\det \hat{A}^{2,3} \\ \det \hat{A}^{3,1} & -\det \hat{A}^{3,2} & \det \hat{A}^{3,3} \end{bmatrix} = \begin{bmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{bmatrix}.$$

OBSERVAÇÃO. A matriz  $\Delta^T$  é também chamada de *adjunta*. É um péssimo nome, que provavelmente deriva de uma tradução infeliz do inglês *adjugate*. O termo matriz adjunta é utilizado mais comumente como sendo simplesmente a transposta de uma matriz (no caso real). Entretanto, na ANPEC, pode aparecer o termo “matrix adjunta” para denominar  $\Delta^T$ .

Após o cômputo de  $\Delta$  temos que se  $A$  for invertível, então

$$A^{-1} = \frac{\Delta^T}{\det A}.$$

No exemplo acima temos que  $\det A = 0$ , e portanto a matriz não é invertível. Na verdade temos o importante resultado que afirma que  $A$  é invertível se e somente se seu determinante é não nulo.

Note que para conferir se uma matriz é ou não inversa de outra, basta executar a multiplicação matricial e checar se resulta na matriz identidade. Por exemplo, se  $A$  e  $B$  são invertíveis, então  $(AB)^{-1} = B^{-1}A^{-1}$  pois

$$(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AA^{-1} = I, \quad (B^{-1}A^{-1})(AB) = BAA^{-1}B^{-1} = BB^{-1} = I.$$

Voltando ao caso geral, dada  $A \in \mathbb{R}^{m \times n}$  definimos a matriz transposta de  $A$  denotada por  $A'$  (ou  $A^T$ ), onde  $A'_{i,j} = A_{j,i}$ . Neste caso, as linhas se tornam colunas, e as colunas se tornam linhas. Note que se  $A \in \mathbb{R}^{m \times n}$  e  $B \in \mathbb{R}^{n \times o}$ , e além disto,  $C = AB$ , então  $C' = B'A'$  pois

$$C'_{i,j} = C_{j,i} = \sum_{k=1}^n A_{j,k} B_{k,i} = \sum_{k=1}^n B'_{i,k} A'_{k,j}.$$

No caso mais geral, considere dois espaços vetoriais  $V$  e  $W$ , que tenham produtos internos  $\langle \cdot, \cdot \rangle_V$  e  $\langle \cdot, \cdot \rangle_W$ , e seja  $T : V \rightarrow W$  operador linear. Definimos então a transposta de  $T$  como sendo  $T' : W \rightarrow V$  tal que

$$(3.2.1) \quad \langle \mathbf{v}, T'\mathbf{w} \rangle_V = \langle T\mathbf{v}, \mathbf{w} \rangle_W \quad \text{para todo } \mathbf{v} \in V, \mathbf{w} \in W.$$

Na verdade, na definição acima estamos considerando *Espaços de Hilbert*, mas isto é outra conversa. No caso  $V = W = \mathbb{R}^n$  com o produto interno usual, se tomarmos  $\mathbf{v} = \mathbf{e}_i$  e  $\mathbf{w} = \mathbf{e}_j$  em (3.2.1), temos  $[T']'_{i,j} = [T]_{j,i}$  (onde  $[T]$  é a representação matricial de  $T$  na base canônica).

### 3.3. Resolução de sistemas lineares

Seja  $A \in \mathbb{R}^{m \times n}$  e  $\vec{\mathbf{b}} \in \mathbb{R}^m$ . Queremos descobrir se existe, e neste caso, quem é,  $\vec{\mathbf{x}} \in \mathbb{R}^n$  tal que  $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$  (chamado de *sistema linear*). Este problema pode não ter solução ( $0x = 1$ ), ter solução única ( $2x = 1$ ), ou ter infinitas soluções ( $x + y = 1$ ). Note entretanto que se  $A$  for invertível, então o sistema tem solução única dada por  $\vec{\mathbf{x}} = A^{-1}\vec{\mathbf{b}}$ .

Em cálculos manuais, a melhor forma de se descobrir se um sistema tem solução é reduzindo-o a uma forma triangular superior, usando a *matriz ampliada*, como nos mostra o exemplo abaixo [4, pag.33].

Seja

$$\begin{cases} x_1 + 4x_2 + 3x_3 = 1, \\ 2x_1 + 5x_2 + 4x_3 = 4, \\ x_1 - 3x_2 - 2x_3 = 5. \end{cases}$$

Obtemos então a matriz ampliada, que reduzimos a uma forma triangular superior:

$$\begin{bmatrix} 1 & 4 & 3 & 1 \\ 2 & 5 & 4 & 4 \\ 1 & -3 & -2 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 4 & 3 & 1 \\ 0 & -3 & -2 & 2 \\ 0 & -7 & -5 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 4 & 3 & 1 \\ 0 & -3 & -2 & 2 \\ 0 & 0 & -1/3 & -2/3 \end{bmatrix}$$

Voltando a forma de equações, temos da última linha que  $x_3 = 2$ . Usando a segunda linha obtemos  $x_2 = -2$ . Finalmente, da primeira linha temos  $x_1 = 3$ .

### 3.4. Determinantes e a regra de Cramer

O determinante é uma função  $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  tal que se uma matriz  $A \in \mathbb{R}^{n \times n}$  é dada por  $A = [\vec{v}_1 \dots \vec{v}_n]$ , então  $\det(\cdot)$  é a (única) forma “multilinear alternada” definida em  $\vec{v}_1 \times \dots \times \vec{v}_n \mapsto \det(A)$  e tal que  $\det(I) = 1$ . Por multilinear quer-se dizer que é uma função linear em cada uma das colunas de  $A$ . Por alternada quer-se dizer que trocando-se duas linhas de lugar, o determinante é multiplicado por  $-1$ , ver [15]. Denotaremos o determinante de  $A$  por  $\det(A)$  ou  $|A|$ .

Uma outra forma de se definir determinantes é usando-se *permutações*. Seja  $I_n = \{1, \dots, n\}$  e  $\sigma : I_n \rightarrow I_n$  uma bijeção tal que  $\sigma(1, \dots, n) = (\sigma_1, \dots, \sigma_n) \in I_n$ . Considere  $S_n$  o conjunto de todas as permutações de  $I_n$ , e denote por  $\text{sgn}(\sigma)$  o  *sinal* ou  *assinatura* de  $\sigma \in S_n$ , i.e.,  $\text{sgn}(\sigma) = 1$  se é necessário um número par de inversões para se obter  $(\sigma_1, \dots, \sigma_n)$  de  $(1, \dots, n)$ . Analogamente  $\text{sgn}(\sigma) = -1$  se é necessário um número ímpar de inversões. Então

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1,\sigma_1} \dots a_{n,\sigma_n}.$$

Como exemplos, note que se  $n = 2$ , há duas permutações possíveis:

$$\sigma(1, 2) = (1, 2), \quad \sigma(1, 2) = (2, 1).$$

Portanto o determinante de uma matriz  $A$  que seja  $2 \times 2$  é dado por  $\det(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$ . Já uma matriz  $3 \times 3$  tem como permutações

$$\begin{aligned} \sigma(1, 2, 3) &= (1, 2, 3), & \sigma(1, 2, 3) &= (2, 1, 3), & \sigma(1, 2, 3) &= (1, 3, 2), & \sigma(1, 2, 3) &= (3, 2, 1), \\ \sigma(1, 2, 3) &= (3, 1, 2), & \sigma(1, 2, 3) &= (2, 3, 1). \end{aligned}$$

Finalmente, note que são sempre  $n!$  permutações possíveis, no caso de matrizes  $n \times n$ .

Algumas propriedades fundamentais de determinantes são dadas abaixo, supondo-se que  $A = [\vec{v}_1 \dots \vec{v}_n]$ :

- (1) Se existir alguma linha ou coluna zero, então o determinante se anula.
- (2)  $\det A = \det A^T$ , portanto propriedades que valem para linhas, valem para colunas.
- (3)  $|\vec{v}_1 \dots \alpha \vec{v}_j \dots \vec{v}_n| = \alpha |\vec{v}_1 \dots \vec{v}_n|$ .

- (4)  $|\vec{v}_1 \dots \vec{v}_i \dots \vec{v}_j \dots \vec{v}_n| = -|\vec{v}_1 \dots \vec{v}_j \dots \vec{v}_i \dots \vec{v}_n|$ .  
 (5)  $|\vec{v}_1 \dots \vec{v}_i + \vec{w} \dots \vec{v}_n| = |\vec{v}_1 \dots \vec{v}_i \dots \vec{v}_n| + |\vec{v}_1 \dots \vec{w} \dots \vec{v}_n|$   
 (6)  $\det(AB) = \det(A) \det(B)$

É importante notar que  $\det(A + B) \neq \det(A) + \det(B)$ . O contraexemplo mais simples é dado por

$$0 = \det \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \det \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \right) \neq \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \det \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = 2.$$

Uma interessante propriedade de determinantes é dada pelo desenvolvimento de Laplace (ver a notação para  $\hat{A}_{i,j}$  na página 13):

$$(3.4.1) \quad \det A = a_{1,1} \det \hat{A}_{1,1} - a_{1,2} \det \hat{A}_{1,2} + \dots \pm a_{1,N} \det \hat{A}_{N,1} \\ = a_{1,1} \Delta_{1,1} + a_{1,2} \Delta_{1,2} + \dots + a_{1,N} \Delta_{N,1} = a_{i,1} \Delta_{i,1} + a_{i,2} \Delta_{i,2} + \dots + a_{i,N} \Delta_{N,i}.$$

Considere agora o sistema linear  $A\vec{x} = \vec{b}$ , onde  $A \in \mathbb{R}^{n \times n}$  é invertível. Então

$$\vec{x} = A^{-1}\vec{b} = \frac{1}{\det A} \begin{bmatrix} \Delta_{1,1} & \dots & \Delta_{n,1} \\ \vdots & \dots & \vdots \\ \Delta_{1,n} & \dots & \Delta_{n,n} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Usando (3.4.1) temos que

$$x_1 = \frac{1}{\det A} (b_1 \Delta_{1,1} + \dots + b_n \Delta_{n,1}) = \frac{1}{\det A} \begin{vmatrix} b_1 & a_{1,2} & \dots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n,2} & \dots & a_{n,n} \end{vmatrix}$$

Analogamente,

$$x_i = \frac{1}{\det A} \begin{vmatrix} a_{1,1} & \dots & b_1 & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & b_n & \dots & a_{n,n} \end{vmatrix}.$$

para  $i = 1, \dots, n$ . Esta identidade é conhecida como *Regra de Cramer*.

EXEMPLO 3.1 (Tirado de [4]). Seja o sistema linear dado por  $A\vec{x} = \vec{b}$ , onde

$$A = \begin{bmatrix} 2 & -3 & 7 \\ 1 & 0 & 3 \\ 0 & 2 & -1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}.$$

Como  $\det A = -1$ , então, pela Regra de Cramer,

$$x_1 = \frac{1}{\det A} \begin{vmatrix} 1 & -3 & 7 \\ 5 & 0 & 3 \\ 0 & 2 & -1 \end{vmatrix} = -49, \quad x_2 = \begin{vmatrix} 2 & 1 & 7 \\ 1 & 5 & 3 \\ 0 & 0 & -1 \end{vmatrix} = 9, \quad x_3 = \begin{vmatrix} 2 & -3 & 1 \\ 1 & 0 & 5 \\ 0 & 2 & 0 \end{vmatrix} = 18.$$

### 3.5. Espaços vetoriais, subespaços, base e dimensão

O exemplo mais comum e intuitivo de espaço vetorial é o  $\mathbb{R}^n$ , ver Definição 2.0.1. Entretanto, uma definição mais geral é de grande utilidade. A menos que explicitamente mencionado, neste texto nos restringiremos a espaços vetoriais sobre o corpo dos reais.

**DEFINIÇÃO 3.5.1.** *Um espaço vetorial  $V$  sobre os reais é um conjunto cujos elementos chamamos de vetores, com duas operações binárias, soma vetorial e multiplicação por escalar tais que*

- (1)  $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ , para todo  $\mathbf{x}, \mathbf{y} \in V$
- (2)  $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{y} + (\mathbf{x} + \mathbf{z})$ , para todo  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$
- (3) Existe um elemento  $\mathbf{0} \in V$  tal que  $\mathbf{0} + \mathbf{x} = \mathbf{x}$ , para todo  $\mathbf{x} \in V$
- (4) Para todo  $\mathbf{x} \in V$ , existe um elemento  $\mathbf{y} \in V$  tal que  $\mathbf{y} + \mathbf{x} = \mathbf{0}$
- (5)  $1\mathbf{x} = \mathbf{x}$ , para todo  $\mathbf{x} \in V$
- (6)  $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$ , para todo  $\alpha, \beta \in \mathbb{R}$  e para todo  $\mathbf{x} \in V$
- (7)  $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$ , para todo  $\alpha, \beta \in \mathbb{R}$  e para todo  $\mathbf{x} \in V$
- (8)  $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$ , para todo  $\alpha \in \mathbb{R}$  e para todo  $\mathbf{x}, \mathbf{y} \in V$

Alguns resultados podem ser obtidos imediatamente:

**LEMA 3.5.2.** Seja  $V$  um espaço vetorial sobre os reais. Então temos que

- (1) O vetor zero é único
- (2) Todo elemento de  $\mathbf{x} \in V$  tem um único negativo dado por  $(-1)\mathbf{x}$
- (3)  $0\mathbf{x} = \mathbf{0}$  para todo  $\mathbf{x} \in V$
- (4)  $\alpha\mathbf{0} = \mathbf{0}$  para todo  $\alpha \in \mathbb{R}$

**DEMONSTRAÇÃO.** Demonstraremos apenas a primeira afirmativa. As demais ficam como exercícios. Para demonstrar (1), suponha que  $\mathbf{0}_1$  e  $\mathbf{0}_2$  sejam dois zeros de  $V$ . Logo

$$\mathbf{0}_1 = \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2,$$

onde usamos que a hipótese de que  $\mathbf{0}_1$  é zero e a propriedade (3) da Definição 3.5.1, seguida da propriedade (1). Na última igualdade usamos a hipótese de que  $\mathbf{0}_1$  é zero e novamente a propriedade (3) da Definição de 3.5.1.  $\square$

**EXEMPLO 3.2.** O espaço das matrizes  $m \times n$  reais denotado por  $\mathbb{R}^{m \times n}$  é um espaço vetorial com a definição usual de soma de matrizes e multiplicação por escalar.

**EXEMPLO 3.3.** O espaço  $F$  das funções de  $\mathbb{R}$  em  $\mathbb{R}$ , com as operações

$$(u + v)(x) = u(x) + v(x) \text{ para todo } x \in \mathbb{R} \text{ e todas } u, v \in F,$$

$$(\alpha u)(x) = \alpha u(x) \text{ para todo } x \in \mathbb{R}, \text{ toda } u \in F \text{ e todo } \alpha \in \mathbb{R},$$

é espaço vetorial, ver Exercício 3.1.

**3.5.1. Subespaço vetorial.** Seja  $V$  um espaço vetorial e  $W \subseteq V$ . Então dizemos que  $W$  é *subespaço vetorial de  $V$*  se  $W$  for também um espaço vetorial. Para que isto aconteça, basta que

- (1)  $W \neq \emptyset$
- (2) se  $\mathbf{u}, \mathbf{v} \in W$ , então  $\mathbf{u} + \mathbf{v} \in W$
- (3) se  $\alpha \in \mathbb{R}$  e  $\mathbf{u} \in W$ , então  $\alpha\mathbf{u} \in W$

Note se  $W$  é subespaço de  $V$ , então o vetor nulo  $\mathbf{0} \in W$  pois como  $W$  é não vazio, então existe algum  $\mathbf{u} \in W$ . Mas então  $\mathbf{0} = 0\mathbf{u} \in W$ , por causa de (3).

\*\*\*\* por figura mostrando subespaço vetorial e não subespaço vetorial \*\*\*\*

EXEMPLO 3.4. Note que  $\{(0, y, z) : y, z \in \mathbb{R}\}$  é subespaço vetorial do  $\mathbb{R}^3$ , mas que  $\{(0, y, z) \in \mathbb{R}^3 : y \geq 0, z \geq 0\}$  não o é.

EXEMPLO 3.5. O espaço das matrizes diagonais  $n \times n$  é subespaço do espaço vetorial das matrizes  $n \times n$ .

EXEMPLO 3.6. O espaço dos polinômios quadráticos é subespaço vetorial do espaço das funções.

EXEMPLO 3.7. Se  $V$  é espaço vetorial e  $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ , então

$$\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \stackrel{\text{def}}{=} \{\alpha_1\mathbf{v}_1 + \dots + \alpha_k\mathbf{v}_k : \alpha_1, \dots, \alpha_k \in \mathbb{R}\}$$

é subespaço vetorial de  $V$ . Chamamos o termo  $\alpha_1\mathbf{v}_1 + \dots + \alpha_k\mathbf{v}_k$  de *combinação linear de*  $\mathbf{v}_1, \dots, \mathbf{v}_k$ .

Os próximos resultados respondem à pergunta natural: interseções e uniões de subespaços são ainda subespaços?

LEMA 3.5.3. Sejam  $W_1$  e  $W_2$  subespaços vetoriais de um espaço vetorial  $V$ . Então  $W_1 \cap W_2$  é subespaço vetorial de  $V$ .

DEMONSTRAÇÃO. Como  $W_1$  e  $W_2$  são ambos subespaços, então  $\mathbf{0} \in W_1$  e  $\mathbf{0} \in W_2$ . Logo  $\mathbf{0} \in W_1 \cap W_2$  e então  $W_1 \cap W_2 \neq \emptyset$ .

Além disto, se  $\mathbf{u}, \mathbf{v} \in W_1 \cap W_2$ , então  $\mathbf{u}, \mathbf{v} \in W_1$  e  $\mathbf{u}, \mathbf{v} \in W_2$ . Logo, pela propriedade (2) de subespaços vetoriais,  $\mathbf{u} + \mathbf{v} \in W_1$  e  $\mathbf{u} + \mathbf{v} \in W_2$ . Mas então  $\mathbf{u} + \mathbf{v} \in W_1 \cap W_2$ .

Da mesma forma, se  $\alpha \in \mathbb{R}$  e  $\mathbf{u} \in W_1 \cap W_2$ , então, pela propriedade (3) de subespaços vetoriais,  $\alpha\mathbf{u} \in W_1$  e  $\alpha\mathbf{u} \in W_2$ . Logo  $\alpha\mathbf{u} \in W_1 \cap W_2$ .  $\square$

Como podemos ver no exemplo a seguir, a união de subespaços vetoriais *não é* subespaço vetorial.

EXEMPLO 3.8. Sejam  $A_1 = \{(x, 0) : x \in \mathbb{R}\}$  e  $A_2 = \{(0, x) : x \in \mathbb{R}\}$  subespaços de  $\mathbb{R}^2$ . Seja  $A = A_1 \cup A_2 = \{(x, y) : x = 0 \text{ ou } y = 0\}$ . Logo  $(1, 0) \in A$  e  $(0, 1) \in A$ , mas  $(1, 0) + (0, 1) = (1, 1) \notin A$ .

Apesar da união não ser necessariamente subespaços, há uma forma de se “juntar” subespaços vetoriais e obter outro subespaço, como vemos no resultado a seguir.

LEMA 3.5.4. Sejam  $W_1$  e  $W_2$  subespaços vetoriais de um espaço vetorial  $V$ . Seja o conjunto

$$W_1 + W_2 \stackrel{\text{def}}{=} \{\mathbf{w}_1 + \mathbf{w}_2 : \mathbf{w}_1 \in W_1 \text{ e } \mathbf{w}_2 \in W_2\}.$$

Então  $W_1 + W_2$  é subespaço vetorial de  $V$ .



**DEMONSTRAÇÃO.** Note que  $\mathbf{0} \in W_1 + W_2$ , logo  $W_1 + W_2 \neq \emptyset$ . Sejam agora  $\mathbf{u}, \mathbf{v} \in W_1 + W_2$ . Logo existem  $\mathbf{u}_1 \in W_1$  e  $\mathbf{u}_2 \in W_2$  tais que  $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ , pela definição de  $W_1 + W_2$ . Da mesma forma, existem  $\mathbf{v}_1 \in W_1$  e  $\mathbf{v}_2 \in W_2$  tais que  $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ . Mas então

$$\mathbf{u} + \mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{v}_1 + \mathbf{v}_2 = (\mathbf{u}_1 + \mathbf{v}_1) + (\mathbf{u}_2 + \mathbf{v}_2)$$

é a soma de um vetor de  $W_1$ , dado por  $\mathbf{u}_1 + \mathbf{v}_1$ , com outro de  $W_2$ , dado por  $\mathbf{u}_2 + \mathbf{v}_2$ . Logo  $\mathbf{u} + \mathbf{v} \in W_1 + W_2$ .

Analogamente, se  $\alpha \in \mathbb{R}$  e  $\mathbf{u} \in W_1 + W_2$ , então existem  $\mathbf{u}_1 \in W_1$  e  $\mathbf{u}_2 \in W_2$  tais que  $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ . Mas então

$$\alpha \mathbf{u} = \alpha(\mathbf{u}_1 + \mathbf{u}_2) = \alpha \mathbf{u}_1 + \alpha \mathbf{u}_2$$

é a soma de um vetor de  $W_1$  com outro de  $W_2$ . Logo  $\alpha \mathbf{u} \in W_1 + W_2$ .  $\square$

**OBSERVAÇÃO.** Algumas observações quanto a soma de espaços. A primeira é que  $W_1 + W_2$  é apenas uma notação, afinal soma de conjuntos não é uma operação que faça sentido em geral. A segunda observação diz respeito ao importante caso  $W_1 \cap W_2 = \emptyset$ . Neste caso dizemos que a soma é *direta* e a representamos por  $W_1 \oplus W_2$ . Note que podemos estender a noção de soma direta para mais que dois espaços, como em  $W_1 \oplus W_2 \oplus \cdots \oplus W_N$ . Note que  $\mathbb{R}^n = \mathbb{R} \oplus \mathbb{R} \oplus \cdots \oplus \mathbb{R}$  é a soma direta do  $\mathbb{R}$  repetida  $n$  vezes.

**3.5.2. Base e dimensão.** Sejam  $\mathbf{v}_1, \dots, \mathbf{v}_k$  vetores de um espaço vetorial  $V$ . Se

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0} \implies \alpha_1 = \cdots = \alpha_k = 0$$

então dizemos que  $\mathbf{v}_1, \dots, \mathbf{v}_k$  são *linearmente independentes*, ou *L.I.*. Vetores que não são L.I. são chamados de L.D., ou *linearmente dependentes*. Outra forma de dizer que  $\mathbf{v}_1, \dots, \mathbf{v}_k$  são L.D. é quando existirem escalares  $\alpha_1, \dots, \alpha_k$ , nem todos nulos e tais que

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0}.$$

Com o conceito de independência linear, podemos definir o que é uma base de um espaço vetorial. Dado um espaço  $V$ , dizemos que  $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$  é base de  $V$  se

- (1)  $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\} = V$
- (2)  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  é L.I.

**OBSERVAÇÃO.** Trataremos aqui sempre de espaços vetoriais de dimensão finita, e isto quer dizer que existe uma base finita para os espaços.

A seguir enunciamos alguns resultados sobre bases.

**TEOREMA 3.5.5.** *Se  $V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ , então é sempre possível extrair uma base de  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ .*

**TEOREMA 3.5.6.** *Se  $V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  então o conjunto  $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$  é L.D. sempre que  $m > n$ .*

**COROLÁRIO 3.5.7.** Qualquer base de  $V$  tem sempre o mesmo número de elementos.

O corolário acima tem grande importância pois nos diz que existe um número inerente a  $V$ , que não depende da escolha da base. A este número chamamos de *dimensão de  $V$* , ou  $\dim V$ .

**TEOREMA 3.5.8.** *Qualquer conjunto L.I. de vetores pode ser completado a fim de formar uma base.*

COROLÁRIO 3.5.9. Se  $\dim V = n$ , qualquer conjunto L.I. com  $n$  vetores é base.

TEOREMA 3.5.10. *Sejam  $W_1$  e  $W_2$  subespaços de  $V$ . Então*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Pelo resultado acima, se a soma é direta, então  $\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2$ .

TEOREMA 3.5.11. *Seja  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  base de  $V$ . Então todo vetor de  $V$  pode ser escrito de forma única como combinação linear de  $\mathbf{v}_1, \dots, \mathbf{v}_n$ .*

OBSERVAÇÃO. **Caros alunos: rever mudança de bases!**

### 3.6. Produto interno, ortogonalidade e projeções

Duas importantes *ferramentas* matemáticas quando se trabalha em espaços vetoriais são produtos internos e normas.

DEFINIÇÃO 3.6.1. *Seja  $V$  espaço vetorial sobre os reais. Um produto interno é uma função de  $V \times V \rightarrow \mathbb{R}$ , denotado por  $\mathbf{x}, \mathbf{y} \mapsto \mathbf{x} \cdot \mathbf{y}$  e tal que*

- (1)  $\mathbf{x} \cdot \mathbf{x} > 0$  para todo  $\mathbf{x} \in V$  com  $\mathbf{x} \neq \mathbf{0}$
- (2)  $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$  para todo  $\mathbf{x}, \mathbf{y} \in V$
- (3)  $(\alpha \mathbf{x}) \cdot \mathbf{y} = \alpha(\mathbf{x} \cdot \mathbf{y})$  para todo  $\alpha \in \mathbb{R}$  e todo  $\mathbf{x}, \mathbf{y} \in V$
- (4)  $(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$  para todo  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$

Outra notação usual para produtos internos é  $\langle \cdot, \cdot \rangle$ .

Note que da definição acima concluímos imediatamente que para todo  $\mathbf{x} \in V$ ,

$$\mathbf{0} \cdot \mathbf{x} = (\mathbf{00}) \cdot \mathbf{x} = 0(\mathbf{0} \cdot \mathbf{x}) = 0.$$

EXEMPLO 3.9. Em  $\mathbb{R}^2$ , se  $\mathbf{x} = (x_1, x_2)$ , e  $\mathbf{y} = (y_1, y_2)$ , o produto interno canônico é dado por

$$\mathbf{x} \cdot \mathbf{y} = \vec{\mathbf{x}}^T \vec{\mathbf{y}} = x_1 y_1 + x_2 y_2.$$

Em  $\mathbb{R}^n$ , para  $\mathbf{x} = (x_1, \dots, x_n)$ , e  $\mathbf{y} = (y_1, \dots, y_n)$ , definimos

$$\mathbf{x} \cdot \mathbf{y} = \vec{\mathbf{x}}^T \vec{\mathbf{y}} = x_1 y_1 + \dots + x_n y_n.$$

EXEMPLO 3.10. Em  $\mathbb{R}^2$ , a operação

$$(x_1, x_2) \cdot (y_1, y_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 2x_1 y_1 - x_1 y_2 - x_2 y_1 + 4x_2 y_2$$

define um produto interno. De fato, a primeira propriedade (positividade) é verdadeira pois

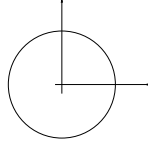
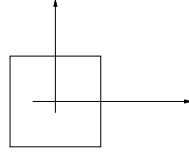
$$(x_1, x_2) \cdot (x_1, x_2) = 2x_1^2 - 2x_1 x_2 + 4x_2^2 = 2[(x_1 - x_2/2)^2 + 7x_2^2/4] > 0,$$

se  $(x_1, x_2) \neq (0, 0)$ . As outras propriedades do produto interno são mais fáceis de serem checadas.

EXEMPLO 3.11. Considere o espaço vetorial das funções contínuas em  $[0, 1]$ , com as operações de multiplicação por escalar e soma como no Exemplo 3.3. Então a operação dada pela integral de Riemann

$$f \cdot g = \int_0^1 f(x)g(x) dx$$

define um produto interno deste espaço.

FIGURA 1. Conjunto  $\{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| = 1\}$ .FIGURA 2. Conjunto  $\{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\|_\infty = 1\}$ 

Introduzimos agora a noção de *norma*. Num espaço vetorial, uma boa forma de se medir distâncias entre vetores é através de normas. Em particular, o conceito normas ajuda na definição canônica de conjuntos abertos e fechados, como veremos a seguir.

**DEFINIÇÃO 3.6.2.** *Dado um espaço vetorial  $V$ , uma norma é uma função de  $V$  em  $\mathbb{R}$ , denotada por  $\mathbf{x} \mapsto \|\mathbf{x}\|$ , e tal que*

- (1)  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$  para todo  $\mathbf{x}, \mathbf{y} \in V$  (desigualdade triangular)
- (2)  $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$  para todo  $\mathbf{x} \in V$ , e para todo  $\alpha \in \mathbb{R}$
- (3)  $\|\mathbf{x}\| > 0$  para todo  $\mathbf{x} \in V$  tal que  $\mathbf{x} \neq \mathbf{0}$

Quando um espaço vetorial  $V$  tem uma norma associada, dizemos que é um *espaço normado*.

**EXEMPLO 3.12.** Em  $\mathbb{R}^2$ ,

$$\|(x_1, x_2)\| = \sqrt{x_1^2 + x_2^2}$$

define uma norma. Na Figura 1 temos que o conjunto de pontos  $\mathbf{x}$  tais que  $\|\mathbf{x}\| = 1$  é dado por um círculo. No caso mais geral, em  $\mathbb{R}^n$ ,

$$\|(x_1, \dots, x_n)\| = \sqrt{x_1^2 + \dots + x_n^2}$$

também define uma norma.

**EXEMPLO 3.13.** Outra norma em  $\mathbb{R}^n$  é dada por

$$\|(x_1, \dots, x_n)\|_\infty = \max_{1 \leq j \leq n} |x_j|.$$

Na Figura 2 vemos que o conjunto de pontos  $\mathbf{x}$  tais que  $\|\mathbf{x}\|_\infty = 1$  é dado por um quadrado. Compare com a Figura 1.

O resultado abaixo é importante pois mostra que *todo* produto interno induz uma norma.

**TEOREMA 3.6.3.** *Seja  $V$  um espaço vetorial com produto interno. Então*

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

*define uma norma em  $V$ . Além disto, vale a desigualdade de Cauchy-Schwartz*

$$(3.6.1) \quad |\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\| \quad \text{para todo } \mathbf{x}, \mathbf{y} \in V.$$

**DEMONSTRAÇÃO.** Como o produto interno garante que sempre teremos  $\mathbf{x} \cdot \mathbf{x} \geq 0$ , então a operação acima está bem definida. Mostraremos primeiro (3.6.1). Seja  $\mathbf{z} = \mathbf{x} - (\mathbf{x} \cdot \mathbf{y})\mathbf{y}/\|\mathbf{y}\|^2$ . Então

$$\mathbf{z} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{y} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2} \mathbf{y} \cdot \mathbf{y} = 0,$$

e

$$0 \leq \|\mathbf{z}\|^2 = \mathbf{z} \cdot \mathbf{z} = \mathbf{z} \cdot \mathbf{x} = \mathbf{x} \cdot \mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2} \mathbf{x} \cdot \mathbf{y}.$$

Logo

$$(\mathbf{x} \cdot \mathbf{y})^2 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2,$$

e (3.6.1) vale.

Para mostrar a propriedade (1) da definição de norma, note que

$$\|\mathbf{x} + \mathbf{y}\|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} + 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y} \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2,$$

e assim temos (1). As propriedades (2) e (3) seguem-se imediatamente da definição e das propriedades do produto interno.  $\square$

**OBSERVAÇÃO.** Note pela demonstração acima que a igualdade  $|\mathbf{x} \cdot \mathbf{y}| = \|\mathbf{x}\|\|\mathbf{y}\|$  vale se e somente se  $\mathbf{x} = \alpha\mathbf{y}$  para algum  $\alpha \in \mathbb{R}$ . Ver exercício 3.3.

Bem como no caso do  $\mathbb{R}^n$ , ver (2.2.3), podemos definir cossenos de “ângulos entre dois vetores” não nulos  $\mathbf{x}$  e  $\mathbf{y} \in V$  por

$$(3.6.2) \quad \cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|},$$

que toma valores entre  $-1$  e  $1$  devido à desigualdade de Cauchy-Schwartz (3.6.1). Dizemos também que  $\mathbf{x}$ ,  $\mathbf{y}$  são ortogonais, ou perpendiculares, se  $\mathbf{x} \cdot \mathbf{y} = 0$ .

Outra generalização interessante é dada por projeções. Dados  $\mathbf{u}$  e  $\mathbf{v} \in V$  não nulos, chamamos  $\mathbf{w}$  de *projeção ortogonal* (ou simplesmente de *projeção*) de  $\mathbf{u}$  em  $\mathbf{v}$  se

- (1)  $\mathbf{w} = \alpha\mathbf{v}$  para algum  $\alpha \in \mathbb{R}$
- (2)  $(\mathbf{u} - \mathbf{w}) \cdot \mathbf{v} = 0$

Note que para  $\mathbf{w}$  ficar bem definido, basta calcular  $\alpha$ . Mas note que de (1) e (2), temos que  $(\mathbf{u} - \alpha\mathbf{v}) \cdot \mathbf{v} = 0$ , e portanto  $\alpha\mathbf{v} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{v}$ . Como  $\mathbf{v}$  é não nulo,

$$\alpha = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{v}\|^2} = \frac{\|\mathbf{u}\| \cos \theta}{\|\mathbf{v}\|}, \quad \mathbf{w} = \frac{\|\mathbf{u}\| \cos \theta}{\|\mathbf{v}\|} \mathbf{v}.$$

onde  $\cos \theta$  é como em (3.6.2). A norma de  $\mathbf{w}$  é dada por  $\|\mathbf{w}\| = \|\mathbf{u}\| \cos \theta$ .

### 3.7. Transformações lineares, núcleo, imagem e representações matriciais

Dados dois espaços vetoriais  $V_1$  e  $V_2$ , dizemos que uma função  $T : V_1 \rightarrow V_2$  é uma *função, transformação* ou *aplicação linear* se

$$T(\mathbf{x} + \alpha\mathbf{y}) = T(\mathbf{x}) + \alpha T(\mathbf{y}) \quad \text{para todo } \mathbf{x}, \mathbf{y} \in V_1 \text{ e todo } \alpha \in \mathbb{R}.$$

Note que em particular, para toda aplicação linear linear temos  $T(\mathbf{0}) = \mathbf{0}$ , pois

$$T(\mathbf{0}) = T(0\mathbf{0}) = 0T(\mathbf{0}) = \mathbf{0}.$$

Seja  $\mathcal{L}(V, W)$  o espaço das aplicações lineares  $T : V \rightarrow W$  para as quais existe  $M \in \mathbb{R}$  tal que

$$\|T\mathbf{x}\|_W \leq M\|\mathbf{x}\|_V,$$

Neste caso dizemos que  $T$  é limitada. Se  $V$  e  $W$  forem de dimensão finita, então toda transformação linear é limitada, ver exercício 3.5. É possível definir operações canônicas de multiplicação por escalar e soma em  $\mathcal{L}(V, W)$  de tal forma que este seja um espaço vetorial, ver exercício 3.2.

O exemplo principal de transformação linear em espaços de dimensões finitas é dado por multiplicação de matrizes. De fato, seja  $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$  definida por  $T(\vec{\mathbf{u}}) = A\vec{\mathbf{u}}$ , onde  $A \in \mathbb{R}^{n \times m}$ . Então  $T$  é linear pois  $T(\vec{\mathbf{u}} + \alpha\vec{\mathbf{v}}) = A(\vec{\mathbf{u}} + \alpha\vec{\mathbf{v}}) = A\vec{\mathbf{u}} + \alpha A\vec{\mathbf{v}} = T(\vec{\mathbf{u}}) + \alpha T(\vec{\mathbf{v}})$ .

Observe que para definir uma aplicação linear qualquer  $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ , basta defini-la numa base  $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$  do  $\mathbb{R}^m$ , i.e., basta conhecer  $T(\mathbf{v}_1), \dots, T(\mathbf{v}_m)$ . De fato, se  $\mathbf{x} = \alpha_1\mathbf{v}_1 + \dots + \alpha_m\mathbf{v}_m$ , então

$$T(\mathbf{x}) = T(\alpha_1\mathbf{v}_1 + \dots + \alpha_m\mathbf{v}_m) = \alpha_1T(\mathbf{v}_1) + \dots + \alpha_mT(\mathbf{v}_m).$$

Num certo sentido, todas as transformações lineares em espaços de dimensões finitas são dadas por matrizes. De forma mais precisa, seja  $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$  base de  $V$  e  $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$  base de  $W$ . Então, se  $\mathbf{x} \in V$  é dado por  $\mathbf{x} = \alpha_1\mathbf{v}_1 + \dots + \alpha_m\mathbf{v}_m$ , então

$$T(\mathbf{x}) = \alpha_1T(\mathbf{v}_1) + \dots + \alpha_mT(\mathbf{v}_m).$$

Seja  $A_{j,i}$  a  $j$ -ésima coordenada de  $T(\mathbf{v}_i)$  na base  $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ , i.e.,

$$(3.7.1) \quad T(\mathbf{v}_i) = A_{1,i}\mathbf{w}_1 + \dots + A_{n,i}\mathbf{w}_n.$$

Logo

$$\begin{aligned} T(\mathbf{x}) &= \alpha_1(A_{1,1}\mathbf{w}_1 + \dots + A_{n,1}\mathbf{w}_n) + \dots + \alpha_m(A_{1,m}\mathbf{w}_1 + \dots + A_{n,m}\mathbf{w}_n) \\ &= (\alpha_1A_{1,1} + \dots + \alpha_mA_{1,m})\mathbf{w}_1 + \dots + (\alpha_1A_{n,1} + \dots + \alpha_mA_{n,m})\mathbf{w}_n = \beta_1\mathbf{w}_1 + \dots + \beta_n\mathbf{w}_n, \end{aligned}$$

onde

$$(3.7.2) \quad \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} A_{1,1} & \dots & A_{1,m} \\ \vdots & \dots & \vdots \\ A_{n,1} & \dots & A_{n,m} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}.$$

Se  $\mathbf{w} = T(\mathbf{u})$ , então a matriz  $A \in \mathbb{R}^{n \times m}$  com coeficientes  $A_{i,j}$  mapeia as coordenadas de  $\mathbf{u}$  nas coordenadas de  $\mathbf{w}$ . Note que a matriz  $A$  depende fortemente das bases de  $V$  e  $W$ . Dizemos que  $A$  é a representação de ou matriz associada a  $T$  nas bases  $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$  e  $\mathcal{W} = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ . Por vezes, esta representação é também escrita como  $[T]_{\mathcal{W}}^{\mathcal{V}}$ .

Um exemplo importante é quando as bases são canônicas. Neste caso, basta ver de (3.7.1) que  $A_{j,i}$  é dada pela  $j$ -ésima coordenada de  $T(\mathbf{e}_i)$ .

Dois importantes conjuntos relacionados a uma aplicação linear  $T : V \rightarrow W$  são seu núcleo e sua imagem, dados por

$$N(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = 0\} \subseteq V, \quad \text{Im}(T) = \{T(\mathbf{v}) : \mathbf{v} \in V\} \subseteq W.$$

O núcleo recebe também a notação  $\ker(T)$ , do inglês *kernel*, e a imagem de  $V$  por  $T$  também recebe a notação  $T(V)$  ou  $R(T)$  (do inglês *range*). É importante notar que tanto o núcleo

como a imagem de uma transformação linear são espaços vetoriais. Para tal, basta checar que estes são subespaços de  $V$  e  $W$  respectivamente. Ver exercício 3.6.

Se  $\text{Im}(V) = W$  dizemos que  $T$  é sobrejetora. Se

$$T(\mathbf{u}) = T(\mathbf{v}) \implies \mathbf{u} = \mathbf{v},$$

então dizemos que  $T$  é injetora, ou 1 – 1. Temos também o seguinte resultado.

**TEOREMA 3.7.1.**  $N(T) = \mathbf{0}$  se e somente se  $T$  é injetiva.

**DEMONSTRAÇÃO.** ( $\implies$ ) Suponha que  $N(T) = \mathbf{0}$ . Sejam  $\mathbf{u}, \mathbf{v} \in V$  tais que  $T(\mathbf{u}) = T(\mathbf{v})$ . Então  $T(\mathbf{u} - \mathbf{v}) = \mathbf{0}$  e portanto  $\mathbf{u} - \mathbf{v} = \mathbf{0}$ . Logo  $T$  é injetiva.

( $\impliedby$ ) Suponha  $T$  injetora e  $T(\mathbf{u}) = \mathbf{0}$ . Então  $T(\mathbf{u}) = \mathbf{0} = T(\mathbf{0})$ . Como  $T$  é injetora, então  $\mathbf{u} = \mathbf{0}$ .  $\square$

Temos a seguir um importante resultado.

**TEOREMA 3.7.2** (Teorema do núcleo e da imagem).  $\dim N(T) + \dim \text{Im}(V) = \dim V$ .

**COROLÁRIO 3.7.3.** Seja  $\dim V = \dim W$ . Então  $T$  é injetora se e somente se  $T$  é sobrejetora.

**COROLÁRIO 3.7.4.** Se  $T$  é injetora então  $T$  leva vetores LI em vetores LI. E se  $\dim W = \dim V$ , então  $T$  leva base em base.

Note que há relação entre as dimensões do núcleo e imagem de uma transformação linear  $T$  e o posto e nulidade da matriz que representa esta transformação (em quaisquer bases):

$$\dim \text{Im}(T) = \text{posto}([T]_{\mathcal{W}}^{\mathcal{V}}), \quad \dim N(T) = \text{nulidade de } ([T]_{\mathcal{W}}^{\mathcal{V}}).$$

Note que pelo Teorema 3.7.2 que

$$\dim N(T) = \text{nulidade de } ([T]_{\mathcal{W}}^{\mathcal{V}}) = \text{número de colunas de } [T]_{\mathcal{W}}^{\mathcal{V}} - \text{posto}([T]_{\mathcal{W}}^{\mathcal{V}}).$$

### 3.8. Autovalores, polinômios característicos e operadores diagonalizáveis

Nesta seção falaremos sobre autovalores, autovetores, e suas propriedades. Por absoluta falta de espaço/tempo, não faremos contas, mas a forte recomendação para quem tem dificuldades ou não se lembra direito como se calcula autovalores, e autovetores é que olhe, por exemplo, o livro [4].

**3.8.1. Autovalores, autovetores e polinômios característicos.** Seja  $T : V \rightarrow V$  transformação linear. Dizemos que  $\lambda \in \mathbb{C}$  é um *autovalor* de  $T$  se existe vetor não nulo, chamado de *autovetor*,  $\mathbf{v} \in V$  tal que  $T\mathbf{v} = \lambda\mathbf{v}$ . Chamamos ainda  $(\lambda, \mathbf{v})$  de *autopar*.

Seja  $\lambda$  autovalor de  $T$ . Então o conjunto

$$E_{\lambda} = \{\mathbf{v} \in V : T\mathbf{v} = \lambda\mathbf{v}\}$$

é um subespaço vetorial de  $V$ , chamado *autoespaço* de  $\lambda$ . Ver Exercício 3.7.

Suponha agora um operador linear dado por uma matriz  $A \in \mathbb{R}^{n \times n}$ , definindo a aplicação linear  $\vec{\mathbf{x}} \rightarrow A\vec{\mathbf{x}}$ . Para achar autovalores e autovetores de  $A$ , basta achar soluções não triviais, i.e., não nulas, para o sistema  $(A - \lambda I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$ . Isto só será possível se  $\det(A - \lambda I) = 0$ . Note que  $\det(A - \lambda I)$  é um polinômio em termos de  $\lambda$ , ao qual damos o nome de *polinômio característico*, e denotamos por  $P(\lambda)$ . O problema de achar autovalores resume-se então

ao problema de encontrar as raízes de  $P(\lambda)$ . Isto é *sempre possível*, segundo o teorema fundamental da Álgebra, desde que admita-se autovalores complexos.

Depois de encontrado um autovalor  $\lambda$ , pode-se encontrar os autovetores correspondentes resolvendo-se  $(A - \lambda I)\vec{x} = \vec{0}$ . Note que este sistema *sempre* tem soluções não triviais, já que  $A - \lambda I$  não é invertível, ver o exercício 3.8.

Vamos agora, nos exemplos abaixo, ver o conceito de multiplicidade *algébrica* e *geométrica*.

**EXEMPLO 3.14.** Seja  $A = 2I$ . Então, se  $\lambda$  for autovalor, temos que  $\det(2I - \lambda I) = 0$ , i.e.,  $(\lambda - 2)^2 = 0$ . Logo  $\lambda = 2$  é o único autovalor. Como autovetores temos que  $(2I - \lambda I)\vec{x} = \vec{0}$ , ou seja,  $0\vec{x} = \vec{0}$ . Portanto todo  $\vec{x} \in \mathbb{R}^2$  é autovetor, e neste caso, o autoespaço tem dimensão dois. Dizemos que  $\lambda$  tem multiplicidade algébrica dois, e multiplicidade geométrica dois.

**EXEMPLO 3.15.** Seja agora a matriz

$$B = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

O polinômio característico é dado por  $P(\lambda) = (2 - \lambda)^2$ , o mesmo do exemplo 3.14, e portanto  $\lambda = 2$  é o único autovalor. Entretanto ao calcular os autovetores vemos que se  $(B - 2I)\vec{x} = \vec{0}$ , então

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Logo  $x_2 = 0$ , e os autovetores são múltiplos de  $[1, 0]^T$ . Dizemos então que  $\lambda$  tem multiplicidade algébrica dois, e multiplicidade geométrica um.

**3.8.2. Operadores diagonalizáveis.** Seja  $T : V \rightarrow V$  operador linear e  $V$  espaço vetorial de dimensão finita. Uma característica interessante de autovalores é que, quando estes formam uma base de  $V$ , a matriz que representa  $T$  é diagonal. De fato, observe em (3.7.1) que se  $\mathbf{v}_i$  é autovetor, então

$$T(\mathbf{v}_i) = \lambda \mathbf{v}_i,$$

onde tomamos  $\mathbf{w}_j = \mathbf{v}_j$  para todo  $j$ . Conclua então que a matriz  $A$  em (3.7.2) é diagonal.

Como é bastante conveniente representar um operador por uma matriz diagonal, é natural perguntar se, dado um operador linear, ele é *diagonalizável*, i.e., se existe uma base tal que sua representação nesta base é uma matriz diagonal. De forma mais simples, dizemos que um operador  $T : V \rightarrow V$  é *diagonalizável* se existe uma base de  $V$  formada por autovetores. Isto não será sempre possível, como pode ser visto no exemplo 3.15.

O resultado abaixo é importante para garantir tal base. Ele garante que autovetores correspondentes a autovetores distintos são LI.

**TEOREMA 3.8.1.** *Sejam  $\lambda_1, \dots, \lambda_k$  autovalores distintos de  $T$ . Então os correspondentes autovetores  $\mathbf{v}_1, \dots, \mathbf{v}_k$  são LI.*

**DEMONSTRAÇÃO.** Suponha que  $\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}$ . Para mostrar que estes vetores são LI, temos que mostrar que  $\alpha_k = \dots = \alpha_1 = 0$ . Aplicando  $T - \lambda_1 I$  à combinação linear acima, obtemos

$$\alpha_1(\lambda_1 - \lambda_1)\mathbf{v}_1 + \alpha_2(\lambda_2 - \lambda_1)\mathbf{v}_2 + \dots + \alpha_k(\lambda_k - \lambda_1)\mathbf{v}_k = \mathbf{0},$$

e portanto  $\alpha_2(\lambda_2 - \lambda_1)\mathbf{v}_2 + \cdots + \alpha_k(\lambda_k - \lambda_1)\mathbf{v}_k = 0$ . Aplicamos agora  $T - \lambda_2 I$  e obtemos

$$\alpha_2(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_2)\mathbf{v}_2 + \alpha_3(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)\mathbf{v}_3 + \cdots + \alpha_k(\lambda_k - \lambda_1)(\lambda_k - \lambda_2)\mathbf{v}_k = 0,$$

e obtemos que  $\alpha_3(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)\mathbf{v}_3 + \cdots + \alpha_k(\lambda_k - \lambda_1)(\lambda_k - \lambda_2)\mathbf{v}_k = 0$ . Procedendo desta forma, aplicando  $T - \lambda_3 I, \dots, T - \lambda_{k-1} I$  temos que

$$\alpha_k(\lambda_k - \lambda_1)(\lambda_k - \lambda_2) \cdots (\lambda_k - \lambda_{k-1})\mathbf{v}_k = 0.$$

Como os autovalores  $\lambda_1, \dots, \lambda_k$  são distintos entre si, obtemos que  $\alpha_k = 0$ . Voltando os passos anteriores, é possível ver que  $\alpha_{k-1} = \cdots = \alpha_1 = 0$ .  $\square$

**COROLÁRIO 3.8.2.** Se uma transformação linear em espaços de dimensão  $n$  tiver  $n$  autovalores distintos, então existe uma base formada por autovetores.

Outra forma de se definir matrizes diagonalizáveis, é exigir que sejam similares a uma matriz diagonal. Dizemos que duas matrizes  $A$  e  $B$  são similares se existe uma matriz  $P$  invertível tal que  $B = P^{-1}AP$ . O resultado abaixo trata de propriedades de matrizes similares.

**TEOREMA 3.8.3.** Matrizes similares têm o mesmo determinante, mesmo traço, mesmo polinômio característico, e mesmos autovalores.

Outro resultado interessante é que similaridade forma uma relação de equivalência, ver exercício 3.10.

Note que se a matriz  $A$  é diagonalizável e  $D = P^{-1}AP$  é diagonal, então as colunas de  $P$  são exatamente os autovetores de  $A$ . Para ver isto, suponha que

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix} = P^{-1}AP,$$

onde  $P = [\vec{v}_1 \dots \vec{v}_n]$ . Então  $PD = AP$  e  $\lambda_i \vec{v}_i = A\vec{v}_i$ .

### 3.9. Operadores auto-adjuntos, operadores ortogonais

Na seção anterior, pouco pudemos dizer a respeito de que matrizes são diagonalizáveis ou não. Em casos especiais é possível conseguir resultados mais interessantes.

Dizemos que uma transformação linear  $T$  é *auto-adjunta* ou *simétrica* se  $T' = T$ . Uma definição equivalente em espaços com produto interno  $V$  é dizer que  $\langle T\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, T\mathbf{v} \rangle$  para todo  $\mathbf{u}, \mathbf{v} \in V$ .

Caso  $T^{-1} = T'$ , dizemos que  $T$  é ortogonal. A mesma definição e terminologia é empregada no caso de matrizes.

Um exemplo de matriz ortogonal e simétrica é a identidade, e de ortogonal e não simétrica é

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix},$$

se  $\theta \neq k\pi$ .

Note que se  $A$  é ortogonal, então

$$(\det A)^2 = \det A \det A' = \det(AA') = \det I = 1,$$



e portanto  $\det A = \pm 1$ .

Segue-se imediatamente da definição de matrizes ortogonais, que os seus vetores colunas e vetores linhas são ortonormais (ou seja, são ortogonais entre si e todos têm norma um).

Outro resultado importante para operadores auto-adjuntos vem abaixo.

**TEOREMA 3.9.1.** *Seja  $T : V \rightarrow V$  auto-adjunto, e  $\lambda_1, \mathbf{v}_1$  e  $\lambda_2, \mathbf{v}_2$  autovalores e autovetores de  $T$ . Se  $\lambda_1 \neq \lambda_2$  então  $\mathbf{v}_1$  e  $\mathbf{v}_2$  são ortogonais:  $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ .*

**DEMONSTRAÇÃO.** Note que

$$\lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 = (T\mathbf{v}_1) \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot (T\mathbf{v}_2) = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2.$$

Logo  $(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ . Como  $\lambda_1 \neq \lambda_2$  então  $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ .  $\square$

Uma importante propriedade de operadores auto-adjuntos é dada pelo resultado a seguir.

**TEOREMA 3.9.2.** *Seja  $T : V \rightarrow V$  auto-adjunto. Então existe uma base ortonormal de  $V$  formada por autovalores de  $T$ .*

Note que o resultado acima *não* diz que toda base de autovalores de operadores auto-adjuntos é ortonormal, apenas que *existe* uma base ortonormal. De fato, por exemplo, para a matriz identidade  $n$ -dimensional, todo vetor de  $\mathbb{R}^n$  é autovetor. Entretanto nem toda base de  $\mathbb{R}^n$  é ortonormal.

Finalmente, o resultado abaixo serve para caracterizar quais matrizes são ortogonais.

**TEOREMA 3.9.3.** *Seja  $T : V \rightarrow V$ . Então as afirmativas abaixo são equivalentes.*

- (1)  $T$  é ortogonal
- (2)  $T$  leva bases ortogonais em bases ortogonais
- (3)  $T$  preserva produtos internos, i.e.,  $\langle T\mathbf{u}, T\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$ , para todo  $\mathbf{u}, \mathbf{v} \in V$ .
- (4)  $T$  preserva normas, i.e.,  $\|T\mathbf{u}\| = \|\mathbf{u}\|$  para todo  $\mathbf{u} \in V$ .

### 3.10. Formas lineares e bilineares

Uma forma linear definida em um espaço vetorial  $V$  é simplesmente um operador linear  $F : V \rightarrow \mathbb{R}$ . Note que se  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  é base de  $V$ , então

$$F(\mathbf{v}) = F(\alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n) = \alpha_1 F(\mathbf{v}_1) + \dots + \alpha_n F(\mathbf{v}_n) = [F(\mathbf{v}_1) \quad \dots \quad F(\mathbf{v}_n)] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix},$$

ou seja, dada uma base, a forma  $F$  num vetor  $\mathbf{v}$  pode ser caracterizada simplesmente como o produto interno do vetor formado pelos valores de  $F$  nos vetores da base com as coordenadas de  $\mathbf{v}$  na base.

Uma forma bilinear  $B$  é um operador definido em  $V \times V$  e tomando valores em  $\mathbb{R}$ , e que seja linear em cada um dos argumentos, i.e.,

$$\begin{aligned} B : V \times V &\rightarrow \mathbb{R} \\ (\mathbf{x}, \mathbf{y}) &\mapsto B(\mathbf{x}, \mathbf{y}), \end{aligned}$$

onde  $B(\mathbf{x} + \alpha \mathbf{y}, \mathbf{z}) = B(\mathbf{x}, \mathbf{z}) + \alpha B(\mathbf{y}, \mathbf{z})$  e  $B(\mathbf{x}, \mathbf{y} + \alpha \mathbf{z}) = B(\mathbf{x}, \mathbf{y}) + \alpha B(\mathbf{x}, \mathbf{z})$ , para todo  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$  e  $\alpha \in \mathbb{R}$ . Se  $B(\mathbf{x}, \mathbf{y}) = B(\mathbf{y}, \mathbf{x})$  para todo  $\mathbf{x}, \mathbf{y} \in V$ , dizemos que a forma  $B$  é simétrica.

Dada uma matriz  $A \in \mathbb{R}^{n \times n}$ , o exemplo usual de forma bilinear é definido por  $(\vec{x}, \vec{y}) \rightarrow \vec{x} \cdot A\vec{y}$ . Se  $A$  for simétrica, esta forma assim definida será simétrica.

Finalmente, assim como no caso de formas bilineares, se  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  é base de  $V$ , então

$$B(\mathbf{u}, \mathbf{v}) = [u_1 \ \dots \ u_n] \begin{bmatrix} B(\mathbf{v}_1, \mathbf{v}_1) & \cdots & B(\mathbf{v}_1, \mathbf{v}_n) \\ \vdots & \cdots & \vdots \\ B(\mathbf{v}_n, \mathbf{v}_1) & \cdots & B(\mathbf{v}_n, \mathbf{v}_n) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix},$$

onde  $\mathbf{u} = u_1\mathbf{v}_1 + \dots + u_n\mathbf{v}_n$  e  $\mathbf{v} = v_1\mathbf{v}_1 + \dots + v_n\mathbf{v}_n$ .

### 3.11. Exercícios

EXERCÍCIO 3.1. Mostre que o espaço das funções, definido no Exemplo 3.3 é de fato um espaço vetorial.

EXERCÍCIO 3.2. Defina operações de multiplicação por escalar e soma em  $\mathcal{L}(V_1, V_2)$ , tais que este seja um espaço vetorial com estas operações.

EXERCÍCIO 3.3. Dado um espaço vetorial com produto interno  $\mathbf{x} \cdot \mathbf{y}$  e norma  $\|\mathbf{x}\| = (\mathbf{x} \cdot \mathbf{x})^{1/2}$ , mostre que  $|\mathbf{x} \cdot \mathbf{y}| = \|\mathbf{x}\| \|\mathbf{y}\|$  se e somente se  $\mathbf{x} = \alpha\mathbf{y}$  para algum  $\alpha \in \mathbb{R}$ .

EXERCÍCIO 3.4. Seja  $V$  espaço vetorial normado e de dimensão finita. Seja  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  base de  $V$ , e  $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_n]^T$  o vetor formado pelas coordenadas de  $\mathbf{v} \in V$  nesta base. Mostre que existem constantes  $c_0, c_1$ , que dependem da base mas não de  $\mathbf{v}$  tais que

$$c_0 \|\boldsymbol{\alpha}\|_{\mathbb{R}^n} \leq \|\mathbf{v}\|_V \leq c_1 \|\boldsymbol{\alpha}\|_{\mathbb{R}^n}.$$

Acima,  $\|\cdot\|_{\mathbb{R}^n}$  denota a norma canônica do  $\mathbb{R}^n$ .

EXERCÍCIO 3.5. Mostre que se  $V$  e  $W$  forem espaços vetoriais de dimensão finita, então toda transformação linear  $T : V \rightarrow W$  é limitada.

EXERCÍCIO 3.6. Sejam  $V$  e  $W$  espaços vetoriais, e  $T : V \rightarrow W$  aplicação linear. Mostre que  $N(T)$  e  $\text{Im}(T)$  são subespaços vetoriais de  $V$  e  $W$  respectivamente.

EXERCÍCIO 3.7. Seja  $T : V \rightarrow W$  operador linear e  $\lambda$  autovalor de  $T$ . Mostre que  $\{\mathbf{v} \in V : T\mathbf{v} = \lambda\mathbf{v}\}$  é um subespaço vetorial de  $V$ .

EXERCÍCIO 3.8. Seja  $A \in \mathbb{R}^{n \times n}$ , e  $\lambda$  autovalor de  $A$ . Mostre que o sistema  $(A - \lambda I)\vec{x} = \vec{\mathbf{0}}$  *sempre* tem soluções não triviais.

EXERCÍCIO 3.9. Faça os detalhes da demonstração do teorema 3.8.1 no caso  $k = 3$ .

EXERCÍCIO 3.10. Mostre que similaridade forma uma relação de equivalência.

## CAPÍTULO 4

### Limites de funções

<sup>1</sup> A fim de discutirmos a noção de *continuidade de funções*, precisamos entender *limites de funções*. Este conceito será importante quando falarmos em derivação. Antes entretanto de definirmos limites de funções, daremos algumas definições básicas relacionadas aos números reais.

#### 4.1. Definições básicas envolvendo funções

Um dos conceitos mais importantes em Matemática é o de funções. Apesar de termos apresentado na página 3 uma definição de função como um caso particular de relação entre conjuntos, a descrição mais usual é dizer que uma função é uma regra que associa elementos entre dois conjuntos de uma forma específica. Para nossos propósitos entretanto, esta “definição” bastará.

Sendo mais específico, considere  $A$  e  $B$  dois conjuntos, e uma função denominada  $f$  entre  $A$  e  $B$ . Então  $f$  é uma regra que associa a cada elemento  $x \in A$ , um elemento  $f(x) \in B$ . Chamamos o conjunto  $A$  de *domínio* da função  $f$  e o denotamos por  $D(f)$ . Chamamos o conjunto  $B$  de *contradomínio* da função  $f$ . Escrevemos  $f : A \rightarrow B$ , ou ainda

$$\begin{aligned} f : A &\rightarrow B \\ x &\mapsto f(x). \end{aligned}$$

Se  $E \subseteq A$ , chamamos de *imagem de  $E$*  ao conjunto

$$f(E) = \{f(x) : x \in E\}.$$

Similarmente, se  $H \subseteq B$ , chamamos de *imagem inversa de  $H$*  o conjunto

$$f^{-1}(H) = \{x : f(x) \in H\}.$$

dizemos que  $f$  é *sobrejetiva* (ou simplesmente, *sobre*) se  $f(A) = B$ , i.e., se para todo  $y \in B$  existe  $x \in A$  tal que  $f(x) = y$ .

Dizemos que  $f$  é *injetiva* (ou *biunívoca* ou *um a um* ou 1-1) quando, para  $a, a'$  no domínio da  $f$ ,

$$f(a) = f(a') \implies a = a'.$$

Outra forma de se definir injetividade é quando

$$a \neq a' \implies f(a) \neq f(a').$$

Se  $f$  é injetiva e sobre, a chamamos de *bijetiva* ou de uma *bijeção*.

Dizemos que  $g : B \rightarrow A$  é *função inversa* de  $f$  se

$$g(f(x)) = x \quad \text{para todo } x \in A, \quad f(g(y)) = y \quad \text{para todo } y \in B.$$

---

<sup>1</sup>Última Atualização: 02/07/2012

Quando esta existir, denotamos a inversa de  $f$  por  $f^{-1}$ . Note que uma função tem inversa se e somente se é sobrejetiva e injetiva, ver problema 4.1. Portanto, se  $f : A \rightarrow B$  é injetiva, então restringindo o contradomínio  $B$  à imagem de  $A$  por  $f$ , teremos  $f$  sobrejetiva. Em outras palavras, seja  $f : A \rightarrow f(A)$ . Então  $f$  é claramente sobre. Se for também injetiva, haverá a inversa  $f^{-1} : f(A) \rightarrow A$ .

EXEMPLO 4.1. Seja  $f : (0, 1) \rightarrow \mathbb{R}$  dada por  $f(x) = x^2 + 1$ . Então  $f$  não tem inversa pois, apesar de ser injetiva, não é sobrejetiva. Mas como  $f((0, 1)) = (1, 2)$ , então se definirmos

$$\begin{aligned} f : (0, 1) &\rightarrow (1, 2) \\ x &\mapsto x^2, \end{aligned}$$

então teremos  $f$  sobrejetiva. Logo existe a inversa  $f^{-1} : (1, 2) \rightarrow (0, 1)$ .

OBSERVAÇÃO. Note que a definição de *imagem inversa* independe de existir ou não a função inversa. Por exemplo, a função  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $f(x) = x^2$  não tem inversa. Entretanto  $f^{-1}(\mathbb{R}) = \mathbb{R}$ .

Existem várias operações com funções, entre elas a composição. Sejam  $A, B$  e  $C$  conjuntos, e  $f : A \rightarrow B$  e  $g : B \rightarrow C$  funções. Então podemos definir uma função  $h : A \rightarrow C$  dada pela *composição* de  $f$  e  $g$ , i.e.,  $h(x) \stackrel{\text{def}}{=} g(f(x))$ , onde  $x \in A$ . Neste caso usamos a notação  $h = g \circ f$ , e dizemos que  $h$  é a composta da  $f$  com a  $g$ .

Outra operação que pode ser muitas vezes executada é soma, diferença, produto, divisão, de funções. Por exemplo, sejam  $A$  conjunto e  $f : A \rightarrow \mathbb{R}$  e  $g : A \rightarrow \mathbb{R}$  funções. Então podemos definir a função  $h = f + g$  tal que

$$h(x) = f(x) + g(x) \text{ para todo } x \in A.$$

Da mesma forma podemos definir  $fg$  por  $(fg)(x) = f(x)g(x)$ . Outras operações são definidas analogamente.

Temos que tomar cuidado entretanto se a definição faz sentido. Por exemplo, se  $f : A \rightarrow \mathbb{R}$  se anula em algum ponto de  $A$ , então não podemos definir  $\phi : A \rightarrow \mathbb{R}$  por  $\phi(x) = 1/f(x)$ . Igualmente, não faz sentido definir  $1/g$ , se  $g : A \rightarrow \mathbb{R}^n$ , para  $n > 1$ , pois não podemos dividir por vetores.

OBSERVAÇÃO. A visualização de funções é possível, via gráficos. Por exemplo, o gráfico de uma equação que depende de  $(x, y)$  é simplesmente o conjunto de pontos do plano que satisfazem esta equação. Por exemplo, o gráfico de uma função  $f : \mathbb{R} \rightarrow \mathbb{R}$  é dado pelo conjunto  $\{(x, y) \in \mathbb{R}^2 : y = f(x)\}$ . Já o gráfico de  $x^2 + y^2 = 3$  é formado por todos os pontos de  $\mathbb{R}^2$  que têm norma igual a  $\sqrt{3}$ .

## 4.2. Intervalos na reta

Neste capítulo, falaremos sobre intervalos na reta. Falaremos também sobre vizinhanças, cuja noção é baseada na função valor absoluto (que nada mais é que uma norma nos reais).

Intervalos na reta serão conjuntos como os abaixo:

- (1) Intervalo aberto:  $(a, b) = \{x \in \mathbb{R} : a < x < b\}$
- (2) Intervalo fechado:  $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$
- (3)  $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$

- (4)  $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$
- (5)  $[a, +\infty) = \{x \in \mathbb{R} : a \leq x\}$
- (6)  $(a, +\infty) = \{x \in \mathbb{R} : a < x\}$
- (7)  $(-\infty, b] = \{x \in \mathbb{R} : x \leq b\}$
- (8)  $(-\infty, b) = \{x \in \mathbb{R} : x < b\}$
- (9)  $(-\infty, +\infty) = \mathbb{R}$

Prosseguimos no sentido de definirmos vizinhanças. Para tal precisamos da noção de distância entre dois pontos  $x$  e  $y$  da reta, que é dada pelo valor absoluto de  $x - y$ , i.e., por  $|x - y|$ . Para um número real  $a$ , o valor absoluto (ou módulo) de  $a$  é dado por

$$|a| = \begin{cases} a & \text{se } a \geq 0, \\ -a & \text{se } a < 0. \end{cases}$$

EXEMPLO 4.2. Por definição  $|5| = 5$ , e  $|-5| = -(-5) = 5$ .

LEMA 4.2.1. Algumas propriedades dos números reais:

- (1)  $|-a| = |a|$  para todo  $a \in \mathbb{R}$ .
- (2)  $|ab| = |a||b|$  para todo  $a, b \in \mathbb{R}$ .
- (3) Dados  $a, k \in \mathbb{R}$  temos que  $|a| \leq k$  se e somente se  $-k \leq a \leq k$ .
- (4)  $-|a| \leq a \leq |a|$  para todo  $a \in \mathbb{R}$ .

DEMONSTRAÇÃO. (1) Se  $a = 0$ , então  $|0| = 0 = |-0|$ . Se  $a > 0$ , então  $-a < 0$  e logo  $|-a| = -(-a) = a = |a|$ . Se  $a < 0$ , então  $-a > 0$  e  $|-a| = -a = |a|$ .

(2) Exercício.

(3) Exercício.

(4) Tome  $k = |a|$  no item (3) do lema. Então  $|a| \leq |a| \implies -|a| \leq a \leq |a|$ . □

LEMA 4.2.2 (Desigualdade Triangular). Para todo  $a, b \in \mathbb{R}$  temos

$$|a + b| \leq |a| + |b|.$$

DEMONSTRAÇÃO. Sabemos que  $-|a| \leq a \leq |a|$  e  $-|b| \leq b \leq |b|$ . Logo,  $-|a| - |b| \leq a + b \leq |a| + |b|$ . Pelo item (3) do Lema 4.2.1 temos que  $|a + b| \leq |a| + |b|$ , como queríamos demonstrar. □

Seja  $a \in \mathbb{R}$  e considere o intervalo

$$B_\epsilon(a) = \{x \in \mathbb{R} : |x - a| < \epsilon\} = (a - \epsilon, a + \epsilon).$$

Uma vizinhança de  $a$  é qualquer conjunto contendo  $B_\epsilon(a)$  para algum  $\epsilon > 0$ .

### 4.3. Funções inversas

Investigaremos mais de perto agora quando uma função é invertível. Em particular nos concentraremos nas funções monótonas, i.e., funções que são crescentes ou decrescentes, que definiremos a seguir.

Seja  $I \subset \mathbb{R}$ . Diremos que uma função  $f : I \rightarrow \mathbb{R}$  é crescente se dados dois pontos  $x, y$  de  $I$  tais que  $x < y$  tem-se que  $f(x) < f(y)$ . Se  $x < y$  implica em  $f(x) \leq f(y)$ , diremos que  $f$  é não decrescente.

Analogamente,  $g : I \rightarrow \mathbb{R}$  é decrescente se  $x < y$  implica em  $f(x) > f(y)$ , e não crescente se  $x < y$  implica em  $f(x) \geq f(y)$ . Se uma função é crescente ou decrescente, a chamamos de *monótona*.

**OBSERVAÇÃO.** A terminologia acima não é unânime. Alguns autores preferem chamar funções crescentes como definidas acima como estritamente crescentes, funções não-decrescentes como crescentes.

Note que toda função monótona é injetiva. A volta vale também, mas sob a condição da função ser contínua. Isto é, toda função contínua definida num intervalo e injetiva é monótona, ver Exercício 4.3.

**TEOREMA 4.3.1.** *Seja  $I$  um intervalo e  $f : I \rightarrow \mathbb{R}$  contínua. Então*

- (1) *se  $f$  for monótona, então  $f(I)$  é intervalo, e a função inversa  $f^{-1} : f(I) \rightarrow \mathbb{R}$  é contínua*
- (2) *se  $f$  for injetiva então é monótona.*

#### 4.4. Limites de funções

Seja  $I = (a, b)$  um intervalo não vazio e  $f : I \rightarrow \mathbb{R}$ , e seja  $c \in [a, b]$ . Dizemos que  $L$  é o limite de  $f$  em  $c$  se para todo  $\epsilon > 0$  existir  $\delta > 0$  tal que

$$x \in (c - \delta, c + \delta) \cap I, \quad x \neq c \implies f(x) \in (L - \epsilon, L + \epsilon).$$

Neste caso, escrevemos  $L = \lim_{x \rightarrow c} f(x)$ , e dizemos que  $f$  converge para  $L$  no ponto  $c$ . Outra forma de escrever a definição acima é dizendo que para todo  $\epsilon > 0$ , existe  $\delta > 0$  tal que

$$x \in I, \quad 0 < |x - c| < \delta \implies |f(x) - L| < \epsilon.$$

Uma observação a respeito da definição acima é que o valor do limite em  $c$  independe do valor que  $f$  assume em  $c$ . Na verdade,  $f$  não precisa nem estar definida neste ponto. Somente quando discutirmos continuidade é que o valor em  $c$  será importante, mas isto fica para o próximo capítulo.

Antes de começarmos a calcular limites, é interessante também ver que as seguintes definições são equivalentes, e qualquer uma delas pode ser usada no estudo de limites.

**LEMA 4.4.1** (Critérios equivalentes para limites). Seja  $I = (a, b)$  e  $f : I \rightarrow \mathbb{R}$ , e seja  $c \in [a, b]$ . Então as afirmativas são equivalentes:

- (1)  $\lim_{x \rightarrow c} f(x) = L$ .
- (2) Seja  $(x_n)$  sequência em  $I$  com  $x_n \neq c$  para todo  $n$  e  $\lim_{n \rightarrow \infty} x_n = c$ . Então  $(f(x_n))$  converge e  $\lim_{n \rightarrow \infty} (f(x_n)) = L$ .

**DEMONSTRAÇÃO.** (1)  $\implies$  (2) Seja  $\epsilon > 0$ , e  $(x_n)$  em  $I \setminus \{c\}$  tal que  $\lim_{n \rightarrow \infty} x_n = c$ . Por hipótese existe  $\delta$  tal que

$$(4.4.1) \quad x \in I, \quad 0 < |x - c| < \delta \implies |f(x) - L| < \epsilon.$$

Seja  $N \in \mathbb{N}$  tal que  $|x_n - c| < \delta$  se  $n \geq N$ . Então, por (4.4.1) tem-se  $|f(x_n) - L| < \epsilon$  e conclui-se que a sequência  $(f(x_n))$  converge para  $L$ .

(2)  $\implies$  (1) (*por contradição*) Assuma que (2) valha, e que (1) seja falso. Logo existe vizinhança  $\delta > 0$  tal que para todo  $n \in \mathbb{N}$  existe  $x_n \in (c - 1/n, c + 1/n) \cap I$ , com  $x_n \neq c$  e  $f(x_n) \notin (L - \delta, L + \delta)$ . Isto é uma contradição pois por (2) teríamos que  $\lim_{n \rightarrow \infty} (f(x_n)) = L$ .  $\square$

EXEMPLO 4.3. Seja  $f(x) = x$ . Então  $\lim_{x \rightarrow c} f(x) = c$ . De fato, dado  $\epsilon > 0$ , tome  $\delta = \epsilon$ , pois

$$0 < |x - c| < \delta \implies |f(x) - f(c)| = |x - c| < \delta = \epsilon.$$

EXEMPLO 4.4. Seja  $f : \mathbb{R} \rightarrow \mathbb{R}$  definida por

$$f(x) = \begin{cases} x, & \text{se } x \in \mathbb{Q} \\ 0, & \text{se } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Então  $f$  tem limite bem definido em  $c = 0$ , mas não nos demais pontos. Dado  $\epsilon > 0$ , seja  $\delta = \epsilon$ . Se  $|x| < \delta$ , então  $|f(x)| = 0 \leq |x| < \delta = \epsilon$  caso  $x \in \mathbb{R} \setminus \mathbb{Q}$ , e  $|f(x)| = |x| < \delta = \epsilon$  caso  $x \in \mathbb{Q}$ . Logo  $|x - 0| < \delta$  implica em  $|f(x) - f(0)| = |f(x)| < \epsilon$ . Portanto  $f$  tem limite no zero.

Nos demais pontos tal limite não existe pela densidade dos racionais nos irracionais e vice-versa. De fato, dado  $x \in \mathbb{R}$ , existe  $(x_n)$  sequência em  $\mathbb{R} \setminus \mathbb{Q}$  e  $(y_n)$  sequência em  $\mathbb{Q}$ , ambas convergentes para  $x$  com  $x_n \neq x$  e  $y_n \neq x$  para todo  $n \in \mathbb{N}$ . Mas  $\lim_{n \rightarrow \infty} f(x_n) = 0$  e  $\lim_{n \rightarrow \infty} f(y_n) = \lim_{n \rightarrow \infty} (y_n) = x \neq 0$ . Portanto  $f$  não tem limite para  $x \neq 0$ .

EXEMPLO 4.5. Ache  $\lim_{x \rightarrow 0} f(x)$  de  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$  dada por

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{se } x \in \mathbb{Q} \cap \mathbb{R}^+, \\ 0 & \text{se } x \in \mathbb{R}^+ \setminus \mathbb{Q}. \end{cases}$$

Primeiro note para  $x \in \mathbb{R}^+$  que se  $x \in \mathbb{Q}$ , então  $|f(x)| \leq |x|$  pois  $|\sin 1/x| \leq 1$ . Se  $x \notin \mathbb{Q}$ , então  $|f(x)| = 0 \leq |x|$ . Em ambos os casos temos  $|f(x)| \leq |x|$ . Então, dado  $\epsilon > 0$ , seja  $\delta = \epsilon$ . Então se  $x \in \mathbb{R}^+$  e

$$0 < |x| < \delta = \epsilon \implies |f(x) - 0| \leq |x| < \epsilon.$$

Logo  $f$  tem limite no zero e o limite é zero, i.e.,  $\lim_{x \rightarrow 0} f(x) = 0$

LEMA 4.4.2 (Unicidade do limite). Seja  $I = (a, b)$  e  $f : I \rightarrow \mathbb{R}$ , e seja  $c \in [a, b]$ . Então  $f$  pode ter, no máximo, um limite em  $c$ .

DEMONSTRAÇÃO. Sejam  $L_1$  e  $L_2$  dois limites de  $f$  em  $c$ . Portanto, dado  $\epsilon > 0$  existem  $\delta_1$  e  $\delta_2$  tais que

$$\begin{aligned} x \in A, \quad 0 < |x - c| < \delta_1 &\implies |f(x) - L_1| < \frac{\epsilon}{2}, \\ x \in A, \quad 0 < |x - c| < \delta_2 &\implies |f(x) - L_2| < \frac{\epsilon}{2}. \end{aligned}$$

Tome  $\delta = \min\{\delta_1, \delta_2\}$ . Então tomando-se  $0 < |x - c| < \delta$  implica em

$$|L_1 - L_2| \leq |L_1 - f(x)| + |f(x) - L_2| < \epsilon.$$

Como  $\epsilon$  é arbitrário, temos  $L_1 = L_2$ . □

EXEMPLO 4.6. Seja

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

Tomando-se as seqüências  $(-1/n)$  e  $(1/n)$ , ambas convergindo para  $c = 0$  mas nunca atingindo este valor, tem-se  $(f(-1/n)) = -1$  e  $(f(1/n)) = 1$ . Então esta função não tem limite em  $c = 0$ , pois se o limite existe, este tem que ser único.

Assim como no caso de seqüências, podemos definir operações com funções, como soma, subtração, etc. Se  $f : I \rightarrow \mathbb{R}$  e  $g : I \rightarrow \mathbb{R}$ , então definimos  $(f + g) : I \rightarrow \mathbb{R}$  por  $(f + g)(x) = f(x) + g(x)$ . De forma análoga definimos  $(f - g)(x) = f(x) - g(x)$  e  $(fg)(x) = f(x)g(x)$ . Se  $g$  é tal que  $g(x) \neq 0$  para todo  $x \in I$ , definimos também  $(f/g)(x) = f(x)/g(x)$ . Valem então resultados como o limite da soma é a soma do limite, etc.

LEMA 4.4.3. Seja  $I = (a, b)$ . Sejam  $f : I \rightarrow \mathbb{R}$  e  $g : I \rightarrow \mathbb{R}$ , e seja  $c \in [a, b]$ . Suponha que existam os limites  $\lim_{x \rightarrow c} f(x)$  e  $\lim_{x \rightarrow c} g(x)$ . Então

- (1)  $\lim_{x \rightarrow c} (f + g)(x) = \lim_{x \rightarrow c} f(x) + \lim_{x \rightarrow c} g(x)$
- (2)  $\lim_{x \rightarrow c} (f - g)(x) = \lim_{x \rightarrow c} f(x) - \lim_{x \rightarrow c} g(x)$
- (3)  $\lim_{x \rightarrow c} (fg)(x) = \lim_{x \rightarrow c} f(x) \lim_{x \rightarrow c} g(x)$
- (4)  $\lim_{x \rightarrow c} (f/g)(x) = \lim_{x \rightarrow c} f(x) / \lim_{x \rightarrow c} g(x)$ , se  $g$  for tal que  $g(x) \neq 0$  para todo  $x \in I$ , e  $\lim_{x \rightarrow c} g(x) \neq 0$ .

Os resultados acima podem ser estendidos para um número finito de operações.

DEMONSTRAÇÃO. Demonstraremos apenas (1). As demais demonstrações são similares.

Seja  $(x_n)$  seqüência em  $I$  com  $x_n \neq c$  para todo  $n \in \mathbb{N}$  e  $\lim_{n \rightarrow \infty} x_n = c$ . Então  $(f + g)(x_n) = f(x_n) + g(x_n)$  converge pois é soma de seqüências convergentes e  $\lim_{n \rightarrow \infty} ((f + g)(x_n)) = \lim_{n \rightarrow \infty} (f(x_n)) + \lim_{n \rightarrow \infty} (g(x_n))$ .  $\square$

EXEMPLO 4.7. Seja  $n \in \mathbb{N}$ . Então  $\lim_{x \rightarrow c} x^n = (\lim_{x \rightarrow c} x)^n = c^n$ .

EXEMPLO 4.8. Se  $c > 0$ , então  $\lim_{x \rightarrow c} 1/x = 1/(\lim_{x \rightarrow c} x) = 1/c$ .

Uma condição extra tem que ser imposta quando lidamos com composição de funções. É natural perguntar, supondo-se que  $\lim_{x \rightarrow c} g(x) = L$ , quando

$$\lim_{x \rightarrow c} f(g(x)) = f(\lim_{x \rightarrow c} g(x)) = f(L)$$

ocorre. E a resposta é que a igualdade acima é verdadeira se  $\lim_{y \rightarrow L} f(y) = f(L)$ . Em outras palavras, basta que  $f$  seja contínua em  $L$ .

Uma outra propriedade de funções que têm limite em um ponto é a de limitação local, i.e., a função é limitada numa vizinhança do ponto. Observe que uma função localmente limitada não necessariamente é globalmente limitada, como veremos mais a seguir.

DEFINIÇÃO 4.4.4. Seja  $I = (a, b)$  e  $f : I \rightarrow \mathbb{R}$  e  $c \in [a, b]$ . Dizemos que  $f$  é limitada numa vizinhança de  $c$  se existem  $\delta > 0$  e constante  $M$  tais que

$$x \in (c - \delta, c + \delta) \cap I \implies |f(x)| \leq M.$$

Dizemos também que  $f$  é localmente limitada em  $c$ .

Para mostrar que se  $f : I \rightarrow \mathbb{R}$  tem limite em  $c$  de  $I$  então  $f$  é localmente limitada em  $c$ , basta primeiro tomar  $\epsilon = 1$ . Dado  $L = \lim_{x \rightarrow c} f(x)$ , existe  $\delta > 0$  tal que

$$x \in I, \quad 0 < |x - c| < \delta \implies |f(x) - L| < 1.$$



Neste caso, temos  $|f(x)| \leq |f(x) - L| + |L| < 1 + L$ . Se  $c \notin I$  defina  $M = 1 + |L|$ . Se  $c \in I$  defina  $M = \max\{|f(c)|, 1 + |L|\}$ . Em qualquer dos casos temos que

$$x \in I, \quad |x - c| < \delta \implies |f(x)| < M.$$

Da discussão acima concluímos imediatamente que  $f(x) = 1/x$  é localmente limitada em todo ponto  $c \neq 0$ . Além disso concluímos que  $f$  não tem limite em  $c = 0$  pois não é limitada localmente em torno deste ponto.

Alguns resultados que valem para seqüências podem ser estendidos para limites de funções. Por exemplo, do Lema 4.4.5 tiramos o seguinte resultado. Sua demonstração é um exercício.

LEMA 4.4.5 (limite de sanduíches). Sejam  $I = (a, b)$  e  $f, g$  e  $h$  funções de  $I$  em  $\mathbb{R}$ , e seja  $c \in [a, b]$ . Suponha que para todo  $x \in I$  com  $x \neq c$  tivermos  $f(x) \leq g(x) \leq h(x)$ , e que  $\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} h(x) = L$ . Então  $\lim_{x \rightarrow c} g(x) = L$ .

LEMA 4.4.6. Sejam  $I = (a, b)$  e  $f : I \rightarrow \mathbb{R}$ , e seja  $c \in [a, b]$ . Suponha que para todo  $x \in I$  com  $x \neq c$  tivermos  $a \leq g(x) \leq b$ , e que existe o limite de  $f$  em  $c$ . Então  $a \leq \lim_{x \rightarrow c} f(x) \leq b$ .

#### 4.5. Limites laterais, infinitos e no infinito

Assim como na seção anterior, assumimos que  $I = (a, b) \subset \mathbb{R}$  e  $f : I \rightarrow \mathbb{R}$ . Seja agora  $c \in [a, b]$ . Dizemos que  $L$  é limite à direita de  $f$  em  $c$  se para todo  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$x \in I, \quad 0 < x - c < \delta \implies |f(x) - L| < \epsilon.$$

Neste caso escrevemos que  $\lim_{x \rightarrow c^+} f(x) = L$ .

Definição similar vale para limite à esquerda (e escrevemos  $\lim_{x \rightarrow c^-} f(x) = L$ ).

É possível mostrar que se  $c$  ponto de acumulação tanto de  $I \cap (c, +\infty)$  como de  $I \cap (-\infty, c)$ , então

$$(4.5.1) \quad \lim_{x \rightarrow c} f(x) = L \iff \lim_{x \rightarrow c^+} f(x) = \lim_{x \rightarrow c^-} f(x) = L.$$

EXEMPLO 4.9. Seja  $f(x) = \text{sgn}(x)$ , como no exemplo 4.6. Como  $\lim_{x \rightarrow 0^+} f(x) = 1$  e  $\lim_{x \rightarrow 0^-} f(x) = -1$ , então não existe limite de  $f$  no zero.

Outra definição importante é a de *limite infinito*. Dizemos que  $f$  tende a  $+\infty$  em  $c$  se para todo  $\alpha \in \mathbb{R}$  existe  $\delta > 0$  tal que

$$x \in I, \quad 0 < |x - c| < \delta \implies f(x) > \alpha.$$

Escrevemos então que  $\lim_{x \rightarrow c} = +\infty$ .

Definição similar vale para  $f$  tende a  $-\infty$  em  $c$ .

EXEMPLO 4.10.  $\lim_{x \rightarrow 0} 1/x^2 = +\infty$ . De fato, dado  $\alpha > 0$ , tomando  $\delta = 1/\sqrt{\alpha}$  temos

$$0 < |x| < \delta \implies x^2 < \delta^2 = \frac{1}{\alpha} \implies \frac{1}{x^2} > \alpha.$$

EXEMPLO 4.11. Seja  $g : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ . Então  $g$  não tende a  $-\infty$  ou a  $+\infty$  no zero pois  $g(x) < 0$  se  $x < 0$  e  $g(x) > 0$  se  $x > 0$ .

Finalmente definimos limites “no infinito”. Seja  $a \in \mathbb{R}$  e  $f : (a, +\infty) \rightarrow \mathbb{R}$ . Dizemos que  $L \in \mathbb{R}$  é limite de  $f$  quando  $x \rightarrow +\infty$  se para todo  $\epsilon$  existe  $k > a$  tal que

$$x > k \implies |f(x) - L| < \epsilon.$$

Analogamente podemos definir limite de  $f$  quando  $x \rightarrow -\infty$ .

EXEMPLO 4.12.  $\lim_{x \rightarrow -\infty} 1/x = \lim_{x \rightarrow +\infty} 1/x = 0$ .

EXEMPLO 4.13. Nem sempre existe limite “no infinito”. Tome por exemplo  $\sin(x)$ .

#### 4.6. Exercícios

EXERCÍCIO 4.1. Mostre que uma função tem inversa se e somente se é sobrejetiva e injetiva.

EXERCÍCIO 4.2. Demonstre os itens (2) e (3) no Lema 4.2.1.

EXERCÍCIO 4.3. Construa  $f : (0, 1) \rightarrow \mathbb{R}$  injetiva e não monótona.

EXERCÍCIO 4.4. Mostre que se  $x \neq y$  são números reais, então existem vizinhanças  $U$  de  $x$  e  $V$  de  $y$  tais que  $U \cap V = \emptyset$ .

EXERCÍCIO 4.5. Demonstre o Lema 4.4.5.

EXERCÍCIO 4.6. Demonstre o Lema 4.4.6.

EXERCÍCIO 4.7. Demonstre a equivalência 4.5.1.

## CAPÍTULO 5

### Continuidade e Funções Contínuas

<sup>1</sup> A partir das definições de limites de funções do capítulo anterior, fica mais fácil definir continuidade e estudar suas propriedades.

#### 5.1. Introdução e exemplos

Seja  $A \subset \mathbb{R}$  e  $f : A \rightarrow \mathbb{R}$ . Dizemos que  $f$  é *contínua em*  $c \in A$  se para todo  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$x \in (c - \epsilon, c + \epsilon) \cap A \implies f(x) \in (c - \delta, c + \delta).$$

Finalmente, dizemos que  $f$  é contínua em  $B \subset A$  se  $f$  for contínua em todos os pontos de  $B$ .

OBSERVAÇÃO. Note que, quando  $\lim_{x \rightarrow c} f(x)$  está bem definido,

$$f \text{ é contínua em } c \iff f(c) = \lim_{x \rightarrow c} f(x).$$

OBSERVAÇÃO. Note uma diferença na definição de limite de função e continuidade num ponto  $c$ . Para definir limite, a função não precisava nem estar definida em  $c$ , e se estivesse, o valor de  $f(c)$  não tinha importância.

LEMA 5.1.1. Seja  $A \subset \mathbb{R}$  e  $f : A \rightarrow \mathbb{R}$ . Seja  $c \in A$ . Então as afirmativas abaixo são equivalentes.

- (1)  $f$  é contínua em  $c$ .
- (2) Para todo  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$x \in A, \quad |x - c| < \delta \implies |f(x) - f(c)| < \epsilon.$$

- (3) Se  $(x_n)$  é tal que  $x_n \in A$  para todo  $n \in \mathbb{N}$  e  $\lim_{n \rightarrow \infty} x_n = c$ , então  $\lim_{n \rightarrow \infty} f(x_n) = f(c)$ .

Outro resultado útil é o seguinte *critério de descontinuidade*: assumindo as hipóteses do Lema 5.1.1, temos que  $f$  não é contínua em  $c$  se e somente se existe sequência  $(x_n)$  em  $A$  convergindo para  $c$  mas  $(f(x_n))$  não convergindo para  $f(c)$ .

EXEMPLO 5.1.  $g(x) = x$  é contínua em  $\mathbb{R}$ . De fato, para todo  $c \in \mathbb{R}$ , temos  $\lim_{x \rightarrow c} g(x) = c = g(c)$ .

EXEMPLO 5.2. A função  $\text{sgn}(x)$  (ver exemplo 4.6) não é contínua no zero, já que não existe  $\lim_{x \rightarrow 0} \text{sgn}(x)$ .

---

<sup>1</sup>Última Atualização: 23/06/2012

EXEMPLO 5.3. Seja  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por

$$f(x) = \begin{cases} 1 & \text{se } x \in \mathbb{Q}, \\ 0 & \text{caso contrário,} \end{cases}$$

é descontínua para todo  $x \in \mathbb{R}$ . Para mostrar isto, assumamos  $x \in \mathbb{Q}$ , e uma sequência  $(x_n)$  em  $\mathbb{R} \setminus \mathbb{Q}$  convergindo para  $x$ . Neste caso,  $\lim_{n \rightarrow \infty} (f(x_n)) = 0 \neq 1 = f(x)$ . Da mesma forma, se  $x \notin \mathbb{Q}$ , tomamos uma sequência  $(x_n)$  em  $\mathbb{Q}$  convergindo para  $x$ , e temos  $\lim_{n \rightarrow \infty} (f(x_n)) = 1 \neq 0 = f(x)$ .

As vezes, é possível estender uma função de forma contínua para pontos “fora” do domínio original. Por exemplo, seja  $I = (a, c)$  e  $f : I \rightarrow \mathbb{R}$ . Se existir  $\lim_{x \rightarrow c} f(x)$ , então definiremos  $f(c)$  como sendo este limite, e  $f$  será contínua em  $c$ .

EXEMPLO 5.4. Considere a função similar ao problema 4.4, mas desta vez definida apenas para reais positivos:

$$f : \mathbb{R}^+ \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} x, & \text{se } x \in \mathbb{R}^+ \cap \mathbb{Q}, \\ 0, & \text{se } x \in \mathbb{R}^+ \setminus \mathbb{Q}. \end{cases}$$

Então  $\lim_{x \rightarrow 0} f(x) = 0$  e podemos estender  $f$  continuamente no zero definindo

$$g : \mathbb{R}^+ \cup \{0\} \rightarrow \mathbb{R}, \quad g(x) = \begin{cases} f(x), & \text{se } x \in \mathbb{R}^+, \\ 0, & \text{se } x = 0. \end{cases}$$

Então temos  $g$  contínua no zero (e somente no zero).

EXEMPLO 5.5. É claro que nem sempre tal extensão contínua é possível. Por exemplo no caso de  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$  dada por  $f(x) = 1/x$ , não se pode definir  $f(0)$  tal que  $f : \mathbb{R}^+ \cup \{0\} \rightarrow \mathbb{R}$  seja contínua.

**5.1.1. Composição de funções.** Em geral, se  $f$  e  $g$  são contínuas, então  $f + g$ ,  $f - g$ ,  $fg$  também o são. Da mesma forma, se  $h(x) \neq 0$  para todo  $x$  do domínio, então  $f/h$  é contínua. O próximo resultado garante que a *composição* de funções contínuas também é contínua.

TEOREMA 5.1.2. *Sejam  $A, B \subset \mathbb{R}$ , e  $f : A \rightarrow B$  e  $g : B \rightarrow \mathbb{R}$ . Assuma  $f$  contínua em  $c \in A$  e  $g$  contínua em  $f(c) \in B$ . Então a composição  $g \circ f : A \rightarrow \mathbb{R}$  é contínua em  $c$ .*

EXEMPLO 5.6. A função  $g(x) = |x|$  é contínua em  $\mathbb{R}$ . Realmente, como

$$|g(x) - g(y)| = ||x| - |y|| \leq |x - y|,$$

se  $(x_n)$  converge para  $x$  então

$$|g(x_n) - g(x)| \leq |x_n - x| \implies \lim_{n \rightarrow \infty} (g(x_n)) = g(x).$$

Portanto, se  $f : A \rightarrow \mathbb{R}$  é contínua em  $c \in A$ , então  $h(x) = |f(x)|$  também o é, pois  $h = g \circ f$  é composição de funções contínuas.

## 5.2. Funções Contínuas em intervalos fechados e limitados

Um resultado com várias aplicações diz que funções contínuas definidas em conjuntos fechados e limitados são limitadas e atingem seus pontos extremos. Chamamos um intervalo de fechado limitado quando é da forma  $[a, b]$ , para  $a < b$ . Na verdade, todos os resultados abaixo, excetuando-se o Teorema 5.2.5, valem em conjuntos mais gerais, por exemplo em uniões finitas de intervalos fechados e limitados.

**DEFINIÇÃO 5.2.1.** Dizemos que  $f : A \rightarrow \mathbb{R}$  é limitada em  $A$  se existe  $M \in \mathbb{R}$  tal que  $|f(x)| \leq M$  para todo  $x \in A$ .

**EXEMPLO 5.7.**  $\sin x$  é limitada em  $\mathbb{R}$  pois  $|\sin x| \leq 1$  para todo  $x \in \mathbb{R}$ .

**EXEMPLO 5.8.**  $1/x$  não é limitada em  $\mathbb{R}^+$ . Entretanto  $1/x$  é limitada em  $(1/2, +\infty)$  pois  $|1/x| \leq 2$  para todo  $x$  neste intervalo.

**TEOREMA 5.2.2.** Seja  $I = [a, b]$ , e  $f : I \rightarrow \mathbb{R}$  contínua em  $I$ . Então  $f$  é limitada em  $I$ .

Outra noção importante é o de máximos e mínimos. Dizemos que  $f : A \rightarrow \mathbb{R}$  tem valor máximo em  $A$  se existe  $x^* \in A$  tal que  $f(x^*)$  é cota superior de  $f(A)$ . De forma análoga dizemos que  $f$  tem valor mínimo em  $A$  se existe  $x_* \in A$  tal que  $f(x_*)$  é cota inferior de  $f(A)$ . Chamamos  $x^*$  de ponto de valor máximo e  $x_*$  de ponto de valor mínimo.

**OBSERVAÇÃO.** Se uma função  $f$  como acima definida assume seus valores máximo e mínimo em  $A$ , então  $f$  é limitada em  $A$ .

**EXEMPLO 5.9.**  $f : (-1, 1) \rightarrow \mathbb{R}$  dada por  $f(x) = 1/(1 - x^2)$  não é limitada em  $(-1, 1)$ , mas é limitada em  $[-1/2, 1/2]$  por exemplo.

**EXEMPLO 5.10.**  $f(x) = x$  é contínua e limitada em  $(-1, 1)$ , mas não assume valor máximo nem mínimo em  $(-1, 1)$ . Entretanto  $f$  assume seus valores máximo e mínimo em  $[-1, 1]$ .

**EXEMPLO 5.11.**  $h(x) = 1/(1 + x^2)$  é limitada em  $\mathbb{R}$ , assume seu valor máximo em  $x^* = 0$ , mas não assume seu valor mínimo. Isto porque  $\inf h(\mathbb{R}) = 0 \neq h(x)$  para todo  $x \in \mathbb{R}$ .

**OBSERVAÇÃO.** Note que pontos de máximo e mínimo não são únicos em geral. Por exemplo,  $f(x) = x^2$  tem  $-1$  e  $1$  como seus dois pontos de máximo em  $[-1, 1]$ .

**TEOREMA 5.2.3 (Pontos Extremos).** Seja  $I = [a, b]$ , e  $f : I \rightarrow \mathbb{R}$  contínua em  $I$ . Então  $f$  tem pelo menos um ponto de máximo e um de mínimo em  $I$ .

Outro resultado de grande importância é o Teorema do Valor Intermediário que garante a preservação de intervalos por funções contínuas.

**TEOREMA 5.2.4 (Teorema do Valor Intermediário).** Sejam  $a < b$  e  $f : [a, b] \rightarrow \mathbb{R}$  contínua. Se existe  $d \in \mathbb{R}$  tal que  $f(a) < d < f(b)$ , então existe  $c \in (a, b)$  tal que  $f(c) = d$ .

Concluimos esta parte com uma importante consequência dos resultados anteriores.

**TEOREMA 5.2.5.** Seja  $I$  intervalo fechado limitado e  $f : I \rightarrow \mathbb{R}$  função contínua. Então  $f(I)$  é intervalo fechado limitado.

## 5.3. Exercícios

**EXERCÍCIO 5.1.** Determine os pontos de continuidade da função  $[x]$ , que retorna para cada  $x \in \mathbb{R}$  o maior inteiro menor ou igual a  $x$ . Por exemplo,  $[2] = 2$ ,  $[2.5] = 2$ ,  $[-2.5] = -3$ .



## CAPÍTULO 6

### Diferenciação

<sup>1</sup> Neste capítulo vemos a noção de diferenciabilidade e suas aplicações.

#### 6.1. Definições e Exemplos

Seja  $f : I \rightarrow \mathbb{R}$ , onde  $I$  é um intervalo em  $\mathbb{R}$ . Dizemos que  $f$  é diferenciável em  $c \in I$  se existe um número real  $L$  onde dado  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$x \in I, \quad 0 < |x - c| < \delta \implies \left| \frac{f(x) - f(c)}{x - c} - L \right| < \epsilon.$$

Chamamos  $L$  de derivada de  $f$  em  $c$ , e escrevemos  $L = f'(c)$ .

Note que se  $f$  é diferenciável em  $c$ , então

$$f'(c) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c}.$$

Se  $f$  é diferenciável em todo ponto de  $I$  dizemos que  $f$  é diferenciável em  $I$ . Neste caso note que a derivada  $f'$  é uma função de  $I$  em  $\mathbb{R}$ .

EXEMPLO 6.1. Se  $f(x) = x^2$ , então para  $c \in \mathbb{R}$  tem-se

$$f'(c) = \lim_{x \rightarrow c} \frac{x^2 - c^2}{x - c} = \lim_{x \rightarrow c} \frac{(x + c)(x - c)}{x - c} = \lim_{x \rightarrow c} (x + c) = 2c.$$

TEOREMA 6.1.1. Se  $f : I \rightarrow \mathbb{R}$ , onde  $I$  é um intervalo em  $\mathbb{R}$  é diferenciável em  $c \in I$ , então  $f$  é contínua em  $c$ .

DEMONSTRAÇÃO. Seja  $L = f'(c)$ . Dado  $\epsilon > 0$ , existe  $\delta > 0$  tal que

$$x \in I, \quad 0 < |x - c| < \delta \implies L - \epsilon < \left| \frac{f(x) - f(c)}{x - c} \right| < L + \epsilon.$$

Seja  $\bar{\delta} = \min\{\delta, \epsilon/(L + \epsilon)\}$ . Então

$$x \in I, \quad 0 < |x - c| < \bar{\delta} \implies |f(x) - f(c)| = \left| \frac{f(x) - f(c)}{x - c} \right| |x - c| \leq (L + \epsilon)\bar{\delta} \leq \epsilon.$$

Logo  $f$  é contínua em  $c$ . □

OBSERVAÇÃO. Pelo teorema acima, diferenciabilidade implica em continuidade. O inverso entretanto não é verdade em geral. Seja por exemplo  $f : \mathbb{R} \rightarrow \mathbb{R}$  onde  $f(x) = |x|$ . Então  $f$  é contínua em  $\mathbb{R}$  mas não é diferenciável em zero pois para  $x \neq 0$  temos

$$\left| \frac{f(x) - f(0)}{x - 0} \right| = \frac{|x|}{x} = \begin{cases} 1 & \text{se } x > 0, \\ -1 & \text{se } x < 0. \end{cases}$$

---

<sup>1</sup>Última Atualização: 26/06/2012

Logo o limite quando  $x \rightarrow 0$  não existe.

## 6.2. Propriedades da Derivada

Seja  $f$  e  $g$  funções de  $I \rightarrow \mathbb{R}$ , onde  $I$  é um intervalo em  $\mathbb{R}$ , ambas diferenciáveis em  $c \in I$ . Então

(1)  $(\alpha f)'(c) = \alpha f'(c)$ , onde  $\alpha \in \mathbb{R}$ . De fato, se  $x \neq c$ , então

$$\frac{(\alpha f)(x) - (\alpha f)(c)}{x - c} = \alpha \frac{f(x) - f(c)}{x - c}.$$

(2)  $(f + g)'(c) = f'(c) + g'(c)$ .

(3) Se  $p = fg$ , então se  $x \neq c$ ,

$$\begin{aligned} \frac{p(x) - p(c)}{x - c} &= \frac{f(x)g(x) - f(c)g(c)}{x - c} = \frac{f(x)g(x) - f(c)g(x) + f(c)g(x) - f(c)g(c)}{x - c} \\ &= \frac{f(x) - f(c)}{x - c} g(x) + f(c) \frac{g(x) - g(c)}{x - c}. \end{aligned}$$

Logo existe  $\lim_{x \rightarrow c} (p(x) - p(c))/(x - c)$  e

$$\begin{aligned} p'(c) &= \lim_{x \rightarrow c} \frac{p(x) - p(c)}{x - c} = \lim_{x \rightarrow c} \left[ \frac{f(x) - f(c)}{x - c} g(x) \right] + \lim_{x \rightarrow c} \left[ f(c) \frac{g(x) - g(c)}{x - c} \right] \\ &= f'(c)g(c) + f(c)g'(c). \end{aligned}$$

(4) Se  $g(x) \neq 0$  para todo  $x \in I$ , então seja  $h(x) = f(x)/g(x)$ . Logo se  $x \neq c$ ,

$$\begin{aligned} \frac{h(x) - h(c)}{x - c} &= \frac{\frac{f(x)}{g(x)} - \frac{f(c)}{g(c)}}{x - c} = \frac{f(x)g(c) - f(c)g(x)}{(x - c)g(x)g(c)} \\ &= \frac{f(x)g(c) - f(c)g(c)}{(x - c)g(x)g(c)} + \frac{f(c)g(c) - f(c)g(x)}{(x - c)g(x)g(c)} = \frac{f(x) - f(c)}{x - c} \frac{1}{g(x)} - \frac{f(c)}{g(x)g(c)} \frac{g(x) - g(c)}{x - c}. \end{aligned}$$

Logo existe  $\lim_{x \rightarrow c} (h(x) - h(c))/(x - c)$  e

$$h'(c) = \lim_{x \rightarrow c} \frac{h(x) - h(c)}{x - c} = f'(c) \frac{1}{g'(c)} - \frac{f(c)}{g^2(c)} g'(c).$$

**EXEMPLO 6.2.** Pela regra acima temos que se  $f(x) = x^n$ , para  $n \in \mathbb{N}$ , então  $f$  é diferenciável e  $f'(c) = nx^{n-1}$ .

Observe que  $f : I \rightarrow \mathbb{R}$  é diferenciável em  $c \in I$  com  $f'(c) = L$  se e somente se existir uma função  $r$  tal que

$$f(x) = f(c) + (x - c)L + r(x - c), \quad \text{com } \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0.$$

De forma equivalente escrevemos  $h = x - c$  e

$$f(c + h) = f(c) + hL + r(h) \quad \text{com } \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0.$$



**TEOREMA 6.2.1** (Regra da Cadeia). *Sejam  $I$  e  $J$  intervalos em  $\mathbb{R}$  e  $g : I \rightarrow \mathbb{R}$  e  $f : J \rightarrow \mathbb{R}$ , onde  $f(J) \subset I$ . Se  $f$  é diferenciável em  $c \in J$  e  $g$  é diferenciável em  $f(c)$ , então  $g \circ f$  é diferenciável em  $c$  e*

$$(g \circ f)'(c) = g'(f(c))f'(c).$$

**DEMONSTRAÇÃO.** Seja  $d = f(c)$ . Note que para  $h$  tal que  $c+h \in J$  e  $k$  tal que  $d+k \in I$ , temos

$$f(c+h) = f(c) + hf'(c) + r(h) \quad \text{com} \quad \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0.$$

$$g(d+k) = g(d) + kg'(d) + p(k) \quad \text{com} \quad \lim_{k \rightarrow 0} \frac{p(k)}{k} = 0.$$

Definindo  $k = f(c+h) - f(c) = hf'(c) + r(h)$ , temos

$$\begin{aligned} g \circ f(c+h) &= g(f(c+h)) = g(d+k) = g(d) + kg'(d) + p(k) \\ &= g(d) + (hf'(c) + r(h))g'(d) + p(f(c+h) - f(c)) = g(d) + hf'(c)g'(d) + q(h) \end{aligned}$$

onde  $q(h) = r(h)g'(d) + p(f(c+h) - f(c))$ . Finalmente,

$$\lim_{h \rightarrow 0} \frac{q(h)}{h} = g'(d) \lim_{h \rightarrow 0} \frac{r(h)}{h} + \lim_{h \rightarrow 0} \frac{p(f(c+h) - f(c))}{h}.$$

Se  $f(c+h) = f(c)$  numa vizinhança de  $c$ , então  $p(f(c+h) - f(c)) = 0$ . Caso contrário,

$$\lim_{h \rightarrow 0} \frac{p(f(c+h) - f(c))}{h} = \lim_{h \rightarrow 0} \frac{p(f(c+h) - f(c))}{f(c+h) - f(c)} \lim_{h \rightarrow 0} \frac{f(c+h) - f(c)}{h} = 0.$$

De qualquer forma concluímos que

$$\lim_{h \rightarrow 0} \frac{p(f(c+h) - f(c))}{h} = 0.$$

□

**EXEMPLO 6.3.** Seja

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x}, & \text{se } x \neq 0 \\ 0, & \text{se } x = 0. \end{cases}$$

Logo, para  $x \neq 0$  temos  $f'(x) = 2x \sin 1/x - \cos 1/x$ . Em  $x = 0$  usamos a definição:

$$f'(0) = \lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \rightarrow 0} x \sin \frac{1}{x} = 0.$$

Logo  $f$  é diferenciável em  $\mathbb{R}$  mas  $f'$  não é contínua no zero.

**TEOREMA 6.2.2** (Derivada da Função Inversa). *Seja  $I \subset \mathbb{R}$  intervalo,  $f : I \rightarrow \mathbb{R}$  contínua e invertível com inversa  $g : J \rightarrow \mathbb{R}$  contínua, e  $J = f(I)$ . Se  $f$  é diferenciável em  $c \in I$ , então  $g$  é diferenciável em  $d = f(c)$  se e somente se  $f'(c) \neq 0$ . Neste caso,*

$$g'(d) = \frac{1}{f'(c)} = \frac{1}{f'(g(d))}$$

**DEMONSTRAÇÃO.** Tendo que  $g$  é contínua. Al'ém disso, se  $y \in J \setminus \{d\}$ , então  $g(y) \neq c$ . Logo, se  $f'(c) \neq 0$ ,

$$\lim_{y \rightarrow d} \frac{g(y) - g(d)}{y - d} = \lim_{y \rightarrow d} \frac{g(y) - c}{f(g(y)) - f(c)} = \lim_{y \rightarrow d} \left( \frac{f(g(y)) - f(c)}{g(y) - c} \right)^{-1} = \frac{1}{f'(c)}.$$

Logo  $g$  é diferenciável em  $d$  e  $g'(d) = 1/f'(c)$ . Analogamente, se  $g$  é diferenciável em  $d$ , então usando a regra da cadeia e que  $g(f(x)) = x$ , temos

$$g'(f(c))f'(c) = 1,$$

e então  $f'(c) \neq 0$ . □

**EXEMPLO 6.4.** Seja  $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  dada por  $f(x) = x^n$ , onde  $n \in \mathbb{N}$ . Então  $f$  tem inversa  $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , e  $g(y) = \sqrt[n]{y}$ . Para  $y > 0$  temos então

$$g'(y) = \frac{1}{ny^{\frac{n-1}{n}}}.$$

Note que  $g$  não é diferenciável no zero pois  $f'(0) = 0$ .

### 6.3. Aplicações

Uma primeira e importante aplicação diz respeito a pontos extremos locais. Dizemos que uma função  $f : I \rightarrow \mathbb{R}$ , onde  $I \subset \mathbb{R}$  é um intervalo, tem um *máximo local* em  $c \in I$  se existe  $\delta > 0$  tal que

$$x \in (c - \delta, c + \delta) \cap I \implies f(x) \leq f(c).$$

Definição análoga serve para mínimo local. Chamamos um ponto de máximo ou mínimo local de ponto extremo local.

O resultado a seguir descreve condição necessária para um ponto ser extremo local.

**TEOREMA 6.3.1** (Ponto extremo interior). *Seja  $f : I \rightarrow \mathbb{R}$ , onde  $I \subset \mathbb{R}$  é um intervalo, e  $c \in I$  ponto extremo local. Se  $f$  é diferenciável em  $c$ , então  $f'(c) = 0$ .*

**DEMONSTRAÇÃO.** Sem perda de generalidade, assuma  $c$  ponto de máximo local. Então, se  $f'(c) > 0$  temos

$$0 < f'(c) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c} \implies \frac{f(x) - f(c)}{x - c} > 0$$

numa vizinhança de  $c$ . Logo, para  $x > c$  tem-se  $f(x) > f(c)$ , contradição pois  $c$  é ponto de máximo local. De forma semelhante não podemos ter  $f'(c) < 0$ . Logo  $f'(c) = 0$ . □

A seguir apresentamos um resultado com importantes por si e por suas consequências. É o *Teorema do Valor Médio*, que vemos a seguir na sua versão mais simples, o *Teorema de Rolle*.

**TEOREMA 6.3.2** (Teorema de Rolle). *Seja  $a < b \in \mathbb{R}$  e  $f : [a, b] \rightarrow \mathbb{R}$  contínua e diferenciável em  $[a, b]$ . Assuma ainda que  $f(a) = f(b) = 0$ . Então existe  $c \in (a, b)$  tal que  $f'(c) = 0$ .*

**DEMONSTRAÇÃO.** Se  $f$  é identicamente nula em  $[a, b]$ , então o resultado é verdadeiro. Caso contrário, então  $f$  assume algum valor positivo ou negativo em  $(a, b)$ . Sem perda de generalidade, suponha que  $f$  assumo algum valor positivo. Como  $[a, b]$  é intervalo fechado e limitado, então  $f$  atinge seu máximo em algum  $c \in (a, b)$ . Mas pelo Teorema do Ponto extremo interior 6.3.1,  $f'(c) = 0$ , como queríamos demonstrar.  $\square$

**TEOREMA 6.3.3** (Teorema do Valor Médio). *Seja  $a < b \in \mathbb{R}$  e  $f : [a, b] \rightarrow \mathbb{R}$  continua e diferenciável em  $[a, b]$ . Então existe  $c \in (a, b)$  tal que*

$$f(b) - f(a) = f'(c)(b - a).$$

**DEMONSTRAÇÃO.** Seja

$$\phi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Então  $\phi(a) = \phi(b) = 0$ . Como  $f$  é diferenciável em  $[a, b]$ , então  $\phi$  também o é no mesmo intervalo. Logo, pelo Teorema de Rolle 6.3.2 existe  $c \in (a, b)$  tal que  $\phi'(c) = 0$ . Portanto

$$f'(x) = \frac{f(b) - f(a)}{b - a}.$$

$\square$

Uma primeira aplicação do Teorema do Valor Médio garante que se uma função definida num intervalo tem derivada identicamente igual a zero, então a função é constante.

**LEMA 6.3.4.** Assuma que  $f : [a, b] \rightarrow \mathbb{R}$  seja contínua em  $[a, b]$ , onde  $a < b$ , e diferenciável em  $(a, b)$ . Se  $f'(x) = 0$  para todo  $x$ , então  $f$  é constante em  $[a, b]$ .

**DEMONSTRAÇÃO.** Seja  $a < x < b$ . Pelo Teorema do Valor Médio 6.3.3, existe  $c \in (a, x)$  tal que  $f(x) - f(a) = f'(c)(x - a)$ . Como  $f'(c) = 0$ , temos  $f(x) = f(a)$ . Como  $x$  é arbitrário, temos  $f$  constante em  $(a, b)$ . Mas continuidade temos  $f$  constante em  $[a, b]$ .  $\square$

Observe que pelo resultado acima, se  $f, g$  são funções diferenciáveis que tem a mesma derivada, então  $f$  e  $g$  diferem por uma constante.

A aplicação seguinte do Teorema do Valor Médio garante condições necessárias e suficientes para uma função ser crescente num intervalo.

**LEMA 6.3.5.** Seja  $I \subset \mathbb{R}$  intervalo e  $f : I \rightarrow \mathbb{R}$  diferenciável em  $I$ . Então

- (1)  $f$  é crescente em  $I$  se e somente se  $f'(x) \geq 0$  para todo  $x \in I$ .
- (2)  $f$  é decrescente em  $I$  se e somente se  $f'(x) \leq 0$  para todo  $x \in I$ .

**DEMONSTRAÇÃO.** Assuma  $f$  crescente.

( $\implies$ ) Para  $x, c \in I$ ,

$$x < c \text{ ou } x > c \implies \frac{f(x) - f(c)}{x - c} \geq 0.$$

Portanto

$$f'(c) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c} \geq 0.$$

( $\implies$ ) Assuma  $f'(x) \geq 0$  para todo  $x \in I$ . Sejam  $x_1 < x_2$  com  $x_1, x_2 \in I$ . Usando o teorema do valor médio 6.3.3, existe  $c \in (x_1, x_2)$ .  $\square$

**OBSERVAÇÃO.** É possível modificar a demonstração acima e mostrar que  $f'(x) > 0$  implica em  $f$  estritamente crescente. Entretanto, mesmo funções que tem derivada nula em alguns pontos podem ser estritamente crescentes, como por exemplo  $f(x) = x^3$ .

**OBSERVAÇÃO.** Não é verdade que se  $f'(c) > 0$  para algum ponto  $c$  no domínio da  $f$  implique em  $f$  crescente numa vizinhança de  $c$ . Como exemplo considere

$$g(x) = \begin{cases} x + 2x^2 \sin \frac{1}{x} & \text{se } x \neq 0, \\ 0 & \text{se } x = 0, \end{cases}$$

é diferenciável em zero com  $g'(0) = 1$ , mas não é crescente e, nenhuma vizinhança do zero.

Outra aplicação do Teorema do Valor Médio segue no exemplo abaixo.

**EXEMPLO 6.5.** Seja  $f(x) = \exp(x)$ . Então  $f'(x) = \exp(x)$ . Queremos mostrar que

$$(6.3.1) \quad \exp(x) > 1 + x \text{ para todo } x \neq 0.$$

Seja  $x > 0$ . Então aplicando o Teorema do Valor Médio em  $[0, x]$  temos que existe  $c \in (0, x)$  tal que

$$\exp(x) - \exp(0) = \exp(c)(x - 0).$$

Como  $c > 0$ , então  $\exp(c) > \exp(0) = 1$ . Logo

$$\exp(x) > 1 + x.$$

Para  $x < 0$ , os argumentos são semelhantes e portanto a desigualdade (6.3.1) vale.

## 6.4. Teorema de Taylor e Aplicações

Uma ferramenta poderosa em análise com várias consequências é o Teorema de Taylor, que é na verdade também uma aplicação do Teorema do Valor Médio.

A expansão de Taylor aproxima localmente uma função que pode ser complicada por um polinômio. Suponha que  $f : I \rightarrow \mathbb{R}$  onde  $I \subset \mathbb{R}$  tenha  $n \geq 0$  derivadas num ponto  $x_0 \in I$ . Defina

$$P_n(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0) \frac{(x - x_0)^2}{2} + \cdots + f^{(n)}(x_0) \frac{(x - x_0)^n}{n!},$$

onde usamos a notação que  $g^{(k)}(c)$  indica a  $k$ -ésima deriva de  $g$  num ponto  $c$ .

Note que com a definição acima, temos  $f^{(k)}(x_0) = P_n^{(k)}(x_0)$  para  $k = 1, \dots, n$ . Chamamos  $P_n$  de polinômio de Taylor de ordem  $n$  para  $f$  em  $x_0$ , e o resultado abaixo diz o quão boa é a aproximação de uma função por seu polinômio de Taylor.

**TEOREMA 6.4.1 (Taylor).** *Seja  $n \geq 0$  e  $I = [a, b]$ , com  $a < b$ . Seja  $f : I \rightarrow \mathbb{R}$   $n$  vezes diferenciável em  $I$  com  $f^{(n)}$  contínua em  $I$  e tal  $f^{(n+1)}$  exista em  $(a, b)$ . Se  $x_0, x \in I$  então existe  $\xi \in (x_0, x) \cap (x, x_0)$  tal que*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0) \frac{(x - x_0)^2}{2} + \cdots + f^{(n)}(x_0) \frac{(x - x_0)^n}{n!} + f^{(n+1)}(\xi) \frac{(x - x_0)^{n+1}}{(n+1)!}.$$

**DEMONSTRAÇÃO.** Sejam  $x_0, x \in I$ . Sem perda de generalidade, assumamos  $x > x_0$ . Defina  $J = [x_0, x]$  e seja  $F : J \rightarrow \mathbb{R}$  dada por

$$F(t) = f(x) - f(t) - (x-t)f'(t) - \dots - \frac{(x-t)^n}{n!}f^{(n)}(t).$$

Logo

$$F'(t) = -\frac{(x-t)^n}{n!}f^{(n+1)}(t)$$

Definindo  $G : J \rightarrow \mathbb{R}$  por

$$G(t) = F(t) - \left(\frac{x-t}{x-x_0}\right)^{n+1} F(x_0),$$

temos  $G(x_0) = G(x) = 0$ . Pelo Teorema de Rolle 6.3.2 existe  $\xi \in (x_0, x)$  tal que

$$0 = G'(\xi) = F'(\xi) + (n+1)\frac{(x-\xi)^n}{(x-x_0)^{n+1}}F(x_0).$$

Portanto

$$\begin{aligned} F(x_0) &= -\frac{1}{n+1}\frac{(x-x_0)^{n+1}}{(x-\xi)^n}F'(\xi) = \frac{1}{n+1}\frac{(x-x_0)^{n+1}}{(x-\xi)^n}\frac{(x-\xi)^n}{n!}f^{(n+1)}(\xi) \\ &= \frac{(x-x_0)^{n+1}}{(n+1)!}f^{(n+1)}(\xi). \end{aligned}$$

□

**EXEMPLO 6.6.** Seja  $f : I \rightarrow \mathbb{R}$ , onde  $I = [a, b] \subset \mathbb{R}$ , com  $a < b$ . Assumamos que  $f$  e suas derivadas  $f', f'', \dots, f^{(n+1)}$  existam e sejam contínuas em  $I$ . Se  $f^{(n+1)}(x) = 0$  para todo  $x \in I$  e  $f(x_0) = f'(x_0) = \dots = f^{(n)}(x_0) = 0$  para algum  $x_0 \in I$ , então  $f(x) = 0$  para todo  $x \in I$ . De fato, pelo Teorema de Taylor 6.4.1, dado  $x \in I$ , existe  $\xi$  entre  $x$  e  $x_0$  tal que

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x-x_0) + f''(x_0)\frac{(x-x_0)^2}{2} + \dots + f^{(n)}(x_0)\frac{(x-x_0)^n}{n!} \\ &\quad + f^{(n+1)}(\xi)\frac{(x-x_0)^{n+1}}{(n+1)!}. \end{aligned}$$

Mas por hipótese,  $f^{(i)}(x_0) = 0$  para  $i = 0, \dots, n$ , e  $f^{(n+1)} \equiv 0$  em  $I$ . Em particular, como  $\xi \in I$ , temos  $f^{(n+1)}(\xi) = 0$ . Portanto,  $f(x) = 0$  para todo  $x \in I$ .

Uma primeira aplicação refere-se à caracterização de extremos locais.

**TEOREMA 6.4.2.** *Seja  $a < b \in \mathbb{R}$  e  $I = [a, b]$ . Sejam  $x_0 \in (a, b)$  e  $k \geq 2$  número inteiro. Supondo que  $f', \dots, f^{(k)}$  existam, que sejam contínuas em  $I$ , e que  $f'(x_0) = \dots = f^{(k-1)}(x_0) = 0$  mas  $f^{(k)}(x_0) \neq 0$ , temos que*

- (1) *Se  $k$  é par e  $f^{(k)}(x_0) > 0$ , então  $f$  tem mínimo local em  $x_0$ .*
- (2) *Se  $k$  é par e  $f^{(k)}(x_0) < 0$ , então  $f$  tem máximo local em  $x_0$ .*
- (3) *Se  $k$  é ímpar, então  $x_0$  não é máximo nem mínimo local.*

DEMONSTRAÇÃO. Pelo Teorema de Taylor, para  $x \in I$  existe  $\xi$  entre  $x_0$  e  $x$  tal que

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\frac{(x - x_0)^2}{2} + \cdots + f^{(k-1)}(x_0)\frac{(x - x_0)^{k-1}}{(k-1)!} \\ + f^k(\xi)\frac{(x - x_0)^k}{k!} = f(x_0) + f^k(\xi)\frac{(x - x_0)^k}{k!}.$$

Assumindo agora que  $f^{(k)}(x_0) > 0$ , como  $f^{(k)}$  é contínua então existe  $\delta > 0$  tal que  $f^{(k)}(x) > 0$  para todo  $x \in U = (x_0 - \delta, x_0 + \delta)$ . Se  $x \in U$ , então  $\xi \in U$  e então  $f^{(k)}(\xi) > 0$ . Se  $n$  é par, então para  $x \neq x_0$  temos

$$f^k(\xi)\frac{(x - x_0)^k}{k!} > 0.$$

Logo

$$x \in U \setminus \{x_0\} \implies f(x) - f(x_0) > 0 \implies x_0 \text{ é mínimo local,}$$

e portanto (1) está demonstrado.

Para demonstrar (2) o argumento é semelhante.

Finalmente, se  $k$  é ímpar, então  $(x - x_0)/k!$  é positivo para  $x > x_0$  e negativo para  $x < x_0$ . Logo  $f(x) > f(x_0)$  ou  $f(x) < f(x_0)$  dependendo do sinal de  $x - x_0$ . Logo a proposição (3) é verdadeira.  $\square$

Uma segunda aplicação diz respeito às funções convexas. Seja  $I \subset \mathbb{R}$  um intervalo. Dizemos que  $f : I \rightarrow \mathbb{R}$  é convexa em  $I$  se para todo  $t \in [0, 1]$  e  $x_1, x_2 \in I$  temos

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2).$$

Graficamente, uma função é convexa se o gráfico de  $f$  entre  $x_1$  e  $x_2$  está abaixo da reta que une os pontos  $(x_1, f(x_1))$  e  $(x_2, f(x_2))$ .

TEOREMA 6.4.3. *Seja  $I$  intervalo aberto e  $f : I \rightarrow \mathbb{R}$ . Então  $f$  é convexa se e somente se  $f''(x) \geq 0$  para todo  $x \in I$ .*

DEMONSTRAÇÃO. ( $\Leftarrow$ ) Assuma que  $f''(x) \geq 0$  para todo  $x \in I$ . Sejam  $x_1 < x_2 \in I$  e  $0 < t < 1$ . Definindo  $x_0 = (1-t)x_1 + tx_2$ , pelo Teorema de Taylor existe  $\xi_1 \in (x_1, x_0)$  e  $\xi_2 \in (x_0, x_2)$  tais que

$$f(x_1) = f(x_0) + f'(x_0)(x_1 - x_0) + \frac{1}{2}f''(\xi_1)(x_1 - x_0)^2, \\ f(x_2) = f(x_0) + f'(x_0)(x_2 - x_0) + \frac{1}{2}f''(\xi_2)(x_2 - x_0)^2.$$

Como  $f''(\xi_1) \geq 0$  e  $f''(\xi_2) \geq 0$ , então

$$(1-t)f(x_1) + tf(x_2) \\ = f(x_0) + [(1-t)x_1 + tx_2 - x_0]f'(x_0) + \frac{(1-t)}{2}f''(\xi_1)(x_1 - x_0)^2 + \frac{t}{2}f''(\xi_2)(x_2 - x_0)^2 \\ = f(x_0) + \frac{(1-t)}{2}f''(\xi_1)(x_1 - x_0)^2 + \frac{t}{2}f''(\xi_2)(x_2 - x_0)^2 \geq f(x_0).$$

Logo  $f$  é convexa.

( $\implies$ ) Sejam  $x_1 < x < x_2 \in I$ . Então  $x = (1-t)x_1 + tx_2$  para  $t = (f(x_2) - f(x_1))/(x_2 - x_1)$ . Logo, se  $f$  é convexa,

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{(1-t)f(x_1) + tf(x_2) - f(x_1)}{t(x_2 - x_1)} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

e

$$\frac{f(x_2) - f(x)}{x_2 - x} \geq \frac{f(x_2) - [(1-t)f(x_1) + tf(x_2)]}{(1-t)(x_2 - x_1)} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Portanto,

$$x_1 < x < x_2 \implies \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x},$$

e

$$x_1 < x_2 \implies f'(x_1) = \lim_{x \rightarrow x_1} \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \lim_{x \rightarrow x_2} \frac{f(x_2) - f(x)}{x_2 - x} = f'(x_2).$$

Logo  $f'$  é função crescente em  $I$  e então  $f''(x) \geq 0$  para todo  $x \in I$ .  $\square$

**6.4.1. pontos de inflexão e concavidades.** Seja  $f$  uma função real duas vezes diferenciável. Dizemos que um ponto é de inflexão se este “separa” curvas de concavidades contrárias. Se  $c$  é ponto de inflexão então  $f''(c) = 0$ . Para descobrir se um ponto  $c$  onde a segunda derivada se anula é de inflexão, basta checar se  $f''$  muda de sinal no intervalo  $(c - \epsilon, c + \epsilon)$ , para todo  $\epsilon > 0$ .

Quanto a concavidades, dizemos que  $f$  é *côncava* (para baixo) em  $(a, b)$  se  $f'' < 0$  em  $(a, b)$ . Dizemos que  $f$  é *convexa* (côncava para baixo) em  $(a, b)$  se  $f'' > 0$  em  $(a, b)$ .

## 6.5. Regra de L'Hôpital

Considere o problema de achar

$$\lim_{x \rightarrow 0} \frac{\sin x}{x},$$

se este limite existir. Surpreendentemente, vale a regra de que, nestes casos, o limite da razão das funções é igual ao limite da razão das derivadas das funções.

**TEOREMA 6.5.1.** *Sejam  $f$  e  $g$  duas funções reais diferenciáveis definidas na intervalo  $(a, b)$ . Suponha também que  $g$  e  $g'$  seja não nula e que*

$$\lim_{x \rightarrow b} f(x) = \lim_{x \rightarrow b} g(x) = 0, \text{ ou que } \lim_{x \rightarrow b} f(x) = \lim_{x \rightarrow b} g(x) = \pm\infty.$$

*Temos então que*

$$\text{se } \lim_{x \rightarrow b} \frac{f'(x)}{g'(x)} = \alpha, \text{ então } \lim_{x \rightarrow b} \frac{f(x)}{g(x)} = \alpha.$$

*Mesmo se  $\alpha$  for  $-\infty$  ou  $\infty$ , o resultado continua valendo. Vale também se  $x \rightarrow a$ , ou mesmo para pontos interiores, onde basta tomar os dois limites laterais.*

**6.6. Exercícios**

EXERCÍCIO 6.1. Assuma  $f : \mathbb{R} \rightarrow \mathbb{R}$  diferenciável em  $c \in \mathbb{R}$  e  $f(c) = 0$ . Mostre então que  $g(x) = |f(x)|$  é diferenciável em  $c$  se e somente se  $f'(c) = 0$ .

EXERCÍCIO 6.2. Seja  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por

$$f(x) = \sum_{i=1}^n (x - c_i)^2,$$

onde  $c_i \in \mathbb{R}$  para  $i = 1, \dots, n$ , e  $n \in \mathbb{N}$ . Ache um ponto de mínimo relativo de  $f$ . Mostre que é único.

EXERCÍCIO 6.3. Seja  $I \subset \mathbb{R}$  um intervalo e  $f : I \rightarrow \mathbb{R}$  diferenciável. Mostre que se  $f'$  é positiva em  $I$ , i.e.,  $f'(x) > 0$  para todo  $x \in I$ , então  $f$  é estritamente crescente.



## Funções trigonométricas, logarítmicas e exponenciais

1

Neste capítulo descrevemos algumas funções especiais, como as funções trigonométricas, o logaritmo e a exponencial.

### 7.1. Funções trigonométricas

Definimos aqui algumas funções trigonométricas, começando pelas funções seno e cosseno. Nossas definições diferem das definições geométricas “usuais”, pois usamos séries de potências. Entretanto são as mesmas funções, como pode ser visto em [Djairo Figueiredo].

**7.1.1. Senos e cossenos.** Definimos as funções  $\sin$  e  $\cos$  de  $\mathbb{R} \rightarrow \mathbb{R}$  através das séries de potências

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots, \quad \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

A derivação termo a termo nas séries de potências acima é válida, e portanto

$$\sin' x = \sum_{n=0}^{\infty} (-1)^n (2n+1) \frac{x^{2n}}{(2n+1)!} = \cos x, \quad \cos' x = \sum_{n=0}^{\infty} (-1)^n 2n \frac{x^{2n-1}}{(2n)!} = -\sin x.$$

Note que das definições, a função  $\sin$  é ímpar ( $\sin(-x) = -\sin x$ ), e a função  $\cos$  é par ( $\cos(-x) = \cos x$ ). Temos ainda uma igualdade fundamental, dada pelo resultado abaixo.

**LEMA 7.1.1.** Para todo  $x \in \mathbb{R}$ , tem-se que  $\sin^2 x + \cos^2 x = 1$ .

**DEMONSTRAÇÃO.** Seja  $f(x) = \sin^2 x + \cos^2 x$ . Então  $f'(x) = 2 \sin x \cos x - 2 \cos x \sin x = 0$ . Logo  $f$  é constante. Basta agora ver que  $f(0) = 1$ , e então

$$1 = f(0) = f(x) = \sin^2 x + \cos^2 x.$$

□

**COROLÁRIO 7.1.2.** Para todo  $x \in \mathbb{R}$ , temos que  $|\sin x| \leq 1$  e  $|\cos x| \leq 1$ .

Valem também as identidades abaixo.

**LEMA 7.1.3.** Para todo  $x \in \mathbb{R}$ , tem-se que

$$\sin(a+b) = \sin a \cos b + \sin b \cos a, \quad \cos(a+b) = \cos a \cos b - \sin a \sin b.$$

Finalmente, uma propriedade importante destas funções são suas periodicidades. Dizemos que uma função  $f : \mathbb{R} \rightarrow \mathbb{R}$  é *periódica*, com período  $T$  se  $f(x+T) = f(x)$  para todo  $x \in \mathbb{R}$ .

<sup>1</sup>Última Atualização: 02/07/2012

LEMA 7.1.4. As funções  $\sin$  e  $\cos$  são periódicas com período  $2\pi$ .

**7.1.2. Outras funções trigonométricas.** Seja  $\tan x = \sin x / \cos x$  definida em  $\mathbb{R}$  excetuando-se  $\{\pm\pi/2, \pm3\pi/2, \dots\}$ . Note que  $\tan x$  é periódica com período  $\pi$ , pois

$$\tan(x + \pi) = \frac{\sin(x + \pi)}{\cos(x + \pi)} = \frac{-\sin x}{-\cos x} = \tan x.$$

Outras funções trigonométricas são:

$$\cot x = \frac{\cos x}{\sin x}, \quad \sec x = \frac{1}{\cos x}, \quad \csc x = \frac{1}{\sin x}.$$

**7.1.3. Funções trigonométricas inversas.** Note que como a função  $\sin x$  é crescente em  $[-\pi/2, \pi/2]$ , então ela possui uma função inversa, que é denominada de  $\arcsin x$ . Mais especificamente,

$$\begin{aligned} \arcsin : [-1, 1] &\rightarrow [-\pi/2, \pi/2] \\ y &\mapsto \arcsin y. \end{aligned}$$

Esta função é diferenciável e

$$\frac{d}{dy} \arcsin y = \frac{1}{\sin'(\arcsin y)} = \frac{1}{\cos(\arcsin y)}.$$

Mas usando que  $\cos^2(\arcsin y) = 1 - \sin^2(\arcsin y) = 1 - y^2$ , concluímos que

$$\frac{d}{dy} \arcsin y = \frac{1}{\sqrt{1 - y^2}}.$$

Analogamente temos as funções  $\arccos : [-1, 1] \rightarrow [0, \pi]$  com

$$\frac{d}{dy} \arccos y = \frac{-1}{\sqrt{1 - y^2}},$$

e  $\arctan : [-\infty, \infty] \rightarrow [-\pi/2, \pi/2]$ , com

$$\frac{d}{dy} \arctan y = \frac{1}{\tan'(\arctan y)} = \frac{1}{\sec^2(\arctan y)} = \frac{1}{1 + y^2},$$

pois  $\sec^2 x = 1 + \tan^2 x$ . Existem ainda as funções trigonométricas inversas para a  $\cot x$ , a  $\sec x$  e a  $\csc x$ , sobre as quais não nos estenderemos.

## 7.2. Funções log e exponencial

Duas funções que têm importância fundamental na matemática são dadas pelo logaritmo e sua inversa, a função exponencial. Há formas diversas de definirmos estas funções, e escolhemos aquela que nos parece mais direta.

**7.2.1. O logaritmo.** Definimos, para  $x > 0$ ,

$$\log x = \int_1^x \frac{1}{s} ds.$$

Observamos diretamente da definição que se  $x > 0$ , então  $\log x > 0$  e que se  $x \in (0, 1)$ , então

$$\log x = \int_1^x \frac{1}{s} ds = - \int_x^1 \frac{1}{s} ds < 0.$$

Ainda de definição,  $\log 1 = 0$ . Temos ainda os seguintes resultados:

- (1)  $\log x$  é crescente
- (2)  $\log x$  é contínua
- (3)  $\log' x = 1/x$
- (4)  $\log(xy) = \log x + \log y$
- (5)  $\log x^r = r \log x$ , para  $r \in \mathbb{Q}$
- (6)  $\lim_{x \rightarrow \infty} \log x = \infty$
- (7)  $\lim_{x \rightarrow 0} \log x = -\infty$

As demonstrações dos resultados acima não são complicadas. O resultado (1) vem do fato que se  $x > y > 0$ , então  $\log x - \log y = \int_y^x 1/s ds > 0$ . Os fatos dados por (2) e (3) são resultados diretos da definição do log e as propriedades das integrais. Para demonstrar (4), definimos  $f(x) = \log(xy)$ , e portanto  $f'(x) = 1/x = \log' x$ . Logo  $f(x) - \log x$  é constante. Como  $f(1) - \log 1 = \log y$ , obtemos o resultado. A identidade (5) é verdadeira para  $r = 0$  pois  $\log 1 = 0$ . Para  $r$  natural, aplicamos (4)  $r - 1$  vezes, pois  $x^r = x \cdots x$ . Se  $r = 1/n$ , então usamos (4) novamente com  $x = x^{1/n} \cdots x^{1/n}$ . O caso geral para racionais positivos vem de  $x^{m/n} = x^m x^{1/n}$ . Para expoentes negativos, note que se  $r > 0$  por exemplo,  $0 = \log(x^r x^{-r}) = \log x^r + \log x^{-r}$ . Portanto,  $\log x^{-r} = -\log x^r = -r \log x$ .

**OBSERVAÇÃO.** Vale a pena ressaltar que (5) vale também para qualquer  $r$  real, mas esta afirmativa esbarra no fato de que ainda não temos uma definição para  $x^r$ , quando  $r$  não é racional. Este lapso será resolvido somente em (7.2.1).

Finalmente, para (6) (7), basta usar que o logaritmo é função crescente e considerar as seqüências  $\log 2^n$  e  $\log 2^{-n}$ .

**7.2.2. A exponencial.** Como a função log é estritamente crescente, ela é invertível. Denominando esta inversa por  $\exp x$ , onde  $\exp : \mathbb{R} \rightarrow (0, +\infty)$ , note que

- (1)  $\exp(0) = 1$
- (2)  $x > 0 \implies \exp x > 1$
- (3)  $x < 0 \implies \exp x < 1$
- (4)  $\exp x$  é contínua
- (5)  $\exp x$  é diferenciável e  $\exp' x = \exp x$
- (6)  $\exp(x + y) = \exp x \exp y$
- (7)  $\exp(\alpha x) = (\exp x)^\alpha$ , para  $\alpha \in \mathbb{R}$

A demonstração da fórmula em (5) é dada por

$$\exp' x = \frac{1}{\log'(\exp x)} = \exp x.$$

Definimos o número especial  $e = \exp 1$ . Note então de (7) que

$$\exp \alpha = \exp(\alpha 1) = (\exp 1)^\alpha = e^\alpha \text{ para } \alpha \in \mathbb{R}.$$

Com a ajuda das funções acima descritas, podemos definir

$$(7.2.1) \quad a^b = \exp(b \log a) \text{ para } a > 0 \text{ e } b \in \mathbb{R}.$$

## CAPÍTULO 8

### Integração

1

Sem entrar em detalhes a respeito da definição de integral (Riemann), enunciamos algumas propriedades importantes. Consideraremos inicialmente (Seção 8.1) somente funções limitadas em intervalos limitados, mas não necessariamente contínuas e aplicações para determinar áreas (Seção 8.2). Funções não limitadas e/ou intervalos limitados serão considerados na seção seguinte (Seção 8.3)

#### 8.1. Propriedade básicas de integrais de funções limitadas

Primeiros veremos algumas propriedades fundamentais de funções diferenciáveis, e a seguir abordaremos a importante relação entre integrabilidade e diferenciabilidade. Na última seção falaremos um pouco sobre técnicas que podem ajudar nos cálculos de algumas integrais.

**8.1.1. Algumas propriedades fundamentais.** Considere abaixo  $f : [a, b] \rightarrow \mathbb{R}$  e  $g : [a, b] \rightarrow \mathbb{R}$  limitadas,  $a < b$  números reais. As integrais serão sempre no domínio  $[a, b]$ . Temos então os seguintes resultados.

- (1) Se  $f$  for contínua, então é diferenciável
- (2) Se  $f$  for monótona (i.e., for função crescente ou decrescente), ela é integrável
- (3) A integral da soma de funções é a soma das integrais. O mesmo vale para diferença e produto por escalar  $\alpha \in \mathbb{R}$ :

$$\int_a^b \alpha f(x) + g(x) dx = \alpha \int_a^b f(x) dx + \int_a^b g(x) dx,$$

- (4) Se  $f$  e  $g$  são integráveis, então o produto  $fg$  é integrável
- (5) Existem funções não integráveis. Por exemplo:

$$f(x) = \begin{cases} 1 & \text{se } x \in \mathbb{Q} \\ -1 & \text{se } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

- (6) O exemplo do item 5 mostra que  $f^2$  pode ser integrável, mesmo que  $f$  não o seja
- (7) Se  $f$  é integrável, e  $g(x) = f(x)$  a menos de um número finito de pontos, então  $g$  é integrável e

$$\int_a^b f(x) dx = \int_a^b g(x) dx$$

---

<sup>1</sup>Última Atualização: 04/09/2012

- (8) Podemos sempre decompor uma função como a soma de suas partes positivas e negativas. Sejam

$$f^+(x) = \begin{cases} f(x) & \text{se } f(x) \geq 0 \\ 0 & \text{caso contrário} \end{cases}, \quad f^-(x) = \begin{cases} 0 & \text{se } f(x) \geq 0 \\ -f(x) & \text{caso contrário} \end{cases}$$

Note que  $f^+$  e  $f^-$  assumem somente valores positivos, e que, por construção,  $f(x) = f^+(x) - f^-(x)$  e  $|f(x)| = f^+(x) + f^-(x)$ . Quanto a integrabilidade, se  $f$  for integrável, então  $f^+$ ,  $f^-$  e  $|f|$  são integráveis. Note que  $|f|$  ser integrável não implica em  $f$  integrável, como nos mostra o exemplo apresentado no ítem 5.

- (9) Se  $f(x) \geq g(x)$  para todo  $x \in (a, b)$  (ou a menos de um número finito de pontos), então

$$\int_a^b f(x) dx \geq \int_a^b g(x) dx$$

Se  $f(x) > g(x)$  então a integral de  $f$  é estritamente maior que a integral de  $g$ . Uma consequência imediata é que se  $\alpha_0 \leq f(x) \leq \alpha_1$ , onde  $\alpha_0$  e  $\alpha_1$  são números reais, então

$$\alpha_0(b-a) \leq \int_a^b f(x) dx \leq \alpha_1(b-a)$$

- (10) O módulo da integral é menor ou igual a integral de módulo:

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$$

**8.1.2. Primitivas e o Teorema Fundamental do Cálculo.** Provavelmente o resultado mais importante em se tratando de integrais é o *Teorema Fundamental do cálculo*. Suponha que  $f : [a, b] \rightarrow \mathbb{R}$  seja integrável, e defina  $F : [a, b] \rightarrow \mathbb{R}$  por

$$F(x) = \int_a^x f(s) ds \quad \text{para } x \in [a, b].$$

Então  $F$  é contínua em  $[a, b]$ . Além disto, se  $f$  for contínua em  $c \in [a, b]$ , então  $F$  é diferenciável em  $c$  e  $F'(c) = f(c)$ . Se  $f$  for contínua em todos os pontos de seu domínio, então  $F$  é chamada de *primitiva* da  $f$ .

A continuidade da  $f$  é essencial para a diferenciabilidade de  $F$ . Por exemplo, considere  $f : [-1, 1] \rightarrow \mathbb{R}$  dada por

$$f(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 & \text{se } x \geq 0 \end{cases}.$$

Então

$$F(x) = \int_{-1}^x f(x) dx = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } x \geq 0 \end{cases}.$$

Note que  $F$  é contínua em  $[-1, 1]$ , mas não diferenciável em  $x = 0$ , ponto em que  $f$  é descontínua.

Veja também que duas primitivas de uma função diferem por uma constante. Seja  $f : [a, b] \rightarrow \mathbb{R}$  e  $F$  e  $\hat{F}$  suas primitivas. Então  $F' = \hat{F}' = f$  em todos os pontos do domínio. Logo,  $(F' - \hat{F}') = 0$  e portanto  $F' = \hat{F}' + C$  para alguma constante  $C$ .

TEOREMA 8.1.1 (Fundamental do Cálculo). *Seja  $f$  integrável. Se  $F$  é primitiva de  $f$ , então*

$$\int_a^b f(x) dx = \int_a^b F'(x) dx = F(b) - F(a).$$

É muito comum a notação

$$F \Big|_a^b = F(b) - F(a).$$

**8.1.3. Cálculo das integrais.** Pode-se notar depois de algumas tentativas que achar primitivas ou calcular o valor de integrais não é tarefa fácil. Aqui falaremos de duas técnicas comumente usadas nesta tarefa. Consideraremos aqui somente funções integráveis.

A primeira técnica, bem simples, é dada executando-se *integração por partes*. Sejam  $f$  e  $g$  diferenciáveis. Então

$$(fg) \Big|_a^b = \int_a^b (f(x)g(x))' dx = \int_a^b f'(x)g(x) dx + \int_a^b f(x)g'(x) dx$$

Logo

$$\int_a^b f'(x)g(x) dx = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x) dx.$$

Um exemplo onde este truque pode ser usado é no cálculo de  $\int_0^{2\pi} \sin^2 x dx$ . Note que

$$\begin{aligned} \int_0^{2\pi} \sin^2 x dx &= \int_0^{2\pi} \sin x \sin x dx = - \int_0^{2\pi} \cos' x \sin x dx \\ &= - \cos x \sin x \Big|_0^{2\pi} + \int_0^{2\pi} \cos x \sin' x dx = \int_0^{2\pi} \cos^2 x dx = \int_0^{2\pi} 1 - \sin^2 x dx \\ &= 2\pi - \int_0^{2\pi} \sin^2 x dx. \end{aligned}$$

Passando  $\int_0^{2\pi} \sin^2 x dx$  para o lado esquerdo temos que  $\int_0^{2\pi} \sin^2 x dx = \pi$ .

Outro resultado que é bastante útil é a mudança de variáveis no domínio de integração. Seja  $f : [a, b] \rightarrow \mathbb{R}$  integrável, e  $\phi : [c, d] \rightarrow [a, b]$  derivável, com  $\phi(c) = a$  e  $\phi(d) = b$ , e tal que  $\phi'$  seja integrável. Então

$$(8.1.1) \quad \int_a^b f(x) dx = \int_c^d f(\phi(s))\phi'(s) ds.$$

Para lembrar esta fórmula, basta notar que  $x \in [a, b]$  e  $s \in [c, d]$  estão relacionados por  $x(s) = \phi(s)$ . Logo  $dx/ds = \phi'(s)$ , e formalmente escrevemos  $dx = \phi'(s)ds$ .

Em boa parte das aplicações, usamos  $\phi$  invertível. Se denotarmos  $u = \phi^{-1}$ , temos que  $\phi'(s) = 1/u'(\phi(s))$  e portanto, de (8.1.1), temos

$$\int_a^b f(x) dx = \int_c^d \frac{f(u^{-1}(s))}{u'(\phi(s))} ds = \int_{u(a)}^{u(b)} \frac{f(u^{-1}(s))}{u'(u^{-1}(s))} ds.$$

A forma de lembrar é usando  $ds/dx = u'$  implica formalmente em  $dx = ds/u'$ .

No exemplo abaixo, vemos como usar estas identidades.

EXEMPLO 8.1. Para calcular  $\int_0^\pi \frac{\sin x}{\cos^3 x} dx$ , notamos que  $\sin x = -\cos' x$  e usamos  $u(x) = \cos x$ . Então  $du = -\sin x dx$  e

$$\int_0^\pi \frac{\sin x}{\cos^3 x} dx = - \int_{\cos 0}^{\cos \pi} \frac{1}{u^3} du = \frac{1}{2u^2} \Big|_1^{-1} = 0.$$

Na verdade, a integral acima poderia ser calculada diretamente observando-se por simetria que

$$\int_0^{\pi/2} \frac{\sin x}{\cos^3 x} dx = - \int_{\pi/2}^\pi \frac{\sin x}{\cos^3 x} dx$$

---

Como se mostra isto facilmente?

---

## 8.2. Áreas planas

Como as integrais definidas dão a área (com sinal) sob determinadas curvas, nada mais natural que usar integrais para cálculo de áreas [8].

EXEMPLO 8.2. Por exemplo, para calcular a área entre a curva  $f(x) = 2x$ , os pontos  $x = 0$  e  $x = 3$ , e o eixo dado por  $y = 0$ , há duas maneiras. Podemos usar a fórmula da área do triângulo (base  $\times$  altura/2) e ver que  $A = 3 \times 6/2 = 9$ . Usando integrais,

$$A = \int_0^3 2x dx = x^2 \Big|_0^3 = 9,$$

como era de se esperar.

É claro que nem todos cálculos de áreas são tão simples como o do exemplo acima. Podemos por exemplo, calcular áreas determinadas por curvas mais sofisticadas.

EXEMPLO 8.3. Para calcular a área  $A$  determinada pela curva  $f(x) = \sin x$  e os pontos  $x = \pi$  e  $x = 2\pi$  e o eixo dado por  $y = 0$ , basta ver que

$$\int_\pi^{2\pi} \sin x dx = -\cos x \Big|_\pi^{2\pi} = \cos \pi - \cos 2\pi = -2.$$

É claro que uma área não pode ser negativa. O que dá “errado” neste exemplo é que a função  $\sin$  é sempre negativa entre  $\pi$  e  $2\pi$ . A área determinada então é simplesmente o negativo da integral, i.e.,  $A = 2$ .

Um cuidado extra tem que ser tomado se a função tomar valores positivos e negativos no intervalo de interesse. Por exmplo, no exemplo acima, para achar a área de  $\sin$  entre  $x = 0$  e  $x = 2\pi$ , não se pode simplesmente calcular

$$\int_\pi^{2\pi} \sin x dx = 0.$$

Tem que se dividir o domínio que se quer integrar nas partes onde a função é positiva e onde é negativa. A área é dada na verdade por

$$A = \int_0^\pi \sin x dx - \int_\pi^{2\pi} \sin x dx = 4.$$



OBSERVAÇÃO. A área dada pela integral sem considerar os sinais é chamada às vezes de *área algébrica*, que pode ser negativa ou nula. A área que é sempre positiva, como as determinadas acima, é por vezes chamada *área geométrica*.

Outro problema mais interessante relacionado a áreas é o de *áreas entre curvas* (na verdade, os exemplos acima já são deste tipo, mas uma das curvas é dada por  $y = 0$ ). Considere as funções reais  $f$  e  $g$ , definidas em  $\mathbb{R}$ . Pode-se perguntar qual é a área entre as curvas  $f(x)$ ,  $g(x)$ ,  $x = a$  e  $x = b$ . Neste caso, se  $f(x) \geq g(x)$  entre  $a$  e  $b$ , então a área é dada por

$$A = \int_a^b f(x) - g(x) dx.$$

Se  $f$  for maior que  $g$  apenas em parte do domínio, a integral tem que ser “quebrada” em partes para que não surjam “áreas negativas”.

EXEMPLO 8.4. Seja  $f(x) = 2x + 3$ , e  $g(x) = x^2$ . Determine a área compreendida entre  $f$  e  $g$  e entre  $x = 1$  e  $x = 3$ . Note que em  $[1, 3]$ , temos  $f(x) \geq g(x)$ , e portanto podemos integrar  $f - g$  para calcular a área:

$$A = \int_1^3 (2x + 3 - x^2) dx = \left( x^2 + 3x - \frac{x^3}{3} \right) \Big|_1^3 = 9 + 9 - 9 - 1 - 3 + \frac{1}{3} = \frac{16}{3}.$$

### 8.3. Integrais impróprias

Integrais impróprias são integrais de funções ilimitadas, ou em domínios ilimitados, e seus valores são dados através de limites, se estes existirem. Nestes casos, dizemos que as integrais *existem*, ou *convergem*.

Em domínios ilimitados, as integrais podem ser

$$\begin{aligned} \int_a^\infty f dx &= \lim_{b \rightarrow \infty} \int_a^b f dx, & \int_{-\infty}^b f dx &= \lim_{a \rightarrow -\infty} \int_a^b f dx, \\ \int_{-\infty}^\infty f dx &= \int_{-\infty}^0 f dx + \int_0^\infty f dx. \end{aligned}$$

É importante atentar para um detalhe nas integrais em  $(-\infty, \infty)$ . Para esta existir, tem que existir o limite  $\lim_{b \rightarrow \infty} \int_0^b f dx$  e o limite  $\lim_{a \rightarrow -\infty} \int_a^0 f dx$ , *separadamente*. Note que isto é diferente de escever

$$\int_{-\infty}^\infty f dx = \lim_{a \rightarrow \infty} \int_{-a}^a f dx.$$

O valor acima é conhecido como *valor principal de Cauchy*, usado em alguns ramos da matemática. Para ver a diferença entre as duas definições, considere a função sinal de  $x$  dada por  $f(x) = \text{sgn}(x)$ . Isto é,  $f(x) = 1$  para números positivos e  $f(x) = -1$  para números negativos. Então  $f$  não é integrável em  $(-\infty, \infty)$  pois não existe  $\lim_{b \rightarrow \infty} \int_0^b f dx = \lim_{b \rightarrow \infty} b$  nem  $\lim_{a \rightarrow -\infty} \int_a^0 f dx = -\lim_{a \rightarrow -\infty} a$ . Mas o valor principal de Cauchy está bem definido pois  $\int_{-a}^a f dx = 0$  para todo  $a$  e então

$$\int_{-\infty}^\infty f dx = \lim_{a \rightarrow \infty} \int_{-a}^a f dx = 0.$$

Integrais em funções não limitadas são definidas de forma análoga. Seja  $f : (a, b) \rightarrow \mathbb{R}$  e suponha  $f$  integrável em  $[a + \delta, b]$  para todo  $\delta > 0$ . Definimos então

$$\int_a^b f dx = \lim_{\delta \rightarrow 0} \int_{a+\delta}^b f dx$$

quando este limite existir. Quando a função é ilimitada numa vizinhança de  $b$ , a definição é análoga:

$$\int_a^b f dx = \lim_{\delta \rightarrow 0} \int_a^{b-\delta} f dx$$

O último caso é quando a singularidade fica no interior do intervalo. Por exemplo, seja  $c \in (a, b)$  tal que  $f$  seja integrável em  $(a, c - \delta)$  e em  $(c + \delta, b)$ . Definimos então

$$\int_a^b f dx = \int_a^c f dx + \int_c^b f dx = \lim_{\delta \rightarrow 0} \int_a^{c-\delta} f dx + \lim_{\delta \rightarrow 0} \int_{c+\delta}^b f dx.$$

Novamente, no caso acima, os dois limites têm que existir. Por exemplo, a função dada em  $[-1, 1]$  por  $f(0) = 0$  e, se  $x \neq 0$  por  $f(x) = \text{sgn}(x)/x$  não é integrável. Mas o valor principal de Cauchy

$$\lim_{\delta \rightarrow 0} \left( \int_{-1}^{-\delta} \frac{-1}{x} dx + \int_{\delta}^1 \frac{1}{x} dx \right) = 0$$

existe.

EXEMPLO 8.5. A integral imprópria  $\int_0^1 1/\sqrt{x} dx$  está bem definida pois

$$\lim_{\delta \rightarrow 0} \int_{\delta}^1 \frac{1}{\sqrt{x}} dx = \lim_{\delta \rightarrow 0} 2\sqrt{x} \Big|_{\delta}^1 = 2.$$

EXEMPLO 8.6. As integrais impróprias  $\int_0^1 1/x dx$  e  $\int_1^{\infty} 1/x dx$  não existem pois

$$\lim_{\delta \rightarrow 0} \int_{\delta}^1 \frac{1}{x} dx = \lim_{\delta \rightarrow 0} \log x \Big|_{\delta}^1 = +\infty, \quad \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x} dx = \lim_{b \rightarrow \infty} \log x \Big|_1^b = +\infty.$$

Vemos neste exemplo que  $\lim_{x \rightarrow \infty} f(x) = 0$  não garante que  $f$  seja integrável.

A exemplo de séries, dizemos que uma integral imprópria de uma função  $f$  converge absolutamente se  $|f|$  for integrável. Temos então o seguinte resultado.

TEOREMA 8.3.1. *Se a integral de  $f$  converge absolutamente, então a integral de  $f$  converge. Em outras palavras,  $|f|$  integrável implica em  $f$  integrável.*

TEOREMA 8.3.2. *Sejam  $f \geq 0$  e  $g \geq 0$ , com  $f \leq g$  no domínio de integração. Então,  $\int g$  convergente implica em  $\int f$  convergente. e  $\int f$  divergente implica em  $\int g$  divergente.*

## CAPÍTULO 9

### Sequências e Séries

<sup>1</sup> Neste capítulo veremos sequências e séries. Como série nada mais é que um caso particular de sequências, veremos os dois tópicos de forma unificada, chamando atenção para as possíveis diferenças.

#### 9.1. Definição e resultados preliminares

Uma sequência em  $\mathbb{R}$  é simplesmente uma função de  $\mathbb{N}$  em  $\mathbb{R}$ . Portanto  $X : \mathbb{N} \rightarrow \mathbb{R}$  indica uma sequência de números reais, que escrevemos também como  $(x_n)$ , ou ainda  $(x_1, x_2, x_3, \dots)$ . Para indicar o  $n$ -ésimo valor da sequência escrevemos simplesmente  $x_n$ .

Sejam  $c_1, c_2, \dots$  números reais. Uma série

$$(9.1.1) \quad \sum_{i=1}^{\infty} c_i$$

pode ser compreendida através da sequência  $s_n$  definida por

$$(9.1.2) \quad s_n = \sum_{i=1}^n c_i.$$

Em geral, a expressão (9.1.1) nem sempre faz sentido, enquanto (9.1.2) está sempre bem-definida.

EXEMPLO 9.1.  $x_n = (-1)^n$  define a sequência  $(-1, 1, -1, 1, -1, 1, -1, \dots)$ .

EXEMPLO 9.2. A sequência de Fibonacci é definida recursivamente por  $x_1 = 1$ ,  $x_2 = 1$ , e  $x_{n+1} = x_n + x_{n-1}$  para  $n \geq 2$ . Portanto temos  $(x_n) = (1, 1, 2, 3, 5, 8, \dots)$ .

Podemos realizar com sequências várias das operações que realizamos com números reais, como por exemplo somar, subtrair, etc. Sejam por exemplo  $(x_n)$  e  $(y_n)$  duas sequências em  $\mathbb{R}$ , e  $c \in \mathbb{R}$ . Então definimos

$$(x_n) + (y_n) = (x_n + y_n), \quad (x_n) - (y_n) = (x_n - y_n), \quad (x_n) \cdot (y_n) = (x_n y_n), \quad c(x_n) = (cx_n).$$

EXEMPLO 9.3. Se  $x_n = (2, 4, 6, 8, \dots)$  e  $(y_n) = (1, 1/2, 1/3, 1/4, \dots)$ , então  $x_n \cdot y_n = (2, 2, 2, \dots)$ .

A primeira pergunta que surge quando tratamos de sequências é quanto à convergência destas, isto é, se quando  $n$  aumenta, os termos  $x_n$  se aproximam de algum valor real. Note que para isto, não importa o que acontece com finitos termos da sequência, mas sim seu comportamento assintótico com respeito a  $n$ . Em outras palavras queremos determinar o comportamento das sequências no “limite”.

---

<sup>1</sup>Última Atualização: 31/08/2012

**DEFINIÇÃO 9.1.1.** Dizemos que  $x \in \mathbb{R}$  é limite de uma sequência  $(x_n)$ , se para todo  $\epsilon > 0$ , existe  $N \in \mathbb{N}$  tal que  $|x - x_n| < \epsilon$  para todo  $n \geq N$ . Escrevemos neste caso que  $x_n \rightarrow x$ , ou que  $x = \lim x_n$ , ou ainda

$$x = \lim_{n \rightarrow \infty} x_n.$$

De forma resumida,  $x_n \rightarrow x$  se para todo  $\epsilon$  existir  $N \in \mathbb{N}$  tal que

$$n \geq N \implies |x - x_n| < \epsilon.$$

Se uma sequência não tem limite, dizemos que ela diverge ou é divergente.

A definição para séries é análoga.

**DEFINIÇÃO 9.1.2.** Dizemos que  $s \in \mathbb{R}$  é limite de uma série  $\sum_{i=1}^{\infty} c_n$ , se para todo  $\epsilon > 0$ , existe  $N \in \mathbb{N}$  tal que

$$\left| s - \sum_{i=1}^n c_n \right| < \epsilon \quad \text{para todo } n \geq N.$$

Escrevemos neste caso que  $x = \sum_{i=1}^{\infty} c_n$ . Se uma série não tem limite, dizemos que ela diverge ou é divergente.

**EXEMPLO 9.4.** Se  $x_n = 1$ , então  $\lim x_n = 1$ . De fato, dado  $\epsilon > 0$ , para todo  $n \geq 1$  temos  $|x_n - 1| = 0 < \epsilon$ .

**EXEMPLO 9.5.**  $\lim(1/n) = 0$ . De fato, dado  $\epsilon > 0$ , seja  $N$  tal que  $1/N < \epsilon$ . Logo, para todo  $n > N$  temos  $|1/n - 0| = 1/n < 1/N < \epsilon$ .

**EXEMPLO 9.6.**  $(0, 2, 0, 2, 0, 2, 0, 2, \dots)$  não converge para 0. De fato, tome  $\epsilon = 1$ . Então para todo  $N \in \mathbb{N}$  temos  $2N > N$  e  $x_{2N} = 2$ . Portanto  $|x_{2N} - 0| = 2 > \epsilon$ .

Observe que diferentes situações ocorrem nos exemplos acima. No primeiro, a sequência é constante, e a escolha de  $N$  independe de  $\epsilon$ . Já no exemplo seguinte,  $N$  claramente depende de  $\epsilon$ .

A seguir, no exemplo 9.6 o objetivo é mostrar que um certo valor  $x$  não é o limite da sequência  $(x_n)$ . Mostramos então que existe pelo menos um certo  $\epsilon > 0$  tal que para todo  $N$ , conseguimos achar  $n > N$  tal que  $|x_n - x| > \epsilon$ . Note que o que fizemos foi *negar* a convergência.

Talvez a segunda pergunta mais natural em relação aos limites de sequências é quanto a unicidade destes, quando existirem. A resposta é afirmativa, como mostra o resultado abaixo.

**TEOREMA 9.1.3** (Unicidade de limite). *Uma sequência (uma série) pode ter no máximo um limite.*

**DEMONSTRAÇÃO.** Considere que  $(x_n)$  é uma sequência de reais tal que  $x_n \rightarrow x$  e  $x_n \rightarrow x'$ , com  $x \neq x'$ . Sejam  $\epsilon = |x - x'|/2 > 0$ , e sejam  $N$  e  $N' \in \mathbb{N}$  tais que  $|x_n - x| < \epsilon$  para todo  $n > N$  e  $|x_n - x'| < \epsilon$  para todo  $n > N'$ . Logo, se  $n > \max\{N, N'\}$ , então

$$|x - x'| \leq |x - x_n| + |x_n - x'| < 2\epsilon = |x - x'|.$$

Como um número não pode ser estritamente menor que ele mesmo, temos uma contradição. Portanto  $x = x'$  e o limite é único.  $\square$

Para mostrar convergência, podemos usar o resultado seguinte.

TEOREMA 9.1.4. *Seja  $(x_n)$  uma sequência em  $\mathbb{R}$ . Então as afirmativas são equivalentes.*

(1)  $(x_n)$  converge para  $x$ .

(2) Para toda vizinhança  $V$  de  $x$  existe  $N \in \mathbb{N}$  tal que

$$n \geq N \implies x_n \in V.$$

DEMONSTRAÇÃO. Fica como exercício.  $\square$

As vezes, uma sequência se aproxima de algum valor de forma mais lenta que alguma outra sequência que converge para 0. É possível assim garantir convergência, como o resultado a seguir nos mostra.

LEMA 9.1.5. Seja  $(a_n)$  sequência em  $\mathbb{R}$  convergente para 0. Se para  $(x_n)$  sequência em  $\mathbb{R}$  existir  $c > 0$  tal que

$$|x_n - x| \leq c|a_n| \quad \text{para todo } n \in \mathbb{N},$$

então  $x_n \rightarrow x$ .

DEMONSTRAÇÃO. Como  $(a_n)$  converge, dado  $\epsilon > 0$ , seja  $N \in \mathbb{N}$  tal que  $|a_n| < \epsilon/c$  para todo  $n > N$ . Logo

$$|x_n - x| \leq c|a_n| < \epsilon \quad \text{para todo } n > N,$$

e  $\lim x_n = x$ .  $\square$

COROLÁRIO 9.1.6. Seja  $(a_n)$  sequência em  $\mathbb{R}$  convergente para 0. Se para  $(x_n)$  sequência em  $\mathbb{R}$  existir  $c > 0$  e  $N \in \mathbb{N}$  tal que

$$|x_n - x| \leq c|a_n| \quad \text{para todo } n \geq N,$$

então  $x_n \rightarrow x$ .

EXEMPLO 9.7. Seja  $x_n = (2/n) \sin(1/n)$ . Então

$$|x_n - 0| \leq 2 \frac{1}{n}.$$

Como  $1/n \rightarrow 0$ , podemos usar o lema acima para garantir que  $\lim[(2/n) \sin(1/n)] = 0$ .

Uma outra noção importante é o de limitação de uma sequência. Neste caso, mesmo quando a sequência não converge, podemos conseguir alguns resultados parciais, como veremos mais a seguir.

DEFINIÇÃO 9.1.7. *Dizemos que uma sequência  $(x_n)$  é limitada quando existe um número real  $M$  tal que  $|x_n| \leq M$  para todo  $n \in \mathbb{N}$ .*

Um primeiro resultado intuitivo é que toda sequência convergente é limitada. De fato, é razoável pensar que se a sequência converge, ela não pode ter elementos arbitrariamente grandes em valor absoluto.

TEOREMA 9.1.8. *Toda sequência convergente é limitada*

**DEMONSTRAÇÃO.** Seja  $(x_n)$  sequência convergente e seja  $x$  seu limite. Seja  $\epsilon = 1$ . Como  $(x_n)$  converge, existe  $N$  tal que  $|x - x_n| < 1$  para todo  $n > N$ . Logo, usando a desigualdade triangular temos

$$|x_n| \leq |x_n - x| + |x| < 1 + |x| \quad \text{para todo } n > N.$$

Falta agora limitar os  $N$  primeiros termos da sequência. Seja então

$$M = \max\{|x_1|, |x_2|, |x_3|, \dots, |x_N|, 1 + |x|\}.$$

Portanto  $|x_n| \leq M$  para todo  $n \in \mathbb{N}$ . □

Existem situações em que a sequência não converge, mas *tendem* para  $+\infty$  ou  $-\infty$ . Por exemplo, as sequências  $(1, 0, 9, 0, 25, 0, 36, \dots)$  e  $(n^2)$  não convergem pois não são limitadas. Entretanto, para a sequência dada por  $x_n = n^2$ , para todo número  $C > 0$  existe  $N \in \mathbb{N}$  tal que

$$n > N \implies x_n > M.$$

quando isto acontece, dizemos que  $x_n \rightarrow +\infty$  ou que  $\lim x_n = +\infty$ . Definição análoga vale para  $-\infty$ .

Outro resultado importante trata de limites de sequências que são resultados de operações entre sequências. Por exemplo, dadas duas sequências convergente, o limite da soma das sequências é a soma dos limites. E assim por diante.

**LEMA 9.1.9.** Seja  $(x_n)$  e  $(y_n)$  tais que  $\lim x_n = x$  e  $\lim y_n = y$ . Então

- (1)  $\lim(x_n + y_n) = x + y$ .
- (2)  $\lim(x_n - y_n) = x - y$ .
- (3)  $\lim(x_n y_n) = xy$ .
- (4)  $\lim(cx_n) = cx$ , para  $c \in \mathbb{R}$ .
- (5) se  $y_n \neq 0$  para todo  $n$  e  $y \neq 0$ , então  $\lim(x_n/y_n) = x/y$ .

**DEMONSTRAÇÃO.** (1) Dado  $\epsilon > 0$ , seja  $N \in \mathbb{N}$  tal que  $|x_n - x| < \epsilon/2$  e  $|y_n - y| < \epsilon/2$  para todo  $n \geq N$ . Logo

$$|x_n + y_n - (x + y)| \leq |x_n - x| + |y_n - y| < \epsilon \quad \text{para todo } n \geq N.$$

(2) A demonstração é basicamente a mesma de (1), tomando-se o devido cuidado com os sinais.

(3) Para todo  $n \in \mathbb{N}$  temos

$$|x_n y_n - xy| \leq |x_n y_n - x_n y| + |x_n y - xy| = |x_n| |y_n - y| + |y| |x_n - x|.$$

Seja  $M \in \mathbb{R}$  tal que  $|x_n| < M$  e  $|y| < M$ . Tal constante  $M$  existe pois como  $(x_n)$  converge, ela é limitada. Agora, dado  $\epsilon > 0$ , seja  $N$  tal que  $|y_n - y| < \epsilon/(2M)$  e  $|x_n - x| < \epsilon/(2M)$  para todo  $n \geq N$ . Logo,

$$|x_n y_n - xy| \leq M[|y_n - y| + |x_n - x|] < \epsilon,$$

para todo  $n \geq N$ .

Deixamos (4) e (5) como exercícios para o leitor. □

**OBSERVAÇÃO.** Os resultados do lema acima continuam válidos para um número finito de somas, produtos, etc.

Outro resultado interessante nos diz que se  $a_n$  é sequência de números positivos, então  $x_n \rightarrow 0$  se e somente se  $1/x_n \rightarrow \infty$  (resultado análogo vale trocando-se  $+\infty$  por  $-\infty$ ). Ver exercício 9.3.

Outros resultados importantes para tentar achar um “candidato” limite vêm a seguir. O primeiro nos diz que se temos uma sequência de números positivos, então o limite, se existir, tem que ser não negativo, podendo ser zero. A seguir, aprendemos que se temos uma sequência “sanduichadas” entre outras duas sequências convergentes que tem o mesmo limite, então a sequência do meio converge e tem também o mesmo limite.

**LEMA 9.1.10.** Seja  $(x_n)$  convergente com  $\lim x_n = x$ . Se existe  $N \in \mathbb{N}$  tal que  $x_n \geq 0$  para todo  $n > N$ , então  $x \geq 0$ .

**DEMONSTRAÇÃO.** (por contradição) Assuma que  $x < 0$ . Seja então  $\epsilon = -x/2 > 0$ . Como  $(x_n)$  converge para  $x$ , seja  $N \in \mathbb{N}$  tal que  $|x_n - x| < \epsilon$  para todo  $n > N$ . Logo,  $x_{N+1} \in (x - \epsilon, x + \epsilon)$ , isto é,  $x_{N+1} < x + \epsilon = x/2 < 0$ . Obtivemos então uma contradição pois  $x_{N+1}$  não é negativo.  $\square$

**COROLÁRIO 9.1.11.** Se  $(x_n)$  e  $(y_n)$  são convergentes com  $\lim x_n = x$  e  $\lim y_n = y$ , e se existe  $N \in \mathbb{N}$  tal que  $x_n \geq y_n$  para todo  $n > N$ , então  $x \geq y$ .

**DEMONSTRAÇÃO.** Se  $z_n = x_n - y_n$ , então  $\lim z_n = \lim x_n - \lim y_n = x - y$ . O presente resultado segue então do Lema 9.1.10.  $\square$

**LEMA 9.1.12** (sanduíche de sequências). Sejam  $(x_n)$ ,  $(y_n)$  e  $(z_n)$  sequências tais que  $x_n \leq y_n \leq z_n$  para todo  $n > N$ , para algum  $N \in \mathbb{N}$ . Assuma ainda que  $(x_n)$  e  $(z_n)$  convergem com  $\lim x_n = \lim z_n$ . Então  $(y_n)$  converge e  $\lim y_n = \lim x_n = \lim z_n$ .

**DEMONSTRAÇÃO.** Seja  $a = \lim x_n = \lim z_n$ . Dado  $\epsilon > 0$ , existe  $N$  tal que  $|x_n - a| < \epsilon$  e  $|z_n - a| < \epsilon$  para todo  $n > N$ . Logo

$$-\epsilon < x_n - a \leq y_n - a \leq z_n - a < \epsilon \implies |x_n - a| < \epsilon$$

para todo  $n > N$ , como queríamos demonstrar.  $\square$

**EXEMPLO 9.8.**  $(n)$  diverge pois não é limitada.

**EXEMPLO 9.9.** Seja  $S_n = 1 + 1/2 + 1/3 + 1/4 + \dots + 1/n$ . Mostraremos que  $(S_n)$  não é limitada, e portanto divergente. Note que

$$\begin{aligned} x_{2^n} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^{n-1} + 1} + \frac{1}{2^n}\right) \\ &= 1 + \frac{1}{2} + \sum_{i=3}^4 \frac{1}{i} + \sum_{i=5}^8 \frac{1}{i} + \dots + \sum_{i=2^{n-1}+1}^{2^n} \frac{1}{i} > 1 + \frac{1}{2} + \sum_{i=3}^4 \frac{1}{i} + \sum_{i=5}^8 \frac{1}{i} + \dots + \sum_{i=2^{n-1}+1}^{2^n} \frac{1}{i} \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots + \frac{1}{2} = 1 + \frac{n}{2}. \end{aligned}$$

Logo  $(S_n)$  não é limitada, e portanto diverge.

Outra forma de ver que a sequência acima diverge é por indução. Quero mostrar que  $S_{2^n} \geq 1 + n/2$ . Note que  $S_2 = 1 + 1/2$ . Assumindo que  $S_{2^{n-1}} \geq 1 + (n-1)/2$  temos

$$S_{2^n} = S_{2^{n-1}} + \frac{1}{2^{n-1} + 1} + \dots + \frac{1}{2^n} > 1 + \frac{(n-1)}{2} + \frac{1}{2} > 1 + \frac{n}{2},$$

como queríamos demonstrar. Mais uma vez a conclusão é que  $(S_n)$  não é limitada, logo diverge.

EXEMPLO 9.10.  $\lim_{n \rightarrow \infty} ((2n + 1)/n) = 2$ . De fato,

$$\frac{2n + 1}{n} = (2) + \left(\frac{1}{n}\right).$$

Como  $\lim_{n \rightarrow \infty} (2) = 2$  e  $\lim_{n \rightarrow \infty} (1/n) = 0$ , nós obtemos o resultado.

EXEMPLO 9.11.  $\lim_{n \rightarrow \infty} (2n/(n^2 + 1)) = 0$ , pois

$$\frac{2n}{n^2 + 1} = \frac{2/n}{1 + 1/n^2}.$$

Como  $\lim_{n \rightarrow \infty} (2/n) = 0$  e  $\lim_{n \rightarrow \infty} (1 + 1/n^2) = 1 \neq 0$ , podemos aplicar o resultado sobre quociente de seqüências.

EXEMPLO 9.12. A seqüência

$$x_n = \frac{1}{n^2} \sum_{i=1}^n i$$

converge. Primeiro note que

$$(9.1.3) \quad \sum_{i=1}^n i = \frac{n^2 + n}{2}.$$

Para  $n = 1$  o resultado (9.1.3) é trivial. Assuma (9.1.3) verdadeiro para  $n = k$ . Temos então que

$$\sum_{i=1}^{k+1} i = \frac{k^2 + k}{2} + k + 1 = \frac{k^2 + 3k + 2}{2} = \frac{(k + 1)^2 + (k + 1)}{2},$$

e portanto fórmula (9.1.3) é verdadeira. Temos então que

$$x_n = \frac{n^2 + n}{2n^2} = \frac{1}{2} \left(1 + \frac{1}{n}\right) = \frac{1}{2} + \left(\frac{1}{2n}\right).$$

Logo  $(x_n)$  é soma de duas seqüências convergentes,  $(1/2)$  e  $(1/2)(1/n)$  e

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \frac{1}{2} + \lim_{n \rightarrow \infty} \frac{1}{2n} = \frac{1}{2}.$$

EXEMPLO 9.13. Seja  $(x_n)$  seqüência convergente em  $\mathbb{R}$ , e seja  $x \in \mathbb{R}$  seu limite. Então a seqüência definida por

$$\frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

converge e tem  $x$  como seu limite.

Sem perda de generalidade, supomos que  $(x_n)$  converge para zero. Para o caso geral quando  $(x_n)$  converge para  $x$  basta tratar a seqüência  $(x_n - x)$ .

Seja  $S_n = (x_1 + x_2 + \cdots + x_n)/n$ . Como  $(x_n)$  converge, então é limitada. Seja  $M$  tal que  $|x_n| < M$  para todo  $n \in \mathbb{N}$ . Dado  $\epsilon > 0$ , seja  $N^*$  tal que  $M/N^* < \epsilon$  e  $\sup\{|x_n| : n \geq N^*\} < \epsilon$ . Então, temos  $S_n = \check{S}_n + \hat{S}_n$ , onde

$$\check{S}_n = \frac{1}{n}(x_1 + x_2 + \cdots + x_{N^*}), \quad \hat{S}_n = \frac{1}{n}(x_{N^*} + x_{N^*+1} + \cdots + x_n).$$



Então  $(S_n)$  é a soma de duas seqüências convergentes. De fato para  $n > (N^*)^2$ , temos  $|\check{S}_n| \leq N^*M/n \leq M/N^* < \epsilon$ . Além disso,  $|\hat{S}_n| < \epsilon(n - N^*)/n < \epsilon$ . Portanto  $(S_n)$  converge.

EXEMPLO 9.14.  $\lim_{n \rightarrow \infty} ((\sin n)/n) = 0$  pois como  $-1 \leq \sin n \leq 1$ , então

$$-1/n \leq (\sin n)/n \leq 1/n,$$

e o resultado segue do lema 9.1.12.

Outro resultado importante refere-se à convergência do valor absoluto de seqüências: se uma seqüência converge, então a seqüência de valores absolutos também converge. A recíproca *não* é verdadeira. Basta considerar como contra-exemplo a seqüência  $((-1)^n)$ . Neste caso a seqüência diverge mas a seqüência de seus valores absolutos converge.

LEMA 9.1.13. Seja  $(x_n)$  convergente. Então  $(|x_n|)$  também o é.

DEMONSTRAÇÃO. Exercício. □

LEMA 9.1.14 (teste da razão). Seja  $(x_n)$  seqüência de números positivos tal que  $(x_{n+1}/x_n)$  converge e  $\lim_{n \rightarrow \infty} (x_{n+1}/x_n) < 1$ . Então  $(x_n)$  converge e  $\lim_{n \rightarrow \infty} (x_n) = 0$ .

DEMONSTRAÇÃO. Seja  $L = \lim_{n \rightarrow \infty} (x_{n+1}/x_n)$ . Então, por hipótese,  $L < 1$ . Seja  $r$  tal que  $L < r < 1$ . Portanto dado  $\epsilon = r - L > 0$ , existe  $N$  tal que  $x_{n+1}/x_n < L + \epsilon = r$  para todo  $n \geq N$ . Logo,

$$0 < x_{n+1} < x_n r < x_{n-1} r^2 < x_{n-2} r^3 < \dots < x_N r^{n-N+1} \quad \text{para todo } n \geq N.$$

Se  $c = x_N r^{-N}$ , então  $0 < x_{n+1} < c r^{n+1}$ . O resultado segue do Corolário 9.1.6, pois como  $r < 1$ , então  $\lim_{n \rightarrow \infty} r^n = 0$ . □

COROLÁRIO 9.1.15. Seja  $(x_n)$  tal que  $x_n \neq 0$  para todo  $n \in \mathbb{N}$  e

$$L = \lim_{n \rightarrow \infty} \frac{|x_{n+1}|}{|x_n|}$$

existe e  $L > 1$ . Então para todo  $C \in \mathbb{R}$  existe  $N^* \in \mathbb{N}$  tal que

$$n \geq N^* \implies |x_n| > C.$$

DEMONSTRAÇÃO. basta considerar o teste da razão para  $y_n = 1/x_n$ . Neste caso,

$$\lim_{n \rightarrow \infty} \frac{|y_{n+1}|}{|y_n|} = \lim_{n \rightarrow \infty} \frac{|x_n|}{|x_{n+1}|} = \lim_{n \rightarrow \infty} \frac{1}{\frac{|x_{n+1}|}{|x_n|}} = \frac{1}{\lim_{n \rightarrow \infty} \frac{|x_{n+1}|}{|x_n|}} = \frac{1}{L} < 1.$$

Logo  $(y_n)$  converge para zero, e para todo  $C \in \mathbb{R}^+$  existe  $N^*$  tal que

$$n \geq N^* \implies |y_n| < \frac{1}{C}.$$

Portanto para  $n \geq N^*$  temos  $|x_n| > C$  e  $(x_n)$  não é limitada e não converge. □

EXEMPLO 9.15. Seja  $(x_n) = n/2^n$ . Então

$$\lim_{n \rightarrow \infty} \left( \frac{x_{n+1}}{x_n} \right) = \lim_{n \rightarrow \infty} \left( \frac{n+1}{2^{n+1}} \frac{2^n}{n} \right) = \frac{1}{2} \lim_{n \rightarrow \infty} \left( \frac{n+1}{n} \right) = \frac{1}{2}.$$

Pelo teste da razão temos  $\lim_{n \rightarrow \infty} (x_n) = 0$

EXEMPLO 9.16. Note que para  $x_n = 1/n$ , temos  $\lim_{n \rightarrow \infty} x_{n+1}/x_n = 1$  e  $(x_n)$  converge. Entretanto, para  $x_n = n$ , temos  $\lim_{n \rightarrow \infty} x_{n+1}/x_n = 1$  mas  $(x_n)$  não converge. Portanto o teste não é conclusivo quando o limite da razão entre os termos é um.

Um resultado importante em se tratando de séries versa sobre o comportamento assintótico dos termos que a compõem. Temos o seguinte resultado.

LEMA 9.1.16. Seja a série dada por  $\sum_{i=1}^{\infty} c_i$  convergente. Então  $\lim_{i \rightarrow \infty} c_i = 0$ .

DEMONSTRAÇÃO. Note que se a série converge para um valor  $S$ , as sequências parciais  $S_n = \sum_{i=1}^n c_i$  convergem para o mesmo valor  $S$ . Logo

$$\lim_{i \rightarrow \infty} c_i = \lim_{i \rightarrow \infty} (S_i - S_{i-1}) = \lim_{i \rightarrow \infty} S_i - \lim_{i \rightarrow \infty} S_{i-1} = S - S = 0.$$

□

OBSERVAÇÃO. Do resultado acima, concluímos que se  $c_i$  não tem limite, ou se seu limite não é zero, a série não converge.

Uma classe de séries especial é dada pelas *séries geométricas*. Para  $r \in \mathbb{R}$ , esta série é dada por

$$\sum_{i=0}^{\infty} r^i.$$

Note que as somas parciais  $S_n = \sum_{i=0}^n r^i$  divergem se  $|r| \geq 1$ , e convergem se  $|r| < 1$ . De fato, se  $|r| \geq 1$ , então  $r^i$  não converge a zero, e portanto a série não converge. Para  $|r| < 1$ , temos

$$S_n = \frac{1 - r^{n+1}}{1 - r},$$

e portanto a série converge para  $\lim_{n \rightarrow \infty} S_n = 1/(1 - r)$ .

## 9.2. Sequências Monótonas

Um classe muito especial de sequências é a de sequências monótonas. Uma sequência monótona é tal que seus valores não “oscilam”, i.e., eles ou nunca diminuem ou nunca aumentam. Pode-se ver que a definição de sequência monótona é restrita a uma dimensão.

DEFINIÇÃO 9.2.1. Dizemos que uma sequência  $(x_n)$  é crescente se  $x_1 < x_2 < \dots < x_n < \dots$  e não decrescente se  $x_1 \leq x_2 \leq \dots \leq x_n \leq \dots$ . Similarmente, uma sequência  $(x_n)$  é decrescente se  $x_1 > x_2 > \dots > x_n > \dots$  e não crescente se  $x_1 \geq x_2 \geq \dots \geq x_n \geq \dots$ . Finalmente, uma sequência é monótona se for crescente ou decrescente.

EXEMPLO 9.17.  $(1, 2, 3, 4, \dots)$  é crescente, e  $(1, 2, 3, 3, 3, 3, \dots)$  não decrescentes.

EXEMPLO 9.18.  $(1/n)$  é decrescente.

EXEMPLO 9.19.  $(-1, 1, -1, 1, -1, \dots)$  não é monótona.

TEOREMA 9.2.2. Uma sequência não crescente ou não decrescente é convergente se e somente se é limitada.

Além disso, se  $(x_n)$  é não decrescente, então  $\lim_{n \rightarrow \infty} (x_n) = \sup\{x_n : n \in \mathbb{N}\}$ . Da mesma forma, se  $(x_n)$  é não crescente, então  $\lim_{n \rightarrow \infty} (x_n) = \inf\{x_n : n \in \mathbb{N}\}$ .

DEMONSTRAÇÃO. ( $\implies$ ) Já vimos que toda sequência convergente é limitada.

( $\impliedby$ ) Assuma  $(x_n)$  não decrescente e limitada. Seja  $x = \sup x_n : n \in \mathbb{N}$ . Então dado  $\epsilon > 0$ , existe  $N$  tal que  $x - \epsilon < x_N \leq x < x + \epsilon$ , pois  $x$  é o supremo. Logo, para todo  $n > N$  temos  $x - \epsilon < x_N \leq x_n \leq x < x + \epsilon$ , portanto  $x_n$  converge para  $x$ . Se a sequência for não-crescente, a demonstração é análoga.  $\square$

TEOREMA 9.2.3. *Uma sequência de números reais  $(x_n)$ , monótona não decrescente e limitada converge para “seu supremo”, i.e., converge para  $\sup\{x_n : n \in \mathbb{N}\}$ .*

DEMONSTRAÇÃO. Seja  $x = \sup\{x_n : n \in \mathbb{N}\}$  (que existe pois a sequência é limitada). Então pela definição de supremo, para todo  $\epsilon > 0$ , existe  $N^* \in \mathbb{N}$  tal que  $x_{N^*} \in (x - \epsilon, x)$ . Logo como a sequência é monótona não decrescente, temos

$$n > N^* \implies x_n > x_{N^*} > x - \epsilon.$$

Mas para todo  $n \in \mathbb{N}$  temos  $x_n \leq x$  por definição de supremo. Logo

$$n > N^* \implies x_n > x_{N^*} > x - \epsilon \quad \text{e} \quad x_n < x + \epsilon \implies x_n \in (x - \epsilon, x + \epsilon).$$

$\square$

EXEMPLO 9.20.  $(a^n)$  diverge se  $a > 1$  pois é ilimitada.

EXEMPLO 9.21.  $(a^n)$  converge se  $0 < a \leq 1$  pois é monótona decrescente e limitada. Além disso,  $\lim_{n \rightarrow \infty} (a^n) = 0$ , pois  $\inf\{a_n : n \in \mathbb{N}\} = 0$ .

EXEMPLO 9.22. (Bartle?) Seja  $y_1 = 1$  e  $y_{n+1} = (1 + y_n)/3$ . Mostraremos que  $(y_n)$  é convergente e achamos seu limite. Note que  $y_2 = 2/3 < 1 = y_1$ . Vamos mostrar por indução que  $0 < y_{n+1} < y_n$ . Esta afirmativa vale para  $n = 1$ . Assuma verdadeira para  $n = k - 1$ , isto é  $0 < y_k < y_{k-1}$ . Então para  $n = k$  temos

$$y_{k+1} = (1 + y_k)/3 < (1 + y_{k-1})/3 = y_k,$$

e como  $y_k > 0$ , então  $y_{k+1} > 0$ , como queríamos. Portanto a sequência é monótona não crescente e limitada inferiormente por zero. Portanto converge. Seja  $y$  seu limite. Então

$$y = \lim_{n \rightarrow \infty} y_{n+1} = \lim_{n \rightarrow \infty} (1 + y_n)/3 = (1 + y)/3.$$

Logo  $y = 1/2$ .

EXEMPLO 9.23. Seja  $y_1 = 1$ , e  $y_{n+1} = (2y_n + 3)/4$ . Note que  $y_2 = 5/4 > y_1$ . Para mostrar que  $y_{n+1} > y_n$  em geral, usamos indução. Note que para  $n = 1$  o resultado vale. Assuma agora que valha também para  $n = k$  para algum  $k$ , i.e.,  $y_{k+1} > y_k$ . Então

$$y_{k+2} = \frac{1}{4}(2y_{k+1} + 3) > \frac{1}{4}(2y_k + 3) = y_{k+1}.$$

Logo, por indução,  $y_{n+1} > y_n$  para todo  $n \in \mathbb{N}$ , e  $(y_n)$  é não decrescente. Para mostrar que é limitada, note que  $|y_1| < 2$ . Mais uma vez usamos indução a fim de provar que em geral  $|y_n| < 2$ . Assuma que  $|y_k| < 2$ . Logo,

$$|y_{k+1}| = \left| \frac{1}{4}(2y_{k+1} + 3) \right| < \frac{1}{4}(2|y_{k+1}| + 3) < \frac{7}{4} < 2.$$

Por indução, segue-se que  $|y_n| < 2$  para todo  $n \in \mathbb{N}$ . Como  $(y_n)$  é monótona e limitada, então é convergente. Seja  $y = \lim_{n \rightarrow \infty} (y_n)$ . Então

$$y = \lim_{n \rightarrow \infty} (y_n) = \lim_{n \rightarrow \infty} ((2y_n + 3)/4) = ((2y + 3)/4).$$

resolvendo a equação algébrica acima, temos  $y = 3/2$ .

EXEMPLO 9.24. Assuma  $0 < a < b$ , e defina  $a_0 = a$  e  $b_0 = b$ . Seja

$$a_{n+1} = \sqrt{a_n b_n}, \quad b_{n+1} = \frac{1}{2}(a_n + b_n),$$

para  $n \in \mathbb{N}$ . Então  $(a_n)$  e  $(b_n)$  convergem para o mesmo limite.

Vamos mostrar por indução que

$$(9.2.1) \quad a_{i+1} > a_i, \quad a_i < b_i, \quad b_{i+1} < b_i \quad \text{para } i = 0, 1, \dots$$

Para  $i = 0$  temos  $a_0 = a < b = b_0$ . Logo, usando que  $y > x$  implica em  $\sqrt{y} > \sqrt{x}$ , e que  $a_0$  e  $b_0$  são positivos, temos  $a_1 = \sqrt{a_0 b_0} > a_0$ . Além disso,  $b_1 = (a_0 + b_0)/2 < b_0$  pois  $a_0 < b_0$ . Portanto (9.2.1) vale para  $i = 0$ . Assuma que valha também para  $i = n$ . Então  $a_{n+1} = \sqrt{a_n b_n} > a_n$ . Além disso,  $b_{n+1} = (a_n + b_n)/2 < b_n$  e  $b_{n+1} = (a_n + b_n)/2 > a_n$  pois  $a_n < b_n$  por hipótese. Então  $a_{n+1} = \sqrt{a_n b_n} < \sqrt{b_{n+1} b_n} < b_{n+1}$ . Logo (9.2.1) vale também para  $i = n + 1$ . Portanto temos que  $(a_n)$  é monótona não decrescente e limitada superiormente, enquanto  $(b_n)$  é monótona não crescente e limitada superiormente. Ambas então convergem e sejam  $A$  e  $B$  seus limites. Neste caso teremos

$$A = \sqrt{AB}, \quad B = \frac{1}{2}(A + B).$$

e portanto  $A = B$ .

### 9.3. Séries

Devido às suas peculiaridades, as séries possuem algumas propriedades próprias. Detalharemos abaixo a questão de convergência absoluta e condicional, e testes de convergência próprios para séries. Dizemos que a série  $\sum_{i=1}^{\infty} c_i$  converge *absolutamente* se a série  $\sum_{i=1}^{\infty} |c_i|$  converge. E dizemos que  $\sum_{i=1}^{\infty} c_i$  converge *condicionalmente* se  $\sum_{i=1}^{\infty} c_i$  converge, mas não converge absolutamente, i.e.,  $\sum_{i=1}^{\infty} |c_i|$  diverge. O resultado abaixo nos garante que a convergência absoluta é suficiente para uma série convergir. A primeira demonstração usa a noção de s'erie de Cauchy, e apresentamos uma segunda demonstração que dispensa este conceito.

TEOREMA 9.3.1. *Toda série que converge absolutamente é convergente.*

DEMONSTRAÇÃO USANDO SEQUÊNCIAS DE CAUCHY. Seja  $\sum_{i=1}^{\infty} c_i$  tal que  $\sum_{i=1}^{\infty} |c_i|$  converja, e denote por  $s_n = \sum_{i=1}^n c_i$  as somas parciais. Note que, para  $n > m$ ,

$$|s_n - s_m| \leq \sum_{i=m}^n |c_i| = \hat{s}_n - \hat{s}_m,$$

onde  $s_n = \sum_{i=1}^n |c_i|$ . Como  $\hat{s}_n$  é de Cauchy, então  $s_n$  é também de Cauchy, e portanto convergente.  $\square$

DEMONSTRAÇÃO ALTERNATIVA [12]. Seja  $p_i = \max\{a_i, 0\}$ , e  $q_i = \min\{a_i, 0\}$ . Então  $|a_i| = p_i - q_i$  e  $a_i = p_i + q_i$ . Então  $p_i \leq |a_i|$  e  $q_i \leq |a_i|$ . Logo  $\sum_{i=1}^{\infty} p_i$  e  $\sum_{i=1}^{\infty} q_i$  convergem pois são limitadas superiormente por  $\sum_{i=1}^{\infty} |a_i|$ . Logo

$$s_n = \sum_{i=1}^n a_i = \sum_{i=1}^n p_i + \sum_{i=1}^n q_i$$

converge. □

OBSERVAÇÃO. Note que na demonstração acima, que se  $\sum_{i=1}^{\infty} a_i$  é condicionalmente convergente, então  $\sum_{i=1}^{\infty} p_i$  e  $\sum_{i=1}^{\infty} q_i$  divergem. De fato, se apenas uma destas séries convergisse, então

$$\sum_{i=1}^{\infty} a_i = \sum_{i=1}^{\infty} p_i + \sum_{i=1}^{\infty} q_i$$

divergiria. Se as duas convergissem, então

$$\sum_{i=1}^{\infty} |a_i| = \sum_{i=1}^{\infty} p_i - \sum_{i=1}^{\infty} q_i$$

convergiria.

---

comentar aqui sobre os teoremas 8 e 9, [12, pag. 44]

---

Prosseguimos agora descrevendo alguns testes que podem ser usados para definir convergências de séries.

LEMA 9.3.2 (Teste da razão). Suponha que  $\lim_{i \rightarrow \infty} |c_{i+1}/c_i| = L$ . Então se  $L < 1$ , a série  $\sum_{i=1}^{\infty} c_i$  converge absolutamente, e se  $L > 1$ , a série diverge.

DEMONSTRAÇÃO. Faremos aqui a prova no caso em que  $|c_{i+1}/c_i| < L < 1$  para todo inteiro  $i$ . Note que neste caso,

$$|c_i| < L|c_{i-1}| < L^2|c_{i-2}| < \cdots < L^{i-1}|c_1|.$$

Então  $\sum_{i=1}^{\infty} |c_i| < |c_1| \sum_{i=1}^{\infty} L^{i-1}$  converge absolutamente pois  $|L| < 1$ . □

LEMA 9.3.3 (Teste da raiz). Suponha que  $\lim_{i \rightarrow \infty} \sqrt[i]{|c_i|} = L$ . Então se  $L < 1$ , a série  $\sum_{i=1}^{\infty} c_i$  converge absolutamente, e se  $L > 1$ , a série diverge.

DEMONSTRAÇÃO. Consideraremos somente o caso  $\sqrt[i]{|c_i|} < L < 1$  para todo  $i$ . Então  $\sum_{i=1}^{\infty} |c_i| < \sum_{i=1}^{\infty} L^i$  que converge absolutamente pois  $|L| < 1$ . □

OBSERVAÇÃO. Tanto o teste da razão como a da raiz são inconclusivo caso  $L = 1$ , como os exemplos  $\sum_{i=1}^{\infty} 1/i$  e  $\sum_{i=1}^{\infty} 1/i^2$  ilustram.

LEMA 9.3.4 (Teste da integral). Seja  $f : [0, +\infty) \rightarrow \mathbb{R}$  não negativa e monótona decrescente. Então  $\sum_{i=1}^{\infty} f(i)$  converge se e somente se  $\int_0^{\infty} f(x) dx < +\infty$ .

**9.4. Exercícios**

EXERCÍCIO 9.1. Demonstrar o Teorema 9.1.13.

EXERCÍCIO 9.2. Demonstrar o Teorema 9.1.4.

EXERCÍCIO 9.3. Mostre que se  $a_n$  é sequência de números positivos, então  $x_n \rightarrow 0$  se e somente se  $1/x_n \rightarrow \infty$ .

EXERCÍCIO 9.4. Seja  $(x_n)$  tal que  $x_n \neq 0$  para todo  $n \in \mathbb{N}$  e

$$L = \lim_{n \rightarrow \infty} \frac{|x_{n+1}|}{|x_n|}$$

existe e  $L > 1$ . Mostre que para todo  $C \in \mathbb{R}$  existe  $N^* \in \mathbb{N}$  tal que

$$n \geq N^* \implies |x_n| > C.$$

## CAPÍTULO 10

### Funções de várias variáveis

<sup>1</sup> Neste capítulo estudamos as propriedades de funções de várias variáveis. Nos concentramos principalmente nas questões relativas a diferenciabilidade destas funções e aplicações.

#### 10.1. Introdução

Consideramos aqui funções  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ , onde, tipicamente,  $m = 2, 3$  e  $n = 1$ . Assim como no caso unidimensional, dizemos que uma função  $f$  é contínua em  $\mathbf{x}_0$  se dado  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$\|\mathbf{x} - \mathbf{x}_0\| < \delta \implies \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \epsilon.$$

Analogamente ao caso unidimensional, podemos definir o conceito de limite de funções num determinado ponto, ver Exercício 10.1.

Outro conceito importante é o de curvas (em duas dimensões) ou superfícies (em três dimensões) de nível, que chamaremos sempre de curvas de nível. Uma curva de nível é dada pelo conjunto de pontos que têm imagem constante, i.e., dado  $c \in \mathbb{R}$ , definimos

$$\{\mathbf{x} \in \mathbb{R}^m : f(\mathbf{x}) = c\}.$$

#### 10.2. Derivadas parciais e planos tangentes

Seja  $f : \mathbb{R}^m \rightarrow \mathbb{R}$ . Então definimos a  $i$ -ésima derivada parcial de  $f$  num ponto  $\mathbf{x} = (x_1, \dots, x_m)$  por

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_m) - f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m)}{h}$$

quando o limite acima existir. Se cada uma das derivadas parciais existirem, definimos o *vetor gradiente* (também chamado de *gradiente* pelos íntimos) dada por

$$\nabla f(\mathbf{x}) = \left( \frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_m}(\mathbf{x}) \right).$$

Note-se que depois de obtida uma derivada parcial, pode-se derivar novamente para se obter segundas derivadas parciais. Por exemplo, derivando-se na direção  $i$  e depois na direção  $j$  obtemos a função  $\partial^2 f / \partial x_j \partial x_i$ . Pra funções “suaves”, a ordem em que se deriva não importa. Em particular, se  $\partial^2 f / \partial x_j \partial x_i(\mathbf{x}_0)$  e  $\partial^2 f / \partial x_i \partial x_j(\mathbf{x}_0)$  existem e são contínuas numa vizinhança de  $\mathbf{x}_0$ , então

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0).$$

---

<sup>1</sup>Última Atualização: 04/09/2012

Considere agora o caso bidimensional e  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ . Com o conceito de derivadas parciais, é possível, quando a função é suave o suficiente, definir o plano tangente ao gráfico de  $f$  num determinado ponto  $\mathbf{p}_0 = (x_0, y_0) \in \mathbb{R}^2$ . Começamos por lembrar que a reta tangente no caso  $m = 1$  é definida pelos pontos  $(x, z)$  tais que

$$z - f(x_0) - f'(x_0)(x - x_0) = 0.$$

Fixando primeiramente  $y_0$  temos que  $f(x, y_0)$  é uma função de uma dimensão, e a reta tangente na direção  $x$  então é dada pelos pontos  $(x, y_0, z)$  tais que

$$(10.2.1) \quad z - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) = 0.$$

De forma análoga, fixando-se  $x_0$ , obtém-se a reta tangente à  $f(x, y_0)$  na direção  $y$ , e esta é dada pelos pontos  $(x_0, y, z)$  tais que

$$(10.2.2) \quad z - f(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0.$$

Dadas as duas retas acima, tangentes ao gráfico de  $f$ , podemos determinar o plano tangente  $P$  a este mesmo gráfico. Tomando  $x = x_0 + 1$  em (10.2.1) e  $y = y_0 + 1$  em (10.2.2), temos que

$$\left( x_0 + 1, y_0, f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \right), \quad \left( x_0, y_0 + 1, f(x_0, y_0) + \frac{\partial f}{\partial y}(x_0, y_0) \right)$$

pertencem ao plano  $P$ . Como  $(x_0, y_0, f(x_0, y_0))$  também pertence a  $P$ , então os vetores

$$\mathbf{v}_x = \left( 1, 0, \frac{\partial f}{\partial x}(x_0, y_0) \right), \quad \mathbf{v}_y = \left( 0, 1, \frac{\partial f}{\partial y}(x_0, y_0) \right)$$

são tangentes a  $P$ . Portanto

$$\mathbf{N} = \mathbf{v}_x \times \mathbf{v}_y = \left( \frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1 \right)$$

é normal a  $P$ . Como temos que

$$P = \{(x, y, z) : (x - x_0, y - y_0, z - f(x_0, y_0)) \cdot \mathbf{N} = 0\}$$

então  $P$  é definido pela equação

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) - z + f(x_0, y_0) = 0.$$

### 10.3. Diferenciabilidade

A noção de diferenciabilidade e de derivada em dimensões maiores simplesmente generaliza de forma natural a derivada unidimensional. Seja  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  e  $\mathbf{x} \in \mathbb{R}^m$ . Dizemos que  $\mathbf{f}$  é diferenciável em  $\mathbf{x}$  se

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}\|}{\|\mathbf{h}\|} = 0,$$



onde  $\mathbf{f}'(\mathbf{x})$  tem a representação matricial  $[\mathbf{f}'(\mathbf{x})] \in \mathbb{R}^{n \times m}$  dada por

$$[\mathbf{f}'(\mathbf{x})] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix}.$$

A matriz  $[\mathbf{f}'(\mathbf{x})]$  também é chamada de *matriz jacobiana* de  $f$  no ponto  $\mathbf{x}$ . Chamamos  $\mathbf{f}'(\mathbf{x})$  de derivada de  $\mathbf{f}$  em  $\mathbf{x}$ , e que também denotamos por  $D\mathbf{f}(\mathbf{x})$ .

Assim como em uma dimensão,  $\mathbf{f}$  é diferenciável em  $\mathbf{x}$  se e somente se existir uma função  $\mathbf{r} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  tal que

$$(10.3.1) \quad \mathbf{f}(\mathbf{x} + \mathbf{h}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{h}) + \mathbf{r}(\mathbf{h}) \quad \text{com} \quad \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{r}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0.$$

Note que pela identidade acima, temos imediatamente que diferenciabilidade implica em continuidade.

EXEMPLO 10.1. Considere  $\mathbf{A} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  aplicação linear e  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  dada por  $\mathbf{f}(\mathbf{x}) = \mathbf{A}(\mathbf{x}) + \mathbf{c}$ , onde  $\mathbf{c} \in \mathbb{R}^n$  é vetor constante. Então  $\mathbf{f}'(\mathbf{x})(\mathbf{h}) = \mathbf{A}(\mathbf{h})$  para todo  $\mathbf{x}, \mathbf{h} \in \mathbb{R}^m$ . Note que neste caso, a derivada  $\mathbf{f}'(\mathbf{x})$  é na verdade *independente* de  $\mathbf{x}$ .

EXEMPLO 10.2. Seja a matriz  $A \in \mathbb{R}^{m \times n}$  e  $\mathbf{c} = (c_1, \dots, c_m)$  vetor constante. Considere ainda  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  tal que para  $\mathbf{x} = (x_1, \dots, x_m)$  tem-se  $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_n(\mathbf{x}))$  e

$$\vec{f}(\mathbf{x}) = A\vec{x} + \vec{c} \quad \text{i.e.,} \quad f_i(\mathbf{x}) = \sum_{j=1}^m A_{ij}x_j + c_j \quad \text{para } i = 1, \dots, n.$$

Então, para  $\mathbf{h} = (h_1, \dots, h_m)$  tem-se  $\mathbf{f}'(\mathbf{x})(\mathbf{h}) = \mathbf{y}$  onde  $\vec{y} = A\vec{h}$ , i.e.,

$$y_i = \sum_{j=1}^m A_{ij}h_j.$$

Compare com o exemplo 10.1.

Uma interessante forma de analisarmos uma função em várias variáveis é restringindo esta função numa direção e usando propriedades de funções de apenas uma variável. Para tanto, sejam  $\mathbf{u} \in \mathbb{R}^m$  com  $\|\mathbf{u}\| = 1$ , e  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ . Dado  $\mathbf{x} \in \mathbb{R}^m$ , definimos a *derivada direcional de  $\mathbf{f}$  em  $\mathbf{x}$  na direção  $\mathbf{u}$*  por  $D_{\mathbf{u}}\mathbf{f}(\mathbf{x}) \in \mathbb{R}^n$ , onde

$$D_{\mathbf{u}}\mathbf{f}(\mathbf{x}) = \lim_{t \rightarrow 0} \frac{\mathbf{f}(\mathbf{x} + t\mathbf{u}) - \mathbf{f}(\mathbf{x})}{t} = 0.$$

quando o limite acima existir.

No caso em que  $\mathbf{u} = \mathbf{e}_i$ , então temos a *derivada parcial* em relação à  $i$ -ésima coordenada e escrevemos

$$D_{\mathbf{e}_i}\mathbf{f}(\mathbf{x}) = \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{x}).$$

É importante ressaltar que a *existência de derivadas parciais em relação às coordenadas não implica na existência de derivadas direcionais em geral*. Considere o simples exemplo abaixo.

EXEMPLO 10.3. Seja  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por

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

Então

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0,$$

mas a derivada direcional na direção  $(a, b)$  não existe se  $a$  e  $b$  são não nulos, pois não existe o limite quando  $t \rightarrow 0$  de

$$\frac{f(ta, tb) - f(0, 0)}{t} = \frac{1}{t} \frac{a}{b}.$$

A situação muda se supusermos diferenciabilidade, como mostra o resultado a seguir.

TEOREMA 10.3.1. *Seja  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  diferenciável em  $\mathbf{x} \in \mathbb{R}^m$ . Seja  $\mathbf{u} \in \mathbb{R}^m$  com  $\|\mathbf{u}\| = 1$ . Então existe a derivada direcional  $\mathbf{D}_{\mathbf{u}}\mathbf{f}(\mathbf{x})$ , e esta é dada por*

$$\mathbf{D}_{\mathbf{u}}\mathbf{f}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})(\mathbf{u}).$$

DEMONSTRAÇÃO. Como  $\mathbf{f}$  é diferenciável em  $\mathbf{x}$ , então para todo  $\epsilon > 0$  existe  $\delta > 0$  tal que

$$\mathbf{h} \in \mathbb{R}^m, 0 < \|\mathbf{h}\| < \delta \implies \frac{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})(\mathbf{h})\|}{\|\mathbf{h}\|} < \epsilon.$$

Tomando  $\mathbf{h} = t\mathbf{u}$ , com  $|t| \in (0, \delta)$ , temos

$$\left\| \frac{\mathbf{f}(\mathbf{x} + t\mathbf{u}) - \mathbf{f}(\mathbf{x})}{t} - \mathbf{f}'(\mathbf{x})(\mathbf{u}) \right\| < \epsilon.$$

Logo

$$\lim_{t \rightarrow 0} \frac{\mathbf{f}(\mathbf{x} + t\mathbf{u}) - \mathbf{f}(\mathbf{x})}{t} = \mathbf{f}'(\mathbf{x})(\mathbf{u}),$$

e portanto a derivada direcional existe e é dada por  $\mathbf{f}'(\mathbf{x})(\mathbf{u})$ . □

OBSERVAÇÃO. Considerando  $n = 1$ , temos que

$$(10.3.2) \quad \mathbf{D}_{\mathbf{u}}f(\mathbf{x}) = f'(\mathbf{x})(\mathbf{u}) = \nabla f(\mathbf{x}) \cdot \mathbf{u} = \|\nabla f(\mathbf{x})\| \|\mathbf{u}\| \cos \theta \leq \|\nabla f(\mathbf{x})\| \|\mathbf{u}\|,$$

onde  $\theta$  é o ângulo formado por  $\nabla f(\mathbf{x})$  e  $\mathbf{u}$ . Portanto, a derivada direcional atinge seu maior valor se  $\theta = 0$ , i.e., quando  $\mathbf{u}$  aponta na direção do gradiente. Outra forma de se ler a equação (10.3.2) é observar que a direção do gradiente é a direção de crescimento máximo da função, e que a direção contrária ao gradiente é a direção de menor crescimento (menor derivada).

A existência de derivadas direcionais *não implica em diferenciabilidade*. Para ilustrar tal fato, considere a função

$$f(x, y) = \begin{cases} \frac{x^2}{y} & \text{se } y \neq 0, \\ 0 & \text{se } y = 0. \end{cases}$$

Então

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0,$$

mas dado  $(a, b)$  com  $\|(a, b)\|^2 = a^2 + b^2 = 1$  e  $b \neq 0$ , temos

$$\lim_{t \rightarrow 0} \frac{f(ta, tb) - f(0, 0)}{t} = \frac{a^2}{b},$$

e a derivada direcional é dada por

$$(10.3.3) \quad D_{(a,b)}f(0, 0) = \frac{a^2}{b}.$$

Entretanto, se  $f$  fosse diferenciável, teríamos

$$D_{(a,b)}f(0, 0) = \mathbf{f}'(0, 0)(a, b) = \frac{\partial f}{\partial x}(0, 0)a + \frac{\partial f}{\partial y}(0, 0)b = 0,$$

uma contradição com (10.3.3). Logo  $f$  não é diferenciável em  $(0, 0)$  apesar de ter todas as derivadas direcionais neste ponto. Note que  $f(x, x^2) = 1$  para  $x \neq 0$ , e portanto  $f$  é descontínua em  $(0, 0)$ .

Apesar da existência de derivadas direcionais num determinado ponto não garantir a diferenciabilidade neste ponto, a existência e continuidade das derivadas parciais numa *vizinhança* dum ponto garante a diferenciabilidade, como podemos ver no resultado a seguir.

**TEOREMA 10.3.2.** *Seja  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}$ . Se  $\partial \mathbf{f} / \partial x_i$  existir e for contínua numa vizinhança de  $\mathbf{x}$  para  $i = 1, \dots, m$ , então  $\mathbf{f}$  é diferenciável em  $\mathbf{x}$ .*

Outro resultado de grande importância diz respeito à diferenciabilidade de composições de funções, garantindo que se duas funções são diferenciáveis, então a composição também o é.

**TEOREMA 10.3.3 (Regra da Cadeia).** *Sejam  $\mathbf{f} : \mathbb{R}^l \rightarrow \mathbb{R}^m$  e  $\mathbf{g} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ . Se  $\mathbf{f}$  é diferenciável em  $\mathbf{x} \in \mathbb{R}^l$  e  $\mathbf{g}$  é diferenciável em  $\mathbf{f}(\mathbf{x})$ , então  $\mathbf{g} \circ \mathbf{f}$  é diferenciável em  $\mathbf{x}$  e*

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}) = \mathbf{g}'(\mathbf{f}(\mathbf{x})) \circ \mathbf{f}'(\mathbf{x}).$$

Veremos agora várias aplicações da regra da cadeia.

• **Aplicação 1:**

**EXEMPLO 10.4.** Seja  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , e seja a função  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  inversa de  $\mathbf{f}$ , isto é,

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}, \quad \mathbf{f}(\mathbf{g}(\mathbf{y})) = \mathbf{y},$$

para todo  $\mathbf{x}, \mathbf{y}$  em  $\mathbb{R}^n$ . Se  $\mathbf{f}$  é diferenciável em  $\mathbf{x} \in \mathbb{R}^n$ , e  $\mathbf{g}$  é diferenciável em  $\mathbf{y} = \mathbf{f}(\mathbf{x})$ , então  $\mathbf{f}'(\mathbf{x})$  e  $\mathbf{g}'(\mathbf{y})$  são inversas uma da outra, isto é,

$$\mathbf{f}'(\mathbf{x}) \circ \mathbf{g}'(\mathbf{y}) = \mathbf{g}'(\mathbf{y}) \circ \mathbf{f}'(\mathbf{x}) = \mathbf{I},$$

onde  $\mathbf{I}$  é o operador identidade  $\mathbf{I}(\mathbf{x}) = \mathbf{x}$ .

De fato, seja  $\mathbf{h}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}$ . Derivando  $\mathbf{h}(\mathbf{x}) = \mathbf{x}$ , temos  $\mathbf{h}'(\mathbf{x}) = \mathbf{I}$ . Usando a regra da cadeia para  $\mathbf{h}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$ , temos  $\mathbf{h}'(\mathbf{x}) = \mathbf{g}'(\mathbf{y})\mathbf{f}'(\mathbf{x})$ . Logo,  $\mathbf{g}'(\mathbf{y})\mathbf{f}'(\mathbf{x}) = \mathbf{I}$ . De forma análoga segue-se que  $\mathbf{f}'(\mathbf{x})\mathbf{g}'(\mathbf{y}) = \mathbf{I}$ .

- **Aplicação 2:** Uma aplicação imediata da regra da cadeia é dada no seguinte teorema do valor médio para funções de várias variáveis. Na verdade, esta é uma aplicação imediata do teorema do valor médio unidimensional (Teorema 6.3.3) quando restringimos uma função de várias variáveis a um segmento de reta.

**TEOREMA 10.3.4.** *Seja  $f : \mathbb{R}^m \rightarrow \mathbb{R}$ , diferenciável. Sejam  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$  e  $S = \{\mathbf{x} + t(\mathbf{y} - \mathbf{x}) : t \in (0, 1)\}$ . Então existe  $\xi \in S$  tal que*

$$f(\mathbf{y}) - f(\mathbf{x}) = \mathbf{f}'(\xi)(\mathbf{y} - \mathbf{x}).$$

**DEMONSTRAÇÃO.** Este resultado segue-se de uma aplicação do teorema do valor médio unidimensional (Teorema 6.3.3) para a função  $\phi : [0, 1] \rightarrow \mathbb{R}$  dada por  $\phi(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$ . Note ainda que pela regra da cadeia temos que

$$\phi'(t) = \mathbf{f}'(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}).$$

□

É interessante notar que não vale uma “generalização trivial” para o teorema do valor médio quando a imagem de uma função está no  $\mathbb{R}^n$ , para  $n \geq 2$ . Como exemplo, considere a função  $\phi : \mathbb{R} \rightarrow \mathbb{R}^2$  dada por  $\phi(t) = (\sin t, \cos t)$ . Tomando-se os pontos  $t = 0$  e  $t = 2\pi$ , vemos que não existe  $\xi \in [0, 2\pi]$  tal que

$$\mathbf{0} = \phi(0) - \phi(2\pi) = \phi'(\xi)(2\pi - 0) = 2\pi\phi'(\xi).$$

pois  $\phi'(\xi) \neq \mathbf{0}$  para todo  $\xi$ .

Existe entretanto o seguinte resultado para funções em  $\mathbb{R}^n$ .

**TEOREMA 10.3.5.** *Seja  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ , diferenciável. Sejam  $\mathbf{x}_0, \mathbf{x}_1 \in \mathbb{R}^m$  e seja  $S$  o segmento de reta unindo estes pontos. Então existe  $\xi \in S$  tal que*

$$\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)\| \leq \|\mathbf{f}'(\xi)(\mathbf{x}_1 - \mathbf{x}_0)\|.$$

**DEMONSTRAÇÃO.** Seja  $\mathbf{v} = \mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)$ , e  $\phi(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \cdot \mathbf{v}$ . Então

$$\phi(\mathbf{x}_1) - \phi(\mathbf{x}_0) = (\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)) \cdot \mathbf{v} = \|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)\|^2,$$

e  $\phi'(\mathbf{x})(\mathbf{h}) = [\mathbf{f}'(\mathbf{x})(\mathbf{h})] \cdot \mathbf{v}$ . Pelo Teorema do valor médio dado pelo Teorema 10.3.4, existe  $\xi \in S$  tal que  $\phi(\mathbf{x}_1) - \phi(\mathbf{x}_0) = \phi'(\xi)(\mathbf{x}_1 - \mathbf{x}_0)$ , i.e.,

$$\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)\|^2 = [\mathbf{f}'(\xi)(\mathbf{x}_1 - \mathbf{x}_0)] \cdot (\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)) \leq \|\mathbf{f}'(\xi)(\mathbf{x}_1 - \mathbf{x}_0)\| \|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)\|.$$

Finalmente, se  $\mathbf{f}(\mathbf{x}_1) = \mathbf{f}(\mathbf{x}_0)$ , o resultado é trivial. Caso contrário dividimos ambos os lados da desigualdade acima para concluir a demonstração. □

**OBSERVAÇÃO.** Podemos continuar a desigualdade acima mais um passo e obter

$$\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)\| \leq \|\mathbf{f}'(\xi)\| \|\mathbf{x}_1 - \mathbf{x}_0\|.$$

**Aplicação 3:** Gradientes são ortogonais a curvas de níveis. Seja  $\phi : \mathbb{R} \rightarrow \mathbb{R}^2$  parametrização “suave” de uma curva de nível de uma função  $f(x, y)$ . Então  $f(\phi(t))$  é constante, e

$$0 = \frac{d}{dt}f(\phi(t)) = \nabla f(\phi(t)) \cdot \phi'(t).$$

Portanto, como  $\phi'$  é tangente a curva de nível da  $f$  em  $\mathbf{x}$ , temos que  $\nabla f(\mathbf{x})$  é ortogonal a curvas de nível de  $f$  no ponto  $\mathbf{x}$ .

**Aplicação 4:** Uma outra aplicação da regra da cadeia tem a ver com funções homogêneas. Dizemos que  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  é *homogênea de grau  $k$*  se  $f(t\mathbf{x}) = t^k f(\mathbf{x})$  para todo  $t \in \mathbb{R}$ . Por exemplo,  $x^2/(yz)$  é homogênea de grau zero, e  $\sqrt{x^5}$  é homogênea de grau  $5/2$ . Um Teorema devido à Euler afirma que  $f$  é homogênea de grau  $k$  se e somente se

$$(10.3.4) \quad \nabla f(\mathbf{x}) \cdot \mathbf{x} = kf(\mathbf{x}).$$

Para obter (10.3.4) a partir de uma função homogênea  $f$ , basta derivar  $f(t\mathbf{x}) = t^k f(\mathbf{x})$  em relação a  $t$  usando a regra da cadeia:

$$\mathbf{x} \cdot \nabla f(t\mathbf{x}) = kt^{k-1}f(\mathbf{x}),$$

e depois tomar  $t = 1$ .

**Aplicação 5:** É possível obter a regra de Leibniz através da regra da cadeia. Seja

$$F(t_1, t_2, t_3) = \int_{t_1}^{t_2} f(t_3, y) dy.$$

Então

$$\frac{d}{dx}F(u(x), v(x), x) = \frac{\partial F}{\partial t_1} \frac{du}{dx} + \frac{\partial F}{\partial t_2} \frac{dv}{dx} + \frac{\partial F}{\partial t_3},$$

implica em

$$\frac{d}{dx} \int_{u(x)}^{v(x)} f(x, y) dy = -f(x, u(x)) \frac{du}{dx} + f(x, v(x)) \frac{dv}{dx} + \int_{u(x)}^{v(x)} \frac{\partial f(x, y)}{\partial x} dy.$$

## 10.4. Matriz Hessiana, Fórmula de Taylor e pontos críticos

Note que a derivada de uma função de uma função de  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  num determinado ponto  $\mathbf{x}$  foi definida como uma aplicação linear de  $\mathbb{R}^m$  em  $\mathbb{R}$ . No caso, para  $\mathbf{x}$  fixo, teríamos  $\mathbf{f}'(\mathbf{x}) : \mathbb{R}^m \rightarrow \mathbb{R}$  dada por

$$\mathbf{f}'(\mathbf{x})(\mathbf{y}) = \frac{\partial f}{\partial x_1}(\mathbf{x})y_1 + \frac{\partial f}{\partial x_2}(\mathbf{x})y_2 + \cdots + \frac{\partial f}{\partial x_m}(\mathbf{x})y_m,$$

onde  $\mathbf{y} \in \mathbb{R}^m$ .

De forma análoga, definimos a segunda derivada de  $f$  num ponto  $\mathbf{x}$  fixado como sendo a função bilinear  $\mathbf{f}''(\mathbf{x}) : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$  tal que

$$\mathbf{f}''(\mathbf{x})(\mathbf{y}, \mathbf{z}) = \sum_{i,j=1}^m \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} y_i z_j, \quad \text{onde} \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right),$$

e  $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ . Uma forma mais compacta de escrever a definição acima é usando-se a *matriz hessiana*  $H$  dada por  $H_{ij}(\mathbf{x}) = \partial^2 f(\mathbf{x}) / \partial x_i \partial x_j$ . Logo

$$f''(\mathbf{x})(\mathbf{y}, \mathbf{z}) = (\bar{\mathbf{y}})^t H(\mathbf{x}) \bar{\mathbf{z}}.$$

**OBSERVAÇÃO.** Um interessante resultado garante que se  $f$  for suficientemente suave num determinado ponto  $\mathbf{x}_0$  (é suficiente que as segundas derivadas existam e sejam contínuas numa vizinhança aberta de  $\mathbf{x}_0$ ) teremos que *não importa a ordem em que se toma as derivadas*, i.e.,  $\partial^2 f / \partial x_i \partial x_j = \partial^2 f / \partial x_j \partial x_i$ , e portanto a matriz hessiana é simétrica. Este tipo de resultado, com diferentes hipóteses, é atribuído à *Clairaut* em [26], e à *Schwarz* em [2, 15].

Definições para derivadas de ordem mais alta seguem o mesmo formato, sendo estas *aplicações multilineares*. Entretanto para os nossos propósitos, a matriz hessiana basta.

Apresentamos no teorema a seguir a fórmula de Taylor, e nos restringimos ao caso particular de polinômios quadráticos. Este teorema será de fundamental importância para caracterizarmos pontos extremos.

**TEOREMA 10.4.1 (Taylor).** *Seja  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  duas vezes diferenciável em  $\Omega$ , com derivadas contínuas. Para  $\mathbf{x} \in \mathbb{R}^m$  e  $\mathbf{h} \in \mathbb{R}^m$ , existe  $\hat{t} \in (0, 1)$  tal que para  $\boldsymbol{\xi} = \mathbf{x} + \hat{t}\mathbf{h}$  tem-se*

$$(10.4.1) \quad f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{h}) + \frac{1}{2} f''(\boldsymbol{\xi})(\mathbf{h}, \mathbf{h}).$$

**DEMONSTRAÇÃO.** Seja  $\phi : [0, 1] \rightarrow \mathbb{R}$  dada por  $\phi(t) = f(\mathbf{x} + t\mathbf{h})$ . Aplicando o Teorema de Taylor em uma dimensão (Teorema 6.4.1), obtemos que existe  $\hat{t} \in (0, 1)$  tal que

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2} \phi''(\hat{t}).$$

Note que

$$\phi'(t) = \mathbf{f}'(\mathbf{x} + t\mathbf{h})(\mathbf{h}) = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(\mathbf{x} + t\mathbf{h}) h_i, \quad \phi''(t) = \sum_{j=1}^m \sum_{i=1}^m \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x} + t\mathbf{h}) h_i h_j,$$

e usando a definição de  $\phi$  obtemos o resultado diretamente. □

**OBSERVAÇÃO.** Note que exigindo que as segundas derivadas sejam contínuas, podemos usar o fato de que a “ordem” das segundas derivadas não importam.

Assim como em uma dimensão, usaremos o Teorema de Taylor para estudarmos pontos extremos de uma função. Dizemos que  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  tem um *máximo local* em  $\mathbf{x} \in \mathbb{R}^m$  se existe  $\delta > 0$  tal que

$$(10.4.2) \quad \|\mathbf{y} - \mathbf{x}\| < \delta \implies f(\mathbf{y}) \leq f(\mathbf{x}).$$

Dizemos que  $\mathbf{x}$  é *máximo estrito local* se valer a desigualdade estrita em (10.4.2). Definição análoga serve para *mínimo local* e *mínimo estrito local*. Chamamos um ponto de máximo ou mínimo local de *ponto extremo local*, e um ponto de máximo ou mínimo estrito local de *ponto extremo estrito local*.

O resultado que obtemos a seguir, relativo a pontos extremos interiores, é análogo ao caso unidimensional, ver o Teorema 6.3.1, e diz primeiro que pontos extremos interiores são *pontos críticos*, i.e., pontos em que a derivada se anula. O resultado mostra também que

se um ponto  $\mathbf{x}$  é de mínimo local, então a forma bilinear  $f''(\mathbf{x})$  é *semi-definida positiva*, i.e.,  $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) \geq 0$  para todo  $\mathbf{h} \in \mathbb{R}^m$ . De forma análoga se um ponto é de máximo local, então  $f''(\mathbf{x})$  é *semi-definida negativa*, i.e.,  $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) \leq 0$  para todo  $\mathbf{h} \in \mathbb{R}^m$ .

Em termos matriciais,  $f''(\mathbf{x})$  é semi-definida positiva se a matriz hessiana  $H(\mathbf{x})$  o for, i.e., se  $(\vec{\mathbf{h}})^t H(\mathbf{x}) \vec{\mathbf{h}} \geq 0$  para todo  $\mathbf{h} \in \mathbb{R}^m$ , e semi-definida negativa se  $(\vec{\mathbf{h}})^t H(\mathbf{x}) \vec{\mathbf{h}} \leq 0$  para todo  $\mathbf{h} \in \mathbb{R}^m$ .

**TEOREMA 10.4.2** (Ponto extremo interior). *Seja  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  e  $\mathbf{x} \in \mathbb{R}^m$  ponto extremo local. Se  $f$  é diferenciável em  $\mathbf{x}$ , então  $\mathbf{x}$  é ponto crítico, i.e.,  $\mathbf{f}'(\mathbf{x}) = 0$ . Se além disto,  $f$  for duas vezes diferenciável com derivadas segundas contínuas, então temos que*

- (1) *se  $\mathbf{x}$  for ponto de mínimo local, então  $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) \geq 0$  para todo  $\mathbf{h} \in \mathbb{R}^m$ ,*
- (2) *se  $\mathbf{x}$  for ponto de máximo local, então  $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) \leq 0$  para todo  $\mathbf{h} \in \mathbb{R}^m$ .*

**DEMONSTRAÇÃO.** Para mostrar que  $\mathbf{x}$  é ponto crítico, basta usar o Teorema 10.3.1 e mostrar que as derivadas parciais se anulam, pois dado o vetor  $\mathbf{e}_i$  temos que a função  $\phi(t) = f(\mathbf{x} + t\mathbf{e}_i)$  tem ponto extremo local em  $t = 0$ . Usando o Teorema 6.3.1 vemos que  $\phi'(0) = 0$ . Mas então

$$0 = \phi'(0) = f'(\mathbf{x})(\mathbf{e}_i) = \frac{\partial f}{\partial x_i}(\mathbf{x})$$

e concluímos que  $\mathbf{f}'(\mathbf{x}) = 0$ .

Suponha agora que  $f$  seja duas vezes diferenciável com derivadas segundas contínuas, e que  $\mathbf{x}$  seja ponto de mínimo local. Então  $\mathbf{x}$  é ponto crítico, como acabamos de mostrar, e pelo Teorema de Taylor em várias dimensões (Teorema 10.4.1), temos que

$$f(\mathbf{x} + s\mathbf{u}) - f(\mathbf{x}) = \frac{s^2}{2} f''(\boldsymbol{\xi}_s)(\mathbf{u}, \mathbf{u}),$$

para todo  $s$  suficientemente pequeno e  $\mathbf{u} \in \mathbb{R}^m$ , onde  $\boldsymbol{\xi}_s$  é ponto do segmento unindo  $\mathbf{x}$  e  $\mathbf{x} + s\mathbf{u}$ . Quando  $s \rightarrow 0$ , temos que  $\boldsymbol{\xi}_s \rightarrow \mathbf{x}$ , e usando a continuidade de  $f''$  concluímos que

$$f''(\mathbf{x})(\mathbf{u}, \mathbf{u}) = \lim_{s \rightarrow 0} f''(\boldsymbol{\xi}_s)(\mathbf{u}, \mathbf{u}) = 2 \lim_{s \rightarrow 0} \frac{f(\mathbf{x} + s\mathbf{u}) - f(\mathbf{x})}{s^2} \geq 0,$$

pois como  $\mathbf{x}$  é mínimo local, então  $f(\mathbf{x} + s\mathbf{u}) - f(\mathbf{x}) \geq 0$  para todo  $s$  suficientemente pequeno. Portanto  $f''(\mathbf{x})(\mathbf{u}, \mathbf{u}) \geq 0$ , como queríamos demonstrar.  $\square$

Os resultados acima nos dão condições necessárias para um ponto interior ser extremo local, porém estas não são suficientes (vide exemplo  $f(x) = x^3$ ). Dizemos que um ponto é *de sela* quando a derivada se anula mas este não é extremo local. Um caso interessante é quando a função é localmente crescente na direção de uma coordenada e decrescente na direção de outra. Por exemplo,  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por  $f(x, y) = x^2 - y^2$ , ver Figura 1. Ver também a sela de macaco dada por  $f(x, y) = x^3 - 3xy^2$ , Figura 2 (tirada de [22]).

O resultado a seguir nos fornece algumas condições suficientes para um ponto ser de máximo, mínimo ou de sela. Mais precisamente, temos que se um ponto crítico  $\mathbf{x}$  de uma função suave tem  $f''(\mathbf{x})$  *positiva definida*, i.e.,  $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) > 0$  para todo  $\mathbf{h} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ , então ele é mínimo estrito local. De forma análoga, se  $f''(\mathbf{x})$  é *negativa definida*, i.e.,  $f''(\mathbf{x})(\mathbf{h}, \mathbf{h}) < 0$  para todo  $\mathbf{h} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ , então ele é máximo estrito local. O último caso é quando  $f''(\mathbf{x})$  é *indefinida* i.e., existem  $\mathbf{h}, \boldsymbol{\xi}$  em  $\mathbb{R}^m$  tais que  $[f''(\mathbf{x})(\mathbf{h}, \mathbf{h})][f''(\mathbf{x})(\boldsymbol{\xi}, \boldsymbol{\xi})] < 0$ . Aí então  $\mathbf{x}$  é ponto de sela.

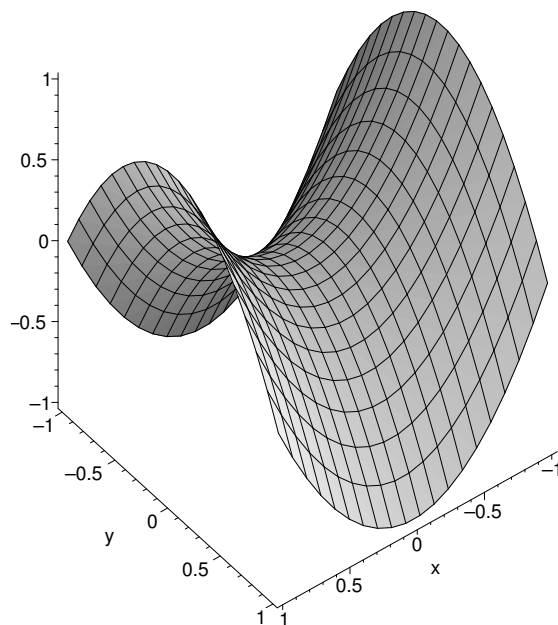


FIGURA 1. Gráfico de  $x^2 - y^2$ , que tem ponto de sela em  $(0, 0)$ .

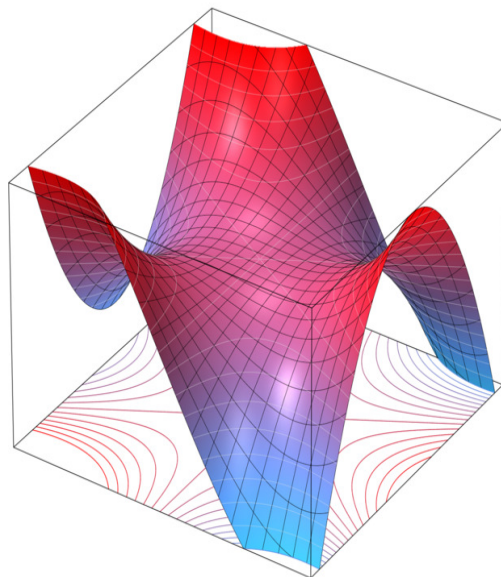


FIGURA 2. Gráfico da sela de macaco dada por  $x^3 - 3xy^2$ , com ponto de sela em  $(0, 0)$ .

**TEOREMA 10.4.3.** *Seja  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  duas vezes diferenciável, com derivadas contínuas, e  $\mathbf{x} \in \mathbb{R}^m$  ponto crítico. Temos então que*

- (1) *se  $f''(\mathbf{x})$  for positiva definida então  $\mathbf{x}$  é mínimo estrito local,*
- (2) *se  $f''(\mathbf{x})$  for negativa definida então  $\mathbf{x}$  é máximo estrito local,*
- (3) *se  $f''(\mathbf{x})$  for indefinida então  $\mathbf{x}$  é ponto de sela.*



**DEMONSTRAÇÃO.** Mostraremos apenas o caso em que  $f''(\mathbf{x})$  é positiva definida. Neste caso, devido à continuidade das segundas derivadas,  $f''(\cdot)$  é positiva definida numa vizinhança aberta de  $\mathbf{x}$ . Para  $\mathbf{y} \neq \mathbf{x}$  satisfazendo as condições do Teorema de Taylor no  $\mathbb{R}^m$  (Teorema 10.4.1), e suficientemente próximo de  $\mathbf{x}$ , temos que existe  $\boldsymbol{\xi}$  pertencente ao segmento de reta entre  $\mathbf{y}$  e  $\mathbf{x}$  e tal que

$$(10.4.3) \quad f(\mathbf{y}) - f(\mathbf{x}) = \frac{1}{2} f''(\boldsymbol{\xi})(\mathbf{y} - \mathbf{x}, \mathbf{y} - \mathbf{x}).$$

Portanto  $\mathbf{x}$  é mínimo estrito local pois a expressão do lado direito de (10.4.3) é estritamente positiva.  $\square$

Note que apesar do teorema anterior dar condições suficientes para determinar se um ponto crítico é ou não extremo local, ainda é preciso descobrir se a  $f''$  é positiva ou negativa definida ou indeterminada. Esta dificuldade é contornável, pois existem vários resultados de álgebra linear que dizem, por exemplo, quando uma matriz é ou não positiva definida. Por exemplo, uma matriz simétrica é positiva definida se e somente se seus autovalores são positivos. A referência [11] apresenta este e vários outros resultados relacionados ao tema.

**EXEMPLO 10.5.** Seja  $F : \mathbb{R}^m \rightarrow \mathbb{R}$  dada por

$$F(\mathbf{x}) = c + \vec{\mathbf{b}}^t \vec{\mathbf{x}} + \frac{1}{2} \vec{\mathbf{x}}^t A \vec{\mathbf{x}},$$

onde  $A \in \mathbb{R}^{m \times m}$  é simétrica positiva definida,  $\vec{\mathbf{b}} \in \mathbb{R}^{m \times 1}$ , e  $c \in \mathbb{R}$ . Então  $\mathbf{x}_*$  é ponto de mínimo estrito de  $F$  se e somente se  $A \vec{\mathbf{x}}_* = -\vec{\mathbf{b}}$ . De fato, se  $\mathbf{x}_*$  é ponto de mínimo estrito de  $F$ , então  $F'(\mathbf{x}_*) = 0$ . Mas a matriz jacobiana  $[F'(\mathbf{x}_*)] \in \mathbb{R}^{m \times 1}$  é dada por

$$[F'(\mathbf{x}_*)] = (\vec{\mathbf{x}}_*)^t A + \vec{\mathbf{b}}^t,$$

e portanto  $A \vec{\mathbf{x}}_* = -\vec{\mathbf{b}}$ . Por outro lado, se  $A \vec{\mathbf{x}}_* = -\vec{\mathbf{b}}$ , então  $F'(\mathbf{x}_*) = 0$ . Como a matriz hessiana de  $F$ , dada por  $A$ , é positiva definida, então  $\mathbf{x}_*$  é ponto de mínimo estrito de  $F$ .

Uma segunda aplicação do Teorema 10.4.1 diz respeito às funções convexas definidas em convexos. Dizemos que  $\Omega \subseteq \mathbb{R}^m$  é convexo se  $\mathbf{x}, \mathbf{y} \in \Omega$  implica em  $(1-t)\mathbf{x} + t\mathbf{y} \in \Omega$  para todo  $t \in [0, 1]$ . Dizemos que  $f : \Omega \rightarrow \mathbb{R}$  é convexa em  $\Omega$  se

$$f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y}).$$

para todo  $t \in [0, 1]$ . Graficamente, uma função é convexa se o gráfico de  $f$  entre  $\mathbf{x}$  e  $\mathbf{y}$  está abaixo da reta que une os pontos  $(\mathbf{x}, f(\mathbf{x}))$  e  $(\mathbf{y}, f(\mathbf{y}))$ , como ilustra a Figura 3.

Existem inúmeros resultados relacionados a convexidade. Em particular, um mínimo local é também global, e se o mínimo local é estrito, segue-se a unicidade de mínimo global [19].

**TEOREMA 10.4.4.** *Seja  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  duas vezes diferenciável, com derivadas contínuas. Então as afirmativas abaixo são equivalentes:*

- (1)  $f$  é convexa
- (2)  $f''(\mathbf{x})$  é semi-definida positiva para todo  $\mathbf{x} \in \mathbb{R}^m$ .

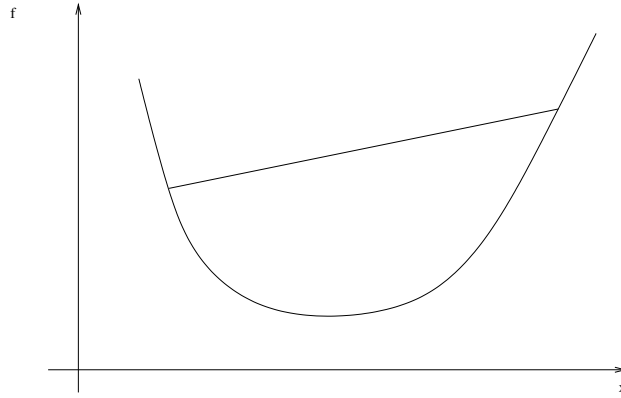


FIGURA 3. Função convexa.

DEMONSTRAÇÃO. ( $\Leftarrow$ ) Suponha que  $f''(\mathbf{x})$  seja semi-definida positiva em  $\mathbb{R}^m$ . Seja  $S$  o segmento de reta unindo  $\mathbf{x}$  e  $\mathbf{y} \in \mathbb{R}^m$ , e seja  $0 < t < 1$ . Definindo  $\mathbf{x}_0 = (1-t)\mathbf{x} + t\mathbf{y}$ , pelo Teorema de Taylor existe  $\boldsymbol{\xi}_1 \in S$  entre  $\mathbf{x}$  e  $\mathbf{x}_0$ , e  $\boldsymbol{\xi}_2 \in S$  entre  $\mathbf{x}_0$  e  $\mathbf{y}$  tais que

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}f''(\boldsymbol{\xi}_1)(\mathbf{x} - \mathbf{x}_0, \mathbf{x} - \mathbf{x}_0), \\ f(\mathbf{y}) &= f(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)(\mathbf{y} - \mathbf{x}_0) + \frac{1}{2}f''(\boldsymbol{\xi}_2)(\mathbf{y} - \mathbf{x}_0, \mathbf{y} - \mathbf{x}_0). \end{aligned}$$

Como  $f''(\boldsymbol{\xi}_1)$  e  $f''(\boldsymbol{\xi}_2)$  são ambas semi-definidas positivas, então

$$\begin{aligned} &(1-t)f(\mathbf{x}) + tf(\mathbf{y}) \\ &= f(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)[(1-t)\mathbf{x} + t\mathbf{y} - \mathbf{x}_0] + \frac{(1-t)}{2}f''(\boldsymbol{\xi}_1)(\mathbf{x} - \mathbf{x}_0, \mathbf{x} - \mathbf{x}_0) + \frac{t}{2}f''(\boldsymbol{\xi}_2)(\mathbf{y} - \mathbf{x}_0, \mathbf{y} - \mathbf{x}_0) \\ &= f(\mathbf{x}_0) + \frac{(1-t)}{2}f''(\boldsymbol{\xi}_1)(\mathbf{x} - \mathbf{x}_0, \mathbf{x} - \mathbf{x}_0) + \frac{t}{2}f''(\boldsymbol{\xi}_2)(\mathbf{y} - \mathbf{x}_0, \mathbf{y} - \mathbf{x}_0) \geq f(\mathbf{x}_0). \end{aligned}$$

Logo  $f$  é convexa.

( $\Rightarrow$ ) Se  $f$  é convexa,

$$f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y})$$

e para  $t \in (0, 1]$  temos que

$$\frac{f((1-t)\mathbf{x} + t\mathbf{y}) - f(\mathbf{x})}{t} \leq f(\mathbf{y}) - f(\mathbf{x}).$$

Tomando o limite  $t \rightarrow 0$  obtemos  $\mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y}) - f(\mathbf{x})$ . Seja  $s = \|\mathbf{x} - \mathbf{y}\|$  e  $\mathbf{h} = (\mathbf{y} - \mathbf{x})/s$ . Usando agora a fórmula de Taylor obtemos que existe  $\hat{s} \in (0, s)$  tal que

$$\frac{1}{2}f''(\mathbf{x} + \hat{s}\mathbf{h})(s\mathbf{h}, s\mathbf{h}) = \frac{1}{2}f''(\mathbf{x} + \hat{s}\mathbf{h})(\mathbf{y} - \mathbf{x}, \mathbf{y} - \mathbf{x}) = f(\mathbf{y}) - f(\mathbf{x}) - \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) \geq 0.$$

Usando a bilinearidade da aplicação  $f''(\mathbf{x} + \hat{s}\mathbf{h})$ , temos

$$f''(\mathbf{x} + \hat{s}\mathbf{h})(\mathbf{h}, \mathbf{h}) \geq 0.$$

para todo  $\mathbf{h} \in B_1(\mathbf{0})$ . Tomando  $\mathbf{y} \rightarrow \mathbf{x}$  temos  $s \rightarrow 0$  e portanto  $\hat{s} \rightarrow 0$ . Usando a continuidade de  $f''$  concluímos a demonstração.  $\square$

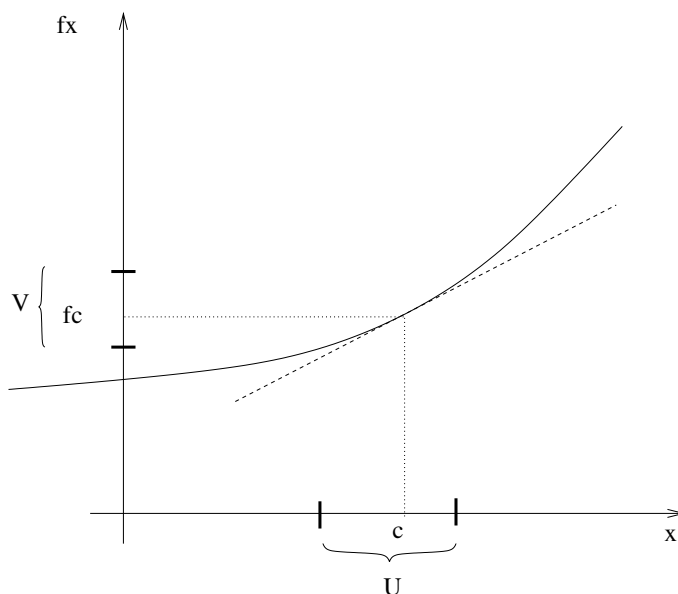


FIGURA 4. Teorema da função inversa.

OBSERVAÇÃO. Note que no processo de demonstração do Teorema 10.4.4, mostramos também que uma função  $f$  ser convexa implica em  $\mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y}) - f(\mathbf{x})$  para todo  $\mathbf{x}, \mathbf{y}$ .

## 10.5. Teorema da Função Inversa e da Função Implícita

**10.5.1. Teorema da Função Inversa.** Como motivação considere primeiro o caso unidimensional, e seja  $f : \mathbb{R} \rightarrow \mathbb{R}$  “suave”. Se  $f'(x) \neq 0$  para algum  $x \in \mathbb{R}$ , então  $f$  é localmente invertível, i.e,  $f$  é injetiva numa vizinhança aberta  $U$  de  $x$  e existe  $g = f^{-1} : V \rightarrow U$ , onde  $U = f(V)$ , tal que

$$g(f(x)) = x, \quad \text{para todo } x \in U.$$

No caso a “suavidade” necessária é que a função tenha derivadas contínuas. Dado  $\Omega \subseteq \mathbb{R}^m$ , dizemos que uma função  $f : \Omega \rightarrow \mathbb{R}^n$  é de classe  $\mathcal{C}^1(\Omega)$  se é diferenciável com derivadas contínuas em  $\Omega$ .

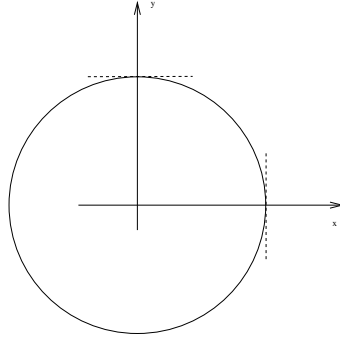
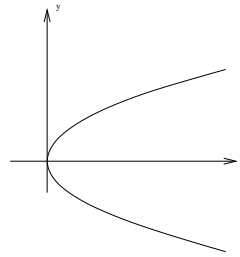
**TEOREMA 10.5.1 (Função Inversa).** *Seja  $\Omega \subseteq \mathbb{R}^m$  e  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$  de classe  $\mathcal{C}^1(\Omega)$ . Seja  $\hat{\mathbf{x}} \in \Omega$  tal que  $D = \mathbf{f}'(\hat{\mathbf{x}})$  é invertível. Então existe uma vizinhança aberta  $U$  de  $\hat{\mathbf{x}}$  tal que*

- (1)  $\mathbf{f} : U \rightarrow V = \mathbf{f}(U)$  é injetiva, e  $V$  é aberto.
- (2) Seja  $\mathbf{g} : V \rightarrow U$  a função inversa de  $\mathbf{f}$  definida por

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x} \quad \text{para todo } \mathbf{x} \in U.$$

Então  $\mathbf{g} \in \mathcal{C}^1(V)$  e para  $\hat{\mathbf{y}} = \mathbf{f}(\hat{\mathbf{x}})$  tem-se  $\mathbf{g}'(\hat{\mathbf{y}}) = [\mathbf{f}'(\hat{\mathbf{x}})]^{-1}$ .

OBSERVAÇÃO. Note que o teorema acima tem caráter *local*. Em particular, é possível construir funções não injetivas em seu domínios que possuem matrizes jacobianas invertíveis em todos os pontos. Entretanto em *uma dimensão*, se a derivada não se anula em nenhum ponto de um intervalo aberto, a função é globalmente invertível.

FIGURA 5. Conjunto  $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ .FIGURA 6. Conjunto  $\{(x, y) \in \mathbb{R}^2 : x = y^2\}$ .

**10.5.2. Teorema da função implícita.** O teorema de função inversa trata da importante questão de solvabilidade de equações dadas de forma implícita. A pergunta é simples: dados os pontos  $(\mathbf{x}, \mathbf{y})$  soluções de uma equação  $F(\mathbf{x}, \mathbf{y}) = 0$ , será que é possível escrever  $\mathbf{y}$  em função de  $\mathbf{x}$ ?

Como uma primeira motivação, considere  $F(x, y) = x^2 + y^2 - 1$ . Então a curva de nível determinada por  $F(x, y) = 0$  é dada pelo círculo de raio unitário, como nos mostra a Figura 5. Seja  $(a, b) \in \mathbb{R}^2$  tal que  $F(a, b) = 0$ . Por exemplo  $(0, 1)$  e  $(-1, 0)$  satisfazem esta condição. Uma pergunta natural é se existe uma função  $\phi$  tal que  $F(x, \phi(x)) = 0$ , e  $\phi(a) = b$ . A resposta é *globalmente, não*. Mas *localmente sim*, se  $\partial F / \partial y(a, b) \neq 0$ .

Um segundo exemplo é dado por  $F(x, y) = x - y^2$ , ver Figura 6. Para se ter  $F(x, \phi(x)) = 0$ , pode-se escolher  $\phi(x) = \sqrt{x}$  ou  $\phi(x) = -\sqrt{x}$ . Entretanto nenhuma das duas funções está definida na vizinhança de  $x = 0$ . Note que  $\partial F / \partial y(0, 0) = 0$ .

Um exemplo final, agora em dimensões maiores. Sejam  $T_1 : \mathbb{R}^m \rightarrow \mathbb{R}^n$  e  $T_2 : \mathbb{R}^n \rightarrow \mathbb{R}^n$  transformações lineares, e  $F : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$  dada por  $F(\mathbf{x}, \mathbf{y}) = T_1\mathbf{x} + T_2\mathbf{y}$ . Então podemos escrever a equação  $F(\mathbf{x}, \mathbf{y}) = 0$  somente em função de  $\mathbf{x}$  se  $T_2$  for invertível. Neste caso temos  $F(\mathbf{x}, -T_2^{-1}T_1\mathbf{x}) = 0$ . Note que se definirmos a aplicação linear  $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  dada por  $L : \mathbf{v} \rightarrow \mathbf{F}'(\mathbf{a}, \mathbf{b})(0, \mathbf{v})$ , teremos  $L = T_2$ . Então a condição de solvabilidade é de  $L$  seja invertível.

**TEOREMA 10.5.2 (Função implícita).** *Seja  $\Omega \subseteq \mathbb{R}^{m+n}$  um aberto, e  $(\mathbf{x}_0, \mathbf{y}_0) \in \Omega$ . Seja  $\mathbf{F} : \Omega \rightarrow \mathbb{R}^n$  de classe  $\mathcal{C}^1(\Omega)$ , e tal que  $\mathbf{F}(\mathbf{x}_0, \mathbf{y}_0) = 0$ . Se a transformação linear de  $\mathbb{R}^n$  em  $\mathbb{R}^n$  definida por  $\mathbf{v} \mapsto \mathbf{F}'(\mathbf{x}_0, \mathbf{y}_0)(0, \mathbf{v})$  for invertível, então existe uma vizinhança aberta  $W$  de  $\mathbf{x}_0$ , e uma única função  $\phi : W \rightarrow \mathbb{R}^n$ , que é  $\mathcal{C}^1(W)$  e tal que  $\mathbf{y}_0 = \phi(\mathbf{x}_0)$  e  $F(\mathbf{x}, \phi(\mathbf{x})) = 0$  para todo  $\mathbf{x} \in W$ .*

**DEMONSTRAÇÃO.** Sem perda de generalidade, suponha  $\mathbf{x}_0 = \mathbf{0}$  e  $\mathbf{y}_0 = \mathbf{0}$ . Seja  $H : \Omega \rightarrow \mathbb{R}^{m+n}$  dada por  $H(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{F}(\mathbf{x}, \mathbf{y}))$ . Então  $H'(\mathbf{0}, \mathbf{0})$  é invertível. Pelo teorema da função inversa (Teorema 10.5.1), existe vizinhança aberta  $U$  de  $(\mathbf{0}, \mathbf{0})$  em  $\mathbb{R}^{m+n}$  tal que  $V = H(U)$  é vizinhança aberta em  $\mathbb{R}^{m+n}$ . Além disto existe  $\Phi : V \rightarrow U$  inversa de  $H$  de classe  $\mathcal{C}^1$ . Escrevendo  $\Phi = (\phi_1, \phi_2)$ , onde  $\phi_1 : V \rightarrow \mathbb{R}^m$  e  $\phi_2 : V \rightarrow \mathbb{R}^n$ , temos

$$(\mathbf{x}, \mathbf{y}) = H \circ \Phi(\mathbf{x}, \mathbf{y}) = H(\phi_1(\mathbf{x}, \mathbf{y}), \phi_2(\mathbf{x}, \mathbf{y})) = (\phi_1(\mathbf{x}, \mathbf{y}), F(\phi_1(\mathbf{x}, \mathbf{y}), \phi_2(\mathbf{x}, \mathbf{y}))).$$

Logo,

$$(10.5.1) \quad \mathbf{x} = \phi_1(\mathbf{x}, \mathbf{y}), \quad \mathbf{y} = F(\mathbf{x}, \phi_2(\mathbf{x}, \mathbf{y})),$$

para todo  $(\mathbf{x}, \mathbf{y}) \in V$ . Então  $W = \{\mathbf{x} \in \mathbb{R}^m : (\mathbf{x}, \mathbf{0}) \in V\}$  é vizinhança aberta de  $\mathbf{x} = \mathbf{0}$  em  $\mathbb{R}^m$ . Definindo  $\phi(\mathbf{x}) = \phi_2(\mathbf{x}, \mathbf{0})$ , temos  $\phi(\mathbf{0}) = \mathbf{0}$ , e segue-se de (10.5.1) que  $F(\mathbf{x}, \phi(\mathbf{x})) = \mathbf{0}$ . Como  $\Phi$  é de classe  $\mathcal{C}^1$ , então  $\phi_2$ , e portanto  $\phi$  também é de classe  $\mathcal{C}^1$ .  $\square$

## 10.6. Minimização com restrições

Para problemas de minimização com restrições, dois importantes resultados nos dão condições suficientes para que um ponto seja extremo. São os teoremas de Lagrange e de Kuhn–Tucker, que apresentamos abaixo.

**10.6.1. Condições de primeira ordem.** Dadas funções reais  $f, g_1, \dots, g_k$  definidas num aberto  $\Omega$  de  $\mathbb{R}^m$ , consideramos o problema de minimizar  $f$  restrita ao conjunto de raízes de  $g_1, \dots, g_k$  em  $\Omega$ . O Teorema de Lagrange nos dá condições necessárias que um candidato a mínimo de tal problema tem que satisfazer.

Começamos com um caso mais simples, de funções definidas no plano e onde há apenas uma restrição. Sejam  $f$  e  $g$  funções do  $\mathbb{R}^2$  em  $\mathbb{R}$ , e suponha que  $\mathbf{x}_* \in \mathbb{R}^2$  é tal que

$$f(\mathbf{x}_*) = \min\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^2, g(\mathbf{x}) = 0\}.$$

Apesar de termos imposto  $\mathbf{x}_*$  como mínimo global, isto é somente uma simplificação. A teoria não se modifica em nada se  $\mathbf{x}_*$  for somente mínimo local.

Suponha agora que  $\nabla g(\mathbf{x}_*) \neq 0$  e considere  $\mathbf{r} : \mathbb{R} \rightarrow \mathbb{R}^2$  uma parametrização para a curva de nível determinada por  $g(x, y) \equiv 0$ , e tal que  $\mathbf{r}(0) = \mathbf{x}_*$ . A existência de tal parametrização (local) pode ser justificada através do teorema da função implícita. Pela regra da cadeia,

$$0 = \frac{d}{dt}g(\mathbf{r}(t)) = \nabla g(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$$

concluimos que  $\nabla g(\mathbf{r}(t))$  é ortogonal a  $\mathbf{r}'(t)$ . De forma análoga, temos que  $f(\mathbf{r}(t))$  atinge seu máximo em  $t = 0$ . Logo,  $df(\mathbf{r}(0))/dt = 0$ , e como

$$\frac{d}{dt}f(\mathbf{r}(t)) = \nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t),$$

temos que  $\nabla f(\mathbf{x}_*)$  é também ortogonal a  $\mathbf{r}'(0)$ . Em duas dimensões, se  $\nabla g(\mathbf{x}_*)$  e  $\nabla f(\mathbf{x}_*)$  são ortogonais a  $\mathbf{r}'(0)$  isto significa que  $\nabla g(\mathbf{x}_*)$  e  $\nabla f(\mathbf{x}_*)$  são colineares, i.e., existe  $\lambda^* \in \mathbb{R}$  tal que

$$\nabla g(\mathbf{x}_*) = \lambda^* \nabla f(\mathbf{x}_*).$$

O caso mais geral, em dimensões maiores e com número arbitrário de restrições, é enunciado no teorema abaixo.

TEOREMA 10.6.1 (Lagrange). *Sejam  $f, g_1, \dots, g_k$  funções reais definidas em  $\mathbb{R}^m$  com derivadas contínuas. Suponha que exista  $\delta > 0$  e  $\mathbf{x}_* \in \mathbb{R}^m$  e tal que*

$$f(\mathbf{x}_*) = \min\{f(\mathbf{x}) : \mathbf{x} \in B_\delta(\mathbf{x}_*) \text{ e } g_1(\mathbf{x}) = \dots = g_k(\mathbf{x}) = 0\}.$$

*Então existem números  $\mu, \lambda_1^*, \dots, \lambda_k^*$  não todos nulos e tais que*

$$(10.6.1) \quad \mu \nabla f(\mathbf{x}_*) = \lambda_1^* \nabla g_1(\mathbf{x}_*) + \dots + \lambda_k^* \nabla g_k(\mathbf{x}_*).$$

*Além disto, se  $\{\nabla g_1(\mathbf{x}_*), \dots, \nabla g_k(\mathbf{x}_*)\}$  é linearmente independente, então pode-se tomar  $\mu = 1$ .*

Os números  $\lambda_1^*, \dots, \lambda_k^*$  acima são conhecidos por *multiplicadores de Lagrange*, e em muitas aplicações têm significado próprio.

Não iremos aqui apresentar uma demonstração do Teorema de Lagrange. Apresentamos, seguindo [7], entretanto alguns argumentos que indicam o porquê do resultado valer. A argumentação é toda baseada na aproximação dada pelo Teorema do Valor Médio:

$$(10.6.2) \quad g_i(\mathbf{x}_* + \mathbf{s}) \approx g_i(\mathbf{x}_*) + \nabla g_i(\mathbf{x}_*) \cdot \mathbf{s},$$

Diremos que  $\mathbf{s} \in \mathbb{R}^m$  é uma *direção factível* se  $\mathbf{x}_* + \mathbf{s}$  satisfizer todas as restrições, i.e.,  $g_i(\mathbf{x}_* + \mathbf{s}) = 0$  para todo  $i = 1, \dots, k$ . Logo, para termos  $\mathbf{s}$  factível, temos  $g_i(\mathbf{x}_* + \mathbf{s}) = g_i(\mathbf{x}_*) = 0$ . Logo, por (10.6.2), impomos

$$(10.6.3) \quad \nabla g_i(\mathbf{x}_*) \cdot \mathbf{s} = 0 \text{ para } i = 1, \dots, k.$$

Note que a derivada direcional da  $f$  em  $\mathbf{x}_*$  na direção factível  $\mathbf{s}$  não pode ser negativa, pois isto indicaria que  $f(\mathbf{x}_* + \mathbf{s}) < f(\mathbf{x}_*)$ , uma contradição com  $\mathbf{x}_*$  ser mínimo local. Logo

$$(10.6.4) \quad \nabla f(\mathbf{x}_*) \cdot \mathbf{s} < 0$$

não pode ocorrer junto com (10.6.3).

A condição necessária e suficiente para que (10.6.3) e (10.6.4) nunca ocorram simultaneamente é que existam  $\lambda_1^*, \dots, \lambda_k^*$  tais que

$$\nabla f(\mathbf{x}_*) = \lambda_1^* \nabla g_1(\mathbf{x}_*) + \dots + \lambda_k^* \nabla g_k(\mathbf{x}_*).$$

É imediato checar que a condição é suficiente. Para ver que é também necessária, basta notar que se  $\nabla f(\mathbf{x}_*)$  não é combinação linear de  $\nabla g_1(\mathbf{x}_*), \dots, \nabla g_k(\mathbf{x}_*)$ , então existem  $\mathbf{u} \in \mathbb{R}^m$  tais que

- $\mathbf{u} \cdot \nabla g_i(\mathbf{x}_*) = 0$  para  $i = 1, \dots, k$
- $\mathbf{v}$  é combinação linear de  $\nabla g_1(\mathbf{x}_*), \dots, \nabla g_k(\mathbf{x}_*)$
- $\nabla f(\mathbf{x}_*) = \mathbf{u} + \mathbf{v}$

Logo,  $\mathbf{s} = -\mathbf{u}$  satisfaz (10.6.3) e (10.6.4), uma contradição.

Note que uma forma compacta de descrever as condições de Lagrange vem através da definição do Lagrangeano  $\mathcal{L} : \mathbb{R}^m \times \mathbb{R}^k \rightarrow \mathbb{R}$  dado por

$$(10.6.5) \quad \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda} \cdot \mathbf{g}(\mathbf{x}),$$

onde  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$  e  $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_k(\mathbf{x})) \in \mathbb{R}^k$ . Então as condições necessárias para  $\mathbf{x}_*$  ser ponto de mínimo são dadas por  $\nabla \mathcal{L}(\mathbf{x}_*, \boldsymbol{\lambda}^*) = 0$ .

Uma outra situação de minimização com restrições ocorre quando as restrições são dadas por desigualdades, e não mais como acima. neste caso temos o Teorema de Kuhn–Tucker, dado abaixo.

**TEOREMA 10.6.2 (Kuhn–Tucker).** *Sejam  $f, h_1, \dots, h_k$  funções reais definidas em  $\mathbb{R}^m$  com derivadas contínuas. Suponha que existam  $\delta > 0$  e  $\mathbf{x}_* \in \mathbb{R}^m$  tais que*

$$f(\mathbf{x}_*) = \min\{f(x) : \mathbf{x} \in B_\delta(\mathbf{x}_*) \text{ e } h_1(\mathbf{x}) \geq 0, \dots, h_k(\mathbf{x}) \geq 0\}.$$

Então as seguintes afirmativas são verdadeiras:

(1) existem números  $\mu, \lambda_1^*, \dots, \lambda_k^*$  não todos nulos e tais que

$$\mu \nabla f(\mathbf{x}_*) = \lambda_1^* \nabla h_1(\mathbf{x}_*) + \dots + \lambda_k^* \nabla h_k(\mathbf{x}_*).$$

(2) seja  $i \in \{1, \dots, k\}$  tal que  $h_i(\mathbf{x}_*) > 0$ . Então pode-se impor  $\lambda_i^* = 0$ .

(3) se conjunto  $V = \{h_i'(\mathbf{x}_*) : h_i(\mathbf{x}_*) = 0, \text{ onde } 1 \leq i \leq k\}$  é linearmente independente, então pode-se tomar  $\mu = 1$  e  $\lambda_1^* \geq 0, \dots, \lambda_k^* \geq 0$ .

Seguiremos aqui apresentar os argumentos que apresentamos justificando o Teorema de Lagrange, seguindo novamente [7]. Podemos supor, sem perda de generalidade, que  $h_i(x_*) = 0$  para todo  $i = 1, \dots, k$ . De fato, se  $h_i(x_*) > 0$ , então podemos tomar uma vizinhança de  $\mathbf{x}_*$  tal que  $h_i$  seja estritamente positiva nesta vizinhança (pois  $h_i$  é positiva). É este o motivo podemos tomar  $\lambda_i^* = 0$  nestes casos.

Se  $\mathbf{s}$  é direção factível, i.e., satisfaz as desigualdades, então,  $h_i(\mathbf{x}_* + \mathbf{s}) \geq 0$ . Logo, como  $h_i(\mathbf{x}_*) = 0$ , por (10.6.2), temos

$$(10.6.6) \quad \nabla h_i(\mathbf{x}_*) \cdot \mathbf{s} \geq 0 \text{ para } i = 1, \dots, k.$$

Desde que a derivada direcional da  $f$  em  $\mathbf{x}_*$  na direção factível  $\mathbf{s}$  não pode ser negativa, temos que

$$(10.6.7) \quad \nabla f(\mathbf{x}_*) \cdot \mathbf{s} < 0$$

não pode ocorrer junto com (10.6.6).

A condição necessária e suficiente para que (10.6.6) e (10.6.7) nunca ocorram simultaneamente é que existam  $\lambda_1^* \geq 0, \dots, \lambda_k^* \geq 0$  tais que

$$\nabla f(\mathbf{x}_*) = \lambda_1^* \nabla h_1(\mathbf{x}_*) + \dots + \lambda_k^* \nabla h_k(\mathbf{x}_*).$$

Novamente, é fácil checar que a condição é suficiente. Para ver que  $\nabla f(\mathbf{x}_*)$  é combinação linear de  $\nabla h_1(\mathbf{x}_*), \dots, \nabla h_k(\mathbf{x}_*)$ , basta argumentar como no caso de restrição estrita. Sem perda de generalidade suponha agora que  $\lambda_1^* < 0$ . Seja  $P$  o espaço vetorial formado pelas combinações lineares de  $\{\nabla h_2(\mathbf{x}_*), \dots, \nabla h_k(\mathbf{x}_*)\}$ . Seja agora  $\mathbf{s}$  vetor ortogonal ao subespaço  $P$ , e “apontando na mesma direção que  $\nabla h_1(\mathbf{x}_*)$ ”, i.e., tal que  $\nabla h_1(\mathbf{x}_*) \cdot \mathbf{s} \geq 0$ . Por construção temos então que (10.6.6) e (10.6.7) ocorrem com tal  $\mathbf{s}$ , uma contradição. Logo temos sempre  $\lambda_i^* \geq 0$ . Note que no argumento acima usamos que  $\nabla h_1(\mathbf{x}_*)$  é linearmente independente de  $\nabla h_2(\mathbf{x}_*), \dots, \nabla h_k(\mathbf{x}_*)$  (caso contrário  $\nabla h_1(\mathbf{x}_*) \in P$ ). Esta hipótese de independência linear é um artifício desta demonstração, e pode ser eliminada [7].

**10.6.2. Condições de segunda ordem.** Considere o caso de restrições de igualdade somente, como no Teorema 10.6.1. Temos por Taylor, para  $\mathbf{s}$  admissível, que

$$(10.6.8) \quad f(\mathbf{x}_* + \mathbf{s}) = \mathcal{L}(\mathbf{x}_* + \mathbf{s}, \boldsymbol{\lambda}^*) \approx \mathcal{L}(\mathbf{x}_*, \boldsymbol{\lambda}^*) + \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}_*, \boldsymbol{\lambda}^*) \cdot \mathbf{s} + \frac{1}{2} \mathbf{s} \cdot W^* \mathbf{s} = f(\mathbf{x}_*) + \frac{1}{2} \mathbf{s} \cdot W^* \mathbf{s}$$

onde  $\mathcal{L}$  foi definida em (10.6.1), e  $W^*$  é a Hessiana de  $\mathcal{L}$  em relação a  $\mathbf{x}$ , i.e.,

$$\mathbf{s} \cdot W^* \mathbf{s} = \sum_{i,j=1}^m \frac{\partial^2 \mathcal{L}(\mathbf{x}_*, \boldsymbol{\lambda}^*)}{\partial x_i \partial x_j} s_i s_j.$$

Como  $\mathbf{x}_*$  é mínimo, obtemos de (10.6.8) a condição necessária  $\mathbf{s} \cdot W^* \mathbf{s} \geq 0$  para todo  $\mathbf{s}$  admissível, i.e., todo  $\mathbf{s}$  tal que satisfaça (10.6.3). Similarmente, de (10.6.8) obtemos a condição suficiente para que  $\mathbf{x}_*$  seja mínimo estrito, que é

$$\mathbf{s} \cdot W^* \mathbf{s} > 0$$

para todo  $\mathbf{s}$  admissível.

Para o caso de restrições com desigualdade, seja  $\mathbf{s}$  admissível, i.e., tal que  $h_i(\mathbf{x}_* + \mathbf{s}) \geq 0$ . Então definimos agora

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda} \cdot \mathbf{h}(\mathbf{x}),$$

onde  $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_k(\mathbf{x})) \in \mathbb{R}^m$ . Portanto

(10.6.9)

$$\begin{aligned} f(\mathbf{x}_* + \mathbf{s}) &= \mathcal{L}(\mathbf{x}_* + \mathbf{s}, \boldsymbol{\lambda}^*) + \boldsymbol{\lambda}^* \cdot \mathbf{h}(\mathbf{x}_* + \mathbf{s}) \approx \mathcal{L}(\mathbf{x}_*, \boldsymbol{\lambda}^*) + \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}_*, \boldsymbol{\lambda}^*) \cdot \mathbf{s} + \frac{1}{2} \mathbf{s} \cdot W^* \mathbf{s} + \boldsymbol{\lambda}^* \cdot \mathbf{h}(\mathbf{x}_* + \mathbf{s}) \\ &= f(\mathbf{x}_*) + \boldsymbol{\lambda}^* \cdot \mathbf{h}(\mathbf{x}_*) + \frac{1}{2} \mathbf{s} \cdot W^* \mathbf{s} + \sum_{i=1}^k \lambda_i^* h_i(\mathbf{x}_* + \mathbf{s}). \end{aligned}$$

Como  $\lambda_i^* \geq 0$ , podemos definir  $\mathcal{A}_+ = \{i \text{ inteiro} : \lambda_i^* > 0, 1 \leq i \leq k\}$  e  $\mathcal{A}_0 = \{i \text{ inteiro} : \lambda_i = 0, 1 \leq i \leq k\}$ . Logo, usando que  $h_i(\mathbf{x}_*) = 0$ ,

$$f(\mathbf{x}_* + \mathbf{s}) - f(\mathbf{x}_*) \approx \frac{1}{2} \mathbf{s} \cdot W^* \mathbf{s} + \sum_{i \in \mathcal{A}_+} \lambda_i^* \nabla h_i(\mathbf{x}_*) \cdot \mathbf{s},$$

Buscamos então condições suficientes para que o lado direito da equação acima seja positivo. Se  $\mathbf{s}$  é factível, então vale (10.6.6). Se  $\nabla h_i(\mathbf{x}_*) \cdot \mathbf{s} > 0$ , para  $\mathbf{s}$  “pequeno o suficiente” temos que

$$\frac{1}{2} \mathbf{s} \cdot W^* \mathbf{s} + \sum_{i \in \mathcal{A}_+} \lambda_i^* \nabla h_i(\mathbf{x}_*) \cdot \mathbf{s} > 0$$

e  $f(\mathbf{x}_* + \mathbf{s}) > f(\mathbf{x}_*)$ . Entretanto, se  $\nabla h_i(\mathbf{x}_*) \cdot \mathbf{s} = 0$ , uma condição suficiente para que  $f(\mathbf{x}_* + \mathbf{s}) \geq f(\mathbf{x}_*)$  é exigir que  $\mathbf{s} \cdot W^* \mathbf{s} > 0$  para todo  $\mathbf{s}$  admissível e ortogonal a  $\nabla h_i(\mathbf{x}_*)$  para  $i \in \mathcal{A}_+$ .



### 10.7. Exercícios

EXERCÍCIO 10.1. Extenda para  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$  o conceito de limite de funções num determinado ponto  $\mathbf{x}_0 \in \mathbb{R}^m$ .

EXERCÍCIO 10.2. Seja  $f$  e  $g$  funções de  $\mathbb{R}^m \rightarrow \mathbb{R}^n$ , diferenciáveis em  $\mathbf{x} \in \mathbb{R}^m$ . Mostre usando a definição de derivadas que  $(f + g)'(\mathbf{x}) = f'(\mathbf{x}) + g'(\mathbf{x})$ . Seja  $h : \mathbb{R}^m \rightarrow \mathbb{R}$  também diferenciável em  $\mathbf{x} \in \mathbb{R}^m$ . Mostre usando a definição de derivadas que  $(hg)'(\mathbf{x}) = h'(\mathbf{x})g(\mathbf{x}) + h(\mathbf{x})g'(\mathbf{x})$ .

EXERCÍCIO 10.3. Sejam  $a < b$  números reais, e  $f : [a, b] \rightarrow \mathbb{R}$  contínua em  $[a, b]$  e diferenciável em  $(a, b)$ . Mostre que entre duas raízes consecutivas de  $f'$  existe no máximo uma raiz de  $f$ .

EXERCÍCIO 10.4. Seja  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4} & \text{para } (x, y) \neq (0, 0), \\ 0 & \text{para } (x, y) = (0, 0). \end{cases}$$

Mostre que a derivada direcional de  $f$  em  $(0, 0)$  com respeito a  $\mathbf{u} = (a, b)$  existe e que

$$D_{\mathbf{u}}f(0, 0) = \frac{b^2}{a}, \quad \text{se } a \neq 0.$$

Mostre que  $f$  não é contínua e portanto não é diferenciável no  $(0, 0)$ .

EXERCÍCIO 10.5 (Kevmasan [18], Example 1.1.1). Mostre que  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por

$$f(x, y) = \begin{cases} \frac{x^5}{(y - x^2) + x^4} & \text{para } (x, y) \neq (0, 0), \\ 0 & \text{para } (x, y) = (0, 0). \end{cases}$$

tem todas as derivadas direcionais em  $(0, 0)$  iguais a zero, mas que  $f$  não é diferenciável no  $(0, 0)$ . (Dica: considere  $\mathbf{h} = (h, h^2)$  em (10.3.1)).

EXERCÍCIO 10.6. Seja  $Q = (0, 1) \times (0, 1)$ . Suponha que  $f : Q \rightarrow \mathbb{R}$ , e  $g : Q \rightarrow \mathbb{R}$  sejam diferenciáveis em  $Q$ . Mostre que se  $\mathbf{f}'(x) = \mathbf{g}'(x)$  para todo  $\mathbf{x} \in Q$ , então existe constante  $c$  tal que  $f(\mathbf{x}) = g(\mathbf{x}) + c$  para todo  $\mathbf{x} \in Q$ .

EXERCÍCIO 10.7. Seja  $\Omega \subset \mathbb{R}^n$ , aberto e com a seguinte propriedade: existe  $\mathbf{x}_* \in \Omega$  tal que para todo  $\mathbf{x} \in \Omega$ , a reta  $S_{\mathbf{x}} = \{t\mathbf{x} + (1-t)\mathbf{x}_* : t \in [0, 1]\}$  está contida em  $\Omega$ , i.e.,  $S_{\mathbf{x}} \subset \Omega$ . Seja  $f : \Omega \rightarrow \mathbb{R}$  função diferenciável em  $\Omega$  e tal que todas as derivadas parciais de  $f(\mathbf{x})$  são nulas, para todo  $\mathbf{x} \in \Omega$ . Mostre que  $f$  é constante.

EXERCÍCIO 10.8. Seja  $B = \{\mathbf{x} \in \mathbb{R}^m : \|\mathbf{x}\| \leq 1\}$  e  $f : B \rightarrow \mathbb{R}$  função contínua em  $B$ , diferenciável no interior de  $B$  e tal que  $f \equiv 0$  na fronteira de  $B$ . Mostre que  $f$  tem ponto crítico no interior de  $B$ .

EXERCÍCIO 10.9 (Mínimos Quadrados). Considere para  $i = 1, \dots, n$  os pontos  $(x_i, y_i) \in \mathbb{R}^2$ , e seja  $p : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $p(x) = ax^2 + bx + c$  tal que  $a, b$  e  $c$  minimizam o erro

$\sum_{i=1}^n |p(x_i) - y_i|^2$ . Mostre que  $a$ ,  $b$  e  $c$  satisfazem as equações

$$\begin{aligned} a \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n x_i^2 y_i, \\ a \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i &= \sum_{i=1}^n x_i y_i, \\ a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + cn &= \sum_{i=1}^n y_i. \end{aligned}$$

EXERCÍCIO 10.10. Seja  $A \subset \mathbb{R}^m$  compacto, e  $A^\circ$  o conjunto dos pontos interiores de  $A$ . Seja  $f : A \rightarrow \mathbb{R}$  duas vezes diferenciável, com derivadas contínuas, em  $A^\circ$ . Suponha ainda que  $f$  se anule em toda a fronteira de  $A$ , e que  $f''$  seja negativa definida para todo ponto em  $A^\circ$ . Mostre que  $f(\mathbf{x}) > 0$  para todo  $\mathbf{x} \in A^\circ$ .

EXERCÍCIO 10.11. Mostre, usando o Teorema 10.4.3, que  $(0, 0)$  é ponto de sela de  $f(x, y) = x^2 - y^2$ , e ponto de mínimo estrito local de  $f(x, y) = x^2 + y^2$ .

EXERCÍCIO 10.12. Sejam as funções  $f : \mathbb{R} \rightarrow \mathbb{R}$  e  $g : \mathbb{R} \rightarrow \mathbb{R}$  duas vezes diferenciáveis, com as segundas derivadas contínuas. Suponha que o zero seja ponto de mínimo estrito de  $f$  e  $g$ , e que  $f(0) = g(0) = 1$ . Seja  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por  $\varphi(x, y) = f(x)g(y)$ . O que podemos afirmar sobre a Hessiana de  $\varphi$  em  $(0, 0)$  (nada pode ser afirmado, ela é indefinida, positiva definida, positiva semi-definida, negativa definida, etc)? O que podemos afirmar sobre o ponto  $(0, 0)$  em relação à  $\varphi$  (nada pode ser afirmado, é ponto de máximo, de máximo estrito, de mínimo, de mínimo estrito, de sela, etc)? Justifique suas respostas.

EXERCÍCIO 10.13. Seja  $\mathbf{f} : \mathbb{R} \rightarrow \mathbb{R}^3$  diferenciável e tal que  $\|\mathbf{f}(t)\| = 1$  para todo  $t \in \mathbb{R}$ . Mostre então que  $\mathbf{f}'(t) \cdot \mathbf{f}(t) = 0$ . O vetor  $\mathbf{f}'(t)$  é o *vetor tangente da curva  $\mathbf{f}$  em  $t$* .

EXERCÍCIO 10.14. Seja  $\Omega \subseteq \mathbb{R}^m$  aberto e  $f : \Omega \rightarrow \mathbb{R}$  diferenciável em  $\mathbf{x} \in \Omega$ . Seja  $\nabla f(\mathbf{x}) = (\partial f / \partial x_1, \dots, \partial f / \partial x_m)(\mathbf{x}) \in \mathbb{R}^m$ . Supondo que  $\mathbf{x}$  não é ponto crítico de  $f$ , mostre que a derivada direcional  $D_{\mathbf{u}}f(\mathbf{x})$  atinge seu máximo quando  $\mathbf{u} = c\nabla f(\mathbf{x})$  para algum  $c > 0$ . O vetor  $\nabla f$  é chamado de *vetor gradiente de  $f$* , e dá a direção de “maior crescimento” da função  $f$  no ponto  $\mathbf{x}$ .

## APÊNDICE A

### Uma introdução não tão formal aos fundamentos da matemática

1

A matemática se baseia na argumentação lógica. Outras áreas do conhecimento, talvez todas, podem também reclamar para si tal propriedade, Entretanto a matemática é o *próprio* desenvolvimento da argumentação formal, é a “lógica aplicada.”

Este aspecto da matemática tem consequências interessantes; seus resultados independem da época, cultura e região em que foram gerados. O Teorema de Pitágoras, demonstrado por fanáticos matemáticos (os pitagóricos), cerca de 500 A.C., será válido em qualquer lugar e época (<http://mathworld.wolfram.com/PythagoreanTheorem.html>).

Outras áreas têm teorias “exatas” que são na verdade aproximações da realidade, com “validade” somente sob determinadas condições (por exemplo, teoria da relatividade versus física quântica). Mesmo certas definições podem mudar. Como exemplo, em 1997 a unidade de tempo *segundo* foi definida mais uma vez (<http://en.wikipedia.org/wiki/Second>). Quanto ao pobre quilograma, bem, este ainda busca uma definição adequada aos nossos tempos (<http://en.wikipedia.org/wiki/Kilogram>).

Parece-me desnecessário comentar sobre a volatilidade de várias teorias econômicas. . .

Nestes rápidos comentários que seguem, pretendo passear por alguns aspectos de como a matemática funciona. Uma ótima referência é o livro do Terence Tao [25].

#### A.1. Argumentação formal

**A.1.1. Afirmativas.** Como funciona a argumentação formal na prática? Objetos fundamentais são as afirmativas (ou afirmações ou expressões lógicas), que sempre são verdadeiras ou falsas, mas nunca verdadeiras e falsas simultaneamente. Por exemplo<sup>2</sup>

$$(A.1.1) \quad 1 + 1 = 2,$$

$$(A.1.2) \quad 1 = 2.$$

Vou me adiantar afirmando que (A.1.1) é verdadeira e (A.1.2) é falsa. Esperando que o leitor já tenha se recuperado da surpresa, cabe aqui comentar que frases sem sentido como

$$= 1 + 3-$$

não são afirmativas. Expressões do tipo  $3+1$  também não. Uma regra usual é que afirmativas têm verbos.

Afirmativas podem ser combinadas com “ou” e “e” gerando outras. Por exemplo, se  $a$  é um número real qualquer, então a afirmativa ( $a > 0$  ou  $a \leq 0$ ) é verdadeira, mas ( $a > 0$  e  $a \leq 0$ ) não o é. A regra geral é que se  $X$  e  $Y$  são afirmativas, então  $(X \text{ e } Y)$  só é verdadeira

---

<sup>1</sup>Última Atualização: 09/01/2008

<sup>2</sup>Suponho, por enquanto, que as propriedades de conjuntos e dos números reais são conhecidas

se  $X$  e  $Y$  forem *ambas* verdadeiras. Similarmente,  $(X \text{ ou } Y)$  só é falsa se  $X$  e  $Y$  forem *ambas* falsas. Note que se apenas uma das afirmativas for verdadeira,  $(X \text{ ou } Y)$  é verdadeira. Note que esta noção pode diferir de um possível uso corriqueiro do *ou*, como na frase *ou eu, ou ele ficamos*. Neste caso quer-se dizer que ou eu fico, ou ele fica, mas não ambos — este é o chamado *ou exclusivo*.<sup>3</sup>

Podemos também negar uma afirmativa. Se  $X$  é uma afirmativa verdadeira, então (*não*  $X$ ) é falsa. Da mesma forma, se  $Y$  é uma afirmativa falsa, então (*não*  $Y$ ) é verdadeira. Negar uma afirmativa pode ser útil pois para concluir que uma afirmativa  $Z$  é falsa, as vezes é mais fácil provar que (*não*  $Z$ ) é verdadeira.

Seguramente, este papo poderia ir bem mais longe com a álgebra de Boole ou booleana ([http://en.wikipedia.org/wiki/Boolean\\_algebra](http://en.wikipedia.org/wiki/Boolean_algebra)).

**A.1.2. Implicações.** Os passos de uma argumentação matemática são dados via implicações. Se de um fato conhecido, por exemplo uma afirmativa verdadeira  $X$ , eu posso concluir uma afirmativa verdadeira  $Y$ , então eu escrevo

$$(A.1.3) \quad X \implies Y,$$

e leio *X implica Y*, ou ainda *se X então Y*. Por exemplo

$$(A.1.4) \quad a > 0 \implies 2a > 0.$$

Abstraindo um pouco mais, note que (A.1.3) e (A.1.4) também são afirmativas. Outros exemplos de afirmativas:

$$(A.1.5) \quad 0 = 0 \implies 0 = 0,$$

$$(A.1.6) \quad 0 = 1 \implies 0 = 0,$$

$$(A.1.7) \quad 0 = 1 \implies 0 = 1,$$

$$(A.1.8) \quad 0 = 0 \implies 0 = 1.$$

As três primeiras afirmativas acima são verdadeiras. Somente a última é falsa. A primeira da lista é uma tautologia (redundância, do grego *tauto*, o mesmo), e é obviamente correta. Já a segunda é correta pois de hipóteses falsas pode-se concluir verdades (multiplique ambos os lados de (A.1.6) por zero). A terceira é verdade pois se a hipótese é verdadeira, a conclusão, sendo uma mera repetição da hipótese, também o é (este tipo de argumento é usado em demonstrações por contradição). Finalmente, (A.1.8) é falsa pois não se pode deduzir uma afirmativa verdadeira partindo-se de uma falsa.

A argumentação (e a demonstração) matemática baseia-se em supor que algumas hipóteses são verdadeiras e em concluir resultados através de implicações.

Note que a implicação não é “reversível”, i.e., se  $X \implies Y$ , não podemos concluir que  $Y \implies X$ . Realmente,  $x = -1 \implies x^2 = 1$ , mas  $x^2 = 1 \not\implies x = -1$  (esta seta cortada é o símbolo de *não implica*), ou seja, não se pode concluir se  $x = -1$  ou não a partir da hipótese  $x^2 = 1$ .

As vezes, tanto a implicação como seu reverso valem. Se por exemplo  $X \implies Y$  e  $Y \implies X$  escrevemos simplesmente  $X \iff Y$ , e lemos *X se e somente se Y*.

<sup>3</sup>Outro termo matemático que pode ter sentido diferente do uso diário é *em geral*. Na matemática, em geral quer dizer *sempre*, enquanto no dia-a-dia quer dizer “quase sempre”

**A.1.3. Axiomas.** E como começar a construção da matemática em si, i.e., quais são as hipóteses *básicas* que são necessariamente verdadeiras? Iso é importante pois, como vimos, partindo-se de hipóteses falsas pode-se chegar a conclusões falsas, *sem comprometer a lógica*. Aqui entram os *axiomas*, premissas verdadeiras consideradas “óbvias.” É uma boa idéia que este conjunto de premissas seja o menor possível, i.e., um axioma do conjunto não pode ser demonstrada a partir dos outros.

A partir dos axiomas contrói-se via implicações toda uma matemática (mudando-se o conjunto de axiomas, muda-se a matemática).

Um exemplo de axioma vem a seguir.

AXIOMA A.1.1 (do conjunto vazio). Existe um conjunto que não contém nenhum elemento.

Suponha que se possa definir o que é uma pessoa careca, e considere o seguinte axioma.

AXIOMA A.1.2 (do fio extra). Um careca que ganhar um fio extra de cabelo continua careca.

Pode-se concluir então o seguinte resultado (tente demonstrá-lo).

*Se o Axioma do fio extra vale, então todos os seres humanos são carecas.*

O alerta que o resultado acima nos fornece é que devemos ter cuidado com os axiomas escolhidos. Resultados “patológicos” podem advir deles. E de fato, resultados “estranhos” permeiam a matemática. . .

**A.1.4. Definições, lemas, teoremas.** Uma das formas de se construir novos objetos matemáticos é através de *definições*. Por exemplo podemos definir o conjunto dos números naturais como  $\mathbb{N} = \{1, 2, 3, \dots\}$ <sup>4</sup>. Outro exemplo: seja

$$\begin{aligned} f : \mathbb{Z} &\rightarrow \mathbb{R} \\ x &\mapsto x^2. \end{aligned}$$

A expressão acima define uma função chamada “f” que associa a cada número inteiro o seu quadrado, levando-o nos reais.

E quanto a proposições dadas por lemas e teoremas<sup>5</sup>? Normalmente, lemas e teoremas são escritos à parte, sendo compostos por hipóteses, e conclusões explicitamente mencionadas.

Exemplos de lema e teorema vêm a seguir.

LEMA A.1.3. Supondo que o Axioma do conjunto vazio vale, então existe somente um conjunto vazio.

<sup>4</sup>Alguns autores utilizam o símbolo  $:=$  no lugar de  $=$  em definições. Esta é provavelmente uma boa idéia pouco utilizada, e eu não a seguirei.

<sup>5</sup>Uma dúvida comum: qual a diferença entre os três? Bom, normalmente *proposição* tem um caráter mais geral, sendo uma sentença lógica verdadeira (na matemática “usual”). Já um *lema* é proposição preliminar, que contribui na demonstração de um resultado principal, um *teorema*. Muitas vezes entretanto, o lema tem interesse próprio. Em geral, o gosto e o estilo do autor determinam o que é proposição, lema ou teorema.

**TEOREMA A.1.4** (de Fermat). <sup>6</sup> *Seja  $n \in \mathbb{N}$ , com  $n > 2$ . Então não existem inteiros positivos  $x, y, z$  tais que  $x^n + y^n = z^n$ .*

A hipótese do lema A.1.3 é o axioma do conjunto vazio (Axioma A.1.1), e a conclusão é de que só existe um conjunto vazio, isto é todos os conjuntos vazios são iguais. Este é um típico resultado de *unicidade*. Já no Teorema de Fermat A.1.4, impondo-se hipóteses sobre a potência  $n$  (ser inteiro e maior que dois), obtém-se um resultado de *não existência*.

Normalmente lemas e teoremas descrevem resultados de interesse e não triviais, i.e., as conclusões não se seguem trivialmente das hipóteses. Algumas vezes entretanto casos importantes particulares são facilmente obtidos de resultados mais gerais. Estes casos particulares são chamados de *corolários*. O Teorema de Fermat por exemplo é um corolário de um outro resultado mais poderoso (chamado Teorema da Modularidade). É claro que “trivialidade” não é um conceito rigoroso e é certamente relativa.

**A.1.5. Prova ou demonstração.** Uma *prova* ou *demonstração* são os passos lógicos para se concluir uma proposição. Algumas demonstrações são simples, outras nem tanto. Por exemplo, a demonstração por Andrew Wiles do Teorema de Fermat fechou com chave de ouro a matemática do século XX. A prova é uma intrincada sequência de resultados publicada num artigo de 109 páginas na mais conceituada revista de matemática, os Anais de Matemática de Princeton [28].

Antes da demonstração de Wiles, o agora “Teorema de Fermat” era “somente” uma conjectura, um resultado que acredita-se verdadeiro mas que ninguém demonstrou. Uma ainda conjectura famosa é a de Goldbach, que afirma que *todo inteiro par maior que dois pode ser escrito como a soma de dois números primos*. Para números menores que  $10^{18}$ , o resultado foi checado computacionalmente, mas o caso geral ainda não está provado.

## A.2. Demonstração por indução e contradição

Primeiro revemos aqui, através de um exemplo, como é possível demonstrar alguns fatos usando argumentos indutivos.

Considere a afirmativa

$$(A.2.1) \quad \sum_{i=1}^n i = \frac{n}{2}(n+1)$$

para todo  $n \in \mathbb{N}$ .

Para demonstrar que (A.2.1) vale para todos os inteiros positivos, começamos observando que para  $n = 1$ , a afirmativa é obviamente verdadeira. Assuma então que (A.2.1) seja verdade para  $n = N^*$ , i.e,

$$(A.2.2) \quad \sum_{i=1}^{N^*} i = \frac{N^*}{2}(N^* + 1).$$

---

<sup>6</sup>Enunciado de Fermat, na margem do livro *Arithmetica* de Diophantus: *Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos eiusdem nominis fas est dividere cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.* (É impossível separar um cubo em dois cubos, ou a quarta potência em quartas potências, ou em geral qualquer potência em duas potências iguais. Eu descobri uma demonstração realmente maravilhosa disto, para a qual esta margem é por demais exígua para caber.)

Para  $n = N^* + 1$  temos

$$\sum_{i=1}^{N^*+1} i = N^* + 1 + \sum_{i=1}^{N^*} i.$$

Usamos a hipótese indutiva (A.2.2) obtemos

$$\sum_{i=1}^{N^*+1} i = N^* + 1 + \frac{N^*}{2}(N^* + 1) = \frac{N^* + 1}{2}(N^* + 2),$$

e podemos concluir que (A.2.1) vale para  $n = N^* + 1$ , e portanto vale para todos os inteiros positivos.

Um dos passos fundamentais, e algumas vezes esquecido, da demonstração por indução é mostrar que o resultado vale para algum valor inicial (na demonstração acima,  $n = 1$ ). De fato, sem isto, podemos erroneamente “provar” que

$$(A.2.3) \quad 2n \text{ é sempre ímpar para todo } n \in \mathbb{N},$$

com uma argumentação obviamente falsa. De fato supondo que  $2N^*$  é ímpar, temos que  $2(N^* + 1) = 2N^* + 2$  também é pois  $2N^*$  é ímpar por hipótese, e somando 2 a um ímpar obtemos um ímpar. O problema desta demonstração é que não se mostrou (A.2.3) para nenhum número natural.

A demonstração por contradição segue os seguintes princípios lógicos: se queremos mostrar que uma afirmativa implica noutra, podemos simplesmente negar este fato e tentar chegar numa contradição. Considere a afirmativa

$$(A.2.4) \quad \emptyset \subseteq A \quad \text{para qualquer conjunto } A.$$

Talvez uma demonstração “direta” não seja tão fácil. Mas suponha que (A.2.4) seja falso. Então existe algum conjunto  $A$  tal que  $\emptyset \not\subseteq A$ . Portanto existe algum elemento no conjunto vazio que não está em  $A$ . Mas isto é um absurdo, pois o vazio não contém nenhum elemento. O que se vemos é que negar (A.2.4) (afirmar que (A.2.4) é falso) nos leva a concluir um absurdo, e portanto (A.2.4) só pode ser verdade.

### A.3. Exercícios

EXERCÍCIO A.1. Mostre por indução que  $n < 2^n$  para todo  $n \in \mathbb{N}$ .

EXERCÍCIO A.2. Prove que, para todo inteiro  $n > 1$  tem-se que

$$1 + \sum_{i=2}^n \frac{1}{\sqrt{i}} = 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} > \sqrt{n}.$$

EXERCÍCIO A.3. Mostre por indução a desigualdade de Bernoulli: se  $x > -1$ , então  $(1 + x)^n \geq 1 + nx$  para todo  $n \in \mathbb{N}$ .

EXERCÍCIO A.4. Mostre usando contradição que  $\sqrt{2}$  não é racional.

EXERCÍCIO A.5. Mostre usando contradição que se  $p_1, \dots, p_n$  são todos os números primos menores ou iguais a  $p_n$ , então  $p_1 \times \cdots \times p_n + 1$  não é divisível por  $p_i$  para nenhum  $i \in \{1, \dots, n\}$ .

EXERCÍCIO A.6. Mostre usando contradição que existem infinitos números primos.

EXERCÍCIO A.7. Usando indução, mostre que existe  $J \in \mathbb{N}$  tal que  $j^2 - 10j > 0$  para todo inteiro  $j > J$ .

EXERCÍCIO A.8. Seja  $\lambda < 1$  e  $n \in \mathbb{N}$ . Mostre que

$$\sum_{i=n}^k \lambda^i = \lambda^n \frac{1 - \lambda^{k-n+1}}{1 - \lambda}$$

para todo inteiro  $k \geq n$ .



## APÊNDICE B

### Soluções das provas da ANPEC

---

Estas soluções foram submetidas por excelentes alunos, mas eu não conferi a exatidão de todas elas.

---

1

#### (2007 - Questão 1)

Seja  $A$  a matriz na base canônica do operador linear  $L: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  dado por

$$L(x, y, z) = (x + 2y + 3z, 4x + 5y + 6z, 7x + 8y + 9z).$$

Denote por  $\lambda_1, \lambda_2, \lambda_3$  os autovalores da matriz  $A$ . Julgue os itens abaixo:

- ① O posto de  $A$  é 2.
- ①  $L(1, -2, 1) = (0, 0, 0)$ .
- ②  $\lambda_1 \lambda_2 \lambda_3 \neq 0$ .
- ③  $\lambda_1 + \lambda_2 + \lambda_3 = 15$ .
- ④  $L$  é diagonalizável.

#### Respostas

- ① Verdadeiro. O posto é maior ou igual a 2 pois  $\{(1, 2, 3), (4, 5, 6)\}$  é LI, mas não pode ser 3 já que  $(7, 8, 9) = -(1, 2, 3) + 2(4, 5, 6)$ . Logo o posto de  $A$  é de fato 2.
- ① Verdadeiro. Verificando

$$L(1, -2, 1) = (1+2(-2)+3, 4+5(-2)+6, 7+8(-2)+9) = (1-4+3, 4-10+6, 7-16+9) = (0, 0, 0)$$

- ② Falso. O núcleo de  $L$  é diferente de  $\{0\}$ , logo  $\det A = 0 \Rightarrow \lambda_1 \lambda_2 \lambda_3 = 0$ . Outra forma de pensar é que como

$$A \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Logo 0 é um dos autovalores da matriz  $A$ .

- ③ Verdadeiro. Como regra de bolso lembre-se que a soma dos autovalores de uma matriz  $A$  é igual ao traço de  $A$ . logo  $\sum_{i=1}^3 \lambda_i = 1 + 5 + 9 = 15$ .
- ④ Verdadeiro. Também como regra de bolso se uma matriz tem  $n$  autovalores distintos, então ela é diagonalizável (repare que não vale a volta, i.e., uma matriz diagonalizável não necessariamente tem autovalores distintos — tome como exemplo a matriz

---

<sup>1</sup>Última Atualização: 27/06/2012

identidade). O polinômio característico será dado por

$$\begin{vmatrix} 1 - \lambda & 2 & 3 \\ 4 & 5 - \lambda & 6 \\ 7 & 8 & 9 - \lambda \end{vmatrix} = (1 - \lambda)(5 - \lambda)(9 - \lambda) + 84 + 96 - 48(1 - \lambda) - 8(9 - \lambda) - 21(5 - \lambda) = P(\lambda)$$

Repare que 0 é um dos autovalores de  $A$  e, além disso,  $P'(0) = -45 - 9 - 5 + 48 + 8 + 21 = 18 > 0$ . Como o coeficiente de  $\lambda^3$  é negativo, então  $P$  tem 3 raízes reais distintas.

Portanto  $A$  tem 3 autovalores distintos, então é diagonalizável.

**(2007 - Questão 2)**

Considere a matriz:

$$A = \begin{bmatrix} 1 & a & b \\ 0 & 2 & c \\ 0 & 0 & 3 \end{bmatrix}$$

em que  $a, b, c$  são constantes. Julgue os itens abaixo:

- Ⓐ O traço de  $A$  é  $\text{tr}(A) = a + b + c + 6$ .
- Ⓑ O determinante de  $A$  é  $\det(A) = 6$ .
- Ⓒ Se  $a, b, c$  são constantes negativas, a matriz  $A'A$  é definida negativa.
- Ⓓ A matriz  $A'A$  é simétrica
- Ⓔ se  $a = b = c = 0$ , a matriz  $A'A$  é definida positiva.

**Respostas**

- Ⓐ Falso. O traço de uma matriz é a soma dos elementos de sua diagonal principal, logo

$$\text{tr}(A) = 1 + 2 + 3 = 6$$

- Ⓑ Verdadeiro. No caso de matrizes diagonais o determinante é dado pela multiplicação dos elementos da diagonal principal. Logo

$$\det(A) = 1 * 2 * 3 = 6$$

- Ⓒ Falso.  $A'A$  nunca é definida negativa, para ver isso note que  $\vec{x}'A'A\vec{x} = \|A\mathbf{x}\|^2 \geq 0$ .
- Ⓓ Verdadeiro. Para qualquer matriz  $A$  temos que  $A'A$  é simétrica pois

$$(A'A)' = A'A'' = A'A$$

- Ⓔ Verdadeiro. Pelo mesmo argumento de Ⓒ vemos que  $A'A$  é definida positiva se, e somente se,  $\det(A) \neq 0$ . Mas também podemos checar diretamente,

$$\vec{x}'A'A\vec{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1^2 + 4x_2^2 + 9x_3^2 \geq 0$$

com igualdade somente se  $x_1 = x_2 = x_3 = 0$ . Logo  $A'A$  é definida positiva.

**(2007 - Questão 3)**

Seja  $\langle \cdot, \cdot \rangle$  o produto escalar usual de  $\mathbb{R}^{n+1}$  e  $V = V_1 \wedge \dots \wedge V_n \in \mathbb{R}^{n+1}$  o produto vetorial de vetores linearmente independentes  $V_1, \dots, V_n \in \mathbb{R}^{n+1}$ . Por definição  $\langle V, W \rangle = \det A_W$ , em que

$$A_W = \begin{bmatrix} W \\ V_1 \\ \vdots \\ V_n \end{bmatrix}$$

é a matriz cujas linhas são os vetores  $W, V_1, \dots, V_n \in \mathbb{R}^{n+1}$ . Julgue os itens abaixo:

- ①  $\langle V, V_i \rangle = 0$ , para todo  $i \in \{1, \dots, n\}$ .
- ②  $\det A_V \neq |V|^2$ .
- ③  $V \neq 0$ .
- ④  $\det(A_V A_V^t) = |V|^2 \det(g_{ij})$ , em que  $g_{ij} = \langle V_i, V_j \rangle$ .
- ⑤  $|V| = \sqrt{\det(g_{ij})}$ .

**Respostas**

- ① Verdadeiro.

$V$  é ortogonal a qualquer  $V_i$ ,  $i \in \mathbb{N}$ , já que  $V$  é resultado do produto vetorial de todos os  $V_i$ 's. Como sabemos o produto escalar usual de 2 vetores ortogonais é igual a 0.

Uma outra forma de se ver que  $\langle V, V_i \rangle = 0$ , é notar que, segundo a definição dada,  $\langle V, V_i \rangle = \det A_{V_i} = 0$  pois  $A_{V_i}$  é matriz que tem duas linhas iguais, e portanto determinante zero.

- ② Falso.

$$\det(A_v) = \langle V, V \rangle = |V|^2$$

- ③ Verdadeiro.

Pensemos intuitivamente:  $V$  é LI em relação aos  $V_i$ 's, pois  $V$  é resultado do produto vetorial desses vetores. Logo,  $V$  tem que ser diferente de 0.

De modo mais rigoroso,  $V$  pode ser escrito na base canônica:

$$V = \sum_{i=1}^{n+1} \langle V, e_i \rangle e_i$$

Se  $V = 0$ ,

$$\langle V, e_i \rangle = 0 \Rightarrow \det \begin{bmatrix} e_i \\ V_1 \\ \vdots \\ V_n \end{bmatrix} = 0 \quad \forall i \in \{1, 2, \dots, n+1\}$$

mas repare que  $\{V_1, \dots, V_n\}$  não gera uma base no  $\mathbb{R}^{n+1}$  (não tem  $n+1$  vetores), logo existe um  $W \in \mathbb{R}^{n+1}$  tal que  $\{W, V_1, \dots, V_n\}$  é linearmente independente. Podemos escrever  $W = \sum_{i=1}^{n+1} \alpha_i e_i$  e assim temos  $\det(W, V_1, \dots, V_n) =$

$\sum_{i=1}^{n+1} \alpha_i \det(e_i, V_1, \dots, V_n) = 0$ , um absurdo pois  $W$  é linearmente independente de  $V_1, \dots, V_n$ .

③ Verdadeiro.

$$\det(A_v A_v^T) = \det \left( \begin{bmatrix} - & V & - \\ - & V_1 & - \\ & \vdots & \\ - & V_n & - \end{bmatrix} \begin{bmatrix} | & | & \dots & | \\ V & V_1 & \dots & V_n \\ | & | & & | \end{bmatrix} \right)$$

mas se  $V = (\wedge V_i)_{i \leq n}$ , então  $\langle V, V_i \rangle = \langle V, V_j \rangle = 0 \quad \forall i, j \geq 1$  e  $\langle V, V \rangle = \|V\|^2$ . O que nos leva a construir a seguinte matriz:

$$A_v A_v^T = \begin{bmatrix} \|V\|^2 & 0 & 0 & \dots & 0 \\ 0 & \langle V_1, V_1 \rangle & \langle V_1, V_2 \rangle & \dots & \langle V_1, V_n \rangle \\ 0 & \langle V_2, V_1 \rangle & \langle V_2, V_2 \rangle & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & \langle V_n, V_1 \rangle & \dots & \dots & \langle V_n, V_n \rangle \end{bmatrix}$$

Repare que se tirarmos a primeira linha e a primeira coluna dessa matriz teremos exatamente a matriz  $g_{i,j}$  e aplicando a expansão de Laplace a matriz temos:

$$\det(A_v A_v^T) = \|V\|^2 \det G$$

④ Verdadeiro.

Sabemos que  $\det A_v = \det A_v^T$  e  $\det A_v = \langle V, V \rangle = \|V\|^2$ . Logo

$$\det A_v A_v^T = \det A_v \det A_v^T = \|V\|^2 \|V\|^2$$

retomando a equação do item anterior:

$$\det(A_v A_v^T) = \|V\|^2 \det G \Rightarrow \|V\|^4 = \|V\|^2 \det G \Rightarrow \|V\| = \sqrt{\det G}$$

**(2007 - Questão 4)**

Considere as funções:

$$f(x) = \begin{cases} x^2, & \text{se } x \geq 0 \\ -x^2, & \text{se } x < 0 \end{cases} \quad g(x) = \begin{cases} x, & \text{se } x > 1 \\ x^3, & \text{se } x \leq 1 \end{cases}$$

Com relação aos conceitos de continuidade e diferenciabilidade, julgue os itens abaixo:

- ① A função  $f$  é contínua em  $x = 0$ .
- ① A derivada de  $f$  não é contínua em  $x = 0$ .
- ② A função  $g$  é diferenciável em  $x = 1$ .
- ③ A segunda derivada de  $f$  é diferenciável em  $x = 0$ .
- ④ A função  $h$ , definida por  $h(x) = |f(x)|$ , não é diferenciável em  $x = 0$ .

**Respostas**

- ① Verdadeira.

$$\lim_{x \rightarrow 0^+} f(x) = \lim_{x \rightarrow 0^+} x^2 = 0 = f(0) = 0 = \lim_{x \rightarrow 0^-} -x^2 = \lim_{x \rightarrow 0^-} f(x)$$

- ① Falso.

$$f'(x) = \begin{cases} 2x, & \text{se } x \geq 0 \\ -2x, & \text{se } x < 0 \end{cases}$$

logo

$$\lim_{x \rightarrow 0^+} f'(x) = \lim_{x \rightarrow 0^+} 2x = 0 = f'(0) = 0 = \lim_{x \rightarrow 0^-} -2x = \lim_{x \rightarrow 0^-} f'(x)$$

- ② Falso.

$$\begin{aligned} \lim_{h \rightarrow 0^+} \frac{g(1+h) - g(1)}{h} &= \lim_{h \rightarrow 0^+} \frac{h}{h} = 1 \neq 3 = \lim_{h \rightarrow 0^-} \frac{1 + 3h + 3h^2 + h^3 - 1}{h} \\ &= \lim_{h \rightarrow 0^-} \frac{g(1+h) - g(1)}{h} \end{aligned}$$

- ③ Falso. Como vimos em ①,  $f'(x) = 2|x|$ , que não é diferenciável em  $x = 0$ , logo não faz nem sentido perguntar se a segunda derivada de  $f$  é diferenciável em  $x = 0$ , a função não está nem definida para esse valor.
- ④ Falso.  $h(x) = x^2$  que é diferenciável para qualquer  $x \in \mathbb{R}$ .

**(2010 - Questão 9)**

Considere os sistemas lineares abaixo e julgue as afirmativas:

$$(I) = \begin{cases} x + y + kz = 2 \\ 3x + 4y + 2z = k \\ 2x + 3y - z = 1 \end{cases}$$

$$(II) = \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

- ① Se  $k \neq 3$ , então o sistema (I) tem solução única;
- ① Se  $k = 0$ , o sistema homogêneo associado a (I) tem infinitas soluções;
- ② Para  $k = 1$ , a matriz dos coeficientes de (I) é uma matriz ortogonal;
- ③ Se  $m > n$ , (II) tem sempre solução;
- ④ Se  $b_1 = b_2 = \dots = b_m = 0$ , então o sistema (II) tem sempre solução;

**Respostas**

- ① **Verdadeiro.** Numeremos, antes de mais nada, as equações de acordo com a ordem em que aparecem. Considere, então, o sistema formado pela equação 1, 3 e pela equação correspondente à equação 2 subtraída da equação 3. O novo sistema foi obtido por meio de operações elementares, portanto tem as mesmas soluções que o inicial:

$$(I) \sim \begin{cases} x + y + kz = 2 \\ x + y + 3z = k - 1 \\ 2x + 3y - z = 1 \end{cases}$$

Fazendo-se uma nova operação elementar com as linhas, agora repetindo-se a segunda e a terceira equações e subtraindo-se a segunda da primeira, tem-se:

$$(B.0.1) \quad (I) \sim \begin{cases} (k-3)z = 3-k \\ x + y + 3z = k-1 \\ 2x + 3y - z = 1 \end{cases}$$

Se  $k \neq 3$ , o único  $z$  que satisfaz ao sistema é  $z = -1$ . Nesse caso, os pares  $(x, y)$  satisfazem ao sistema se, e somente se

$$\begin{cases} x + y - 3 = k - 1 \\ 2x + 3y + 1 = 1 \end{cases} \sim \begin{cases} x + y = k + 2 \\ 2x + 3y = 0 \end{cases} \sim \begin{cases} 2x + 2y = 2k + 4 \\ 2x + 3y = 0 \end{cases}$$

Por fim, repetindo-se a primeira linha do último sistema e subtraindo-se a primeira equação da segunda, tem-se:

$$\begin{cases} 2x + 2y = 2k + 4 \\ y = -(2k + 4) \end{cases}$$

E esse sistema, equivalente ao primeiro (supondo-se  $k \neq 3$ ), tem claramente uma única solução.

- ① **Falso.** Observe que, se  $k = 0$ , então  $k \neq 3$ , logo o sistema (I) tem uma única solução, digamos,  $\mathbf{a}^t = (x, y, z)$ . Isso garante que o sistema homogêneo associado tenha somente uma solução. Para provar isso, escrevemos o sistema na forma matricial, com

$$M\mathbf{a} = \begin{bmatrix} 1 & 1 & k \\ 3 & 4 & 1 \\ 2 & 3 & -1 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ k \\ 1 \end{bmatrix} = \mathbf{b}$$

Onde definiu-se como  $M$  a matriz associada ao sistema e  $\mathbf{b}^t = (2, k, 1)$ . Suponha agora que, para dado  $u \in \mathbb{R}^3$ , valha que  $M\mathbf{u} = \mathbf{0}$ . Então, por exemplo,

$$M(\mathbf{a} + \mathbf{u}) = M(\mathbf{a}) + M(\mathbf{u}) = M(\mathbf{a}) = \mathbf{b}$$

Ou seja,  $\mathbf{a} + \mathbf{u}$  é solução do sistema. Mas a solução é única, logo  $\mathbf{a} + \mathbf{u} = \mathbf{a}$ . Conclui-se que  $\mathbf{u} = \mathbf{0}$ . Em particular, há uma única solução do sistema homogêneo.

- ② **Falso.** A definição de uma *matriz ortogonal*  $n \times n$  impõe que as colunas formem uma base ortonormal de  $\mathbb{R}^n$ . Ou seja: as colunas distintas devem ser ortogonais e cada uma delas deve ter norma 1. Mas a coluna  $c_1^t = (1, 3, 2)$  não o tem.
- ③ **Falso.** Considere um sistema sem solução e repita uma das equações até que o número de linhas fique maior que o de colunas. Ou ainda, o sistema

$$\begin{cases} 2x = 5 \\ x = 0 \end{cases}$$

satisfaz às condições do enunciado, e claramente não tem solução.

- ④ **Verdadeiro.** Quando  $b_1 = b_2 = \dots = 0$ , o sistema será homogêneo. Ora, qualquer sistema linear homogêneo admite uma solução, a saber, a solução  $x_1 = x_2 = \dots = x_n = 0$ .



**(2010 - Questão 10)**

- ①  $\{(x, y, x + y) \in \mathbb{R}^3 / x, y \in \mathbb{R}\}$  é um subespaço vetorial de  $\mathbb{R}^3$  e a dimensão de  $S$  é 2;
- ①  $\{(1, 2, 3), (4, 5, 12), (0, 8, 0)\}$  é base de  $\mathbb{R}^3$
- ② Se  $u, v$  e  $w$  são vetores linearmente independentes, então  $v + w$ ,  $u + w$  e  $u + v$  são também linearmente independentes;
- ③ Se  $S$  é um subconjunto de  $\mathbb{R}^3$  formado por vetores linearmente dependentes, então podemos afirmar que  $S$  tem 4 elementos ou mais
- ④ Se o posto da matriz  $\begin{bmatrix} 1 & x & 0 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{bmatrix}$  é 3, então  $x \neq 1$ .

**Respostas**

- ① **Verdadeiro** O conjunto  $S$  é um subespaço vetorial, pois, dados  $(u, v, w) \in S$  e  $(u', v', w') \in S$ , existem  $(x, y) \in \mathbb{R}^2$  e  $(x', y') \in \mathbb{R}^2$  tais que:
- $u = x, v = y$  e  $w = x + y$
  - $u' = x', v' = y'$  e  $w' = x' + y'$

Em particular,

$$(u, v, w) + (u', v', w') = (u + u', v + v', w + w') = (x + x', y + y', (x + x') + (y + y')) \in S$$

Ou seja, a soma de elementos de  $S$  pertence a  $S$ . Por outro lado, dado  $t \in \mathbb{R}$ ,

$$t(u, v, w) = t(x, y, x + y) = (tx, ty, tx + ty) \in S$$

Logo a multiplicação de elementos de  $S$  por escalar os mantém em  $S$ . Isso é suficiente para que  $S$  seja um subespaço. Para avaliar sua dimensão, observe que  $(u, v, w) \in S$  se, e somente se, existem  $x, y \in \mathbb{R}$  tais que

$$\begin{aligned} (u, v, w) = (x, y, x + y) &= x(1, 0, 0) + y(0, 1, 0) + (x + y)(0, 0, 1) \\ &= x[(1, 0, 0) + (0, 0, 1)] + y[(0, 1, 0) + (0, 0, 1)] = \\ &= x(1, 0, 1) + y(0, 1, 1) \end{aligned}$$

Ou seja, o conjunto  $S$  é precisamente o conjunto das combinações lineares dos vetores linearmente independentes  $(1, 0, 1)$  e  $(0, 1, 1)$ ! Logo sua dimensão é 2.

- ① **Falso**. Sejam  $u_1 = (1, 2, 3)$ ,  $u_2 = (4, 5, 12)$  e  $u_3 = (0, 8, 0)$ . Considere agora o vetor  $u = 4u_1 - u_2 = (4, 8, 12) - (4, 5, 12) = (0, 3, 0)$ . Tal vetor é múltiplo de  $u_3$ , logo o conjunto  $\{u_1, u_2, u_3\}$  não é linearmente independente. Não pode, portanto, ser base de  $\mathbb{R}^3$ .
- ② **Verdadeiro**. Sejam  $\alpha, \beta$  e  $\gamma$  números reais tais que

$$\alpha(u + w) + \beta(v + w) + \gamma(u + v) = 0$$

Um teorema importante garante que eles serão linearmente independentes se, e somente se, a equação acima implicar  $\alpha = \beta = \gamma = 0$ . Mas, continuando a conta:

$$\begin{aligned} 0 &= \alpha u + \alpha w + \beta v + \beta w + \gamma u + \gamma v = \\ &= (\alpha + \gamma)u + (\beta + \gamma)v + (\alpha + \beta)w \end{aligned}$$

mencionado teorema, garante-se, pela independência linear de  $\{u, v, w\}$ , que

$$\begin{cases} \alpha + \gamma = 0 \\ \beta + \gamma = 0 \\ \alpha + \beta = 0 \end{cases}$$

Se a única solução do sistema acima for  $\alpha = \beta = \gamma = 0$ , a independência linear de  $\{u, v, w\}$  implicará a independência linear de  $\{u + w, v + w, u + v\}$ . Avaliemos, portanto, o sistema acima.

Subtraindo-se da terceira equação a segunda, tem-se, equivalentemente:

$$\begin{cases} \alpha + \gamma = 0 \\ \beta + \gamma = 0 \\ \alpha + \beta = 0 \end{cases} \sim \begin{cases} \alpha + \gamma = 0 \\ \beta + \gamma = 0 \\ \alpha - \gamma = 0 \end{cases}$$

Conclui-se, somando-se a primeira à terceira equação no novo sistema, que  $\alpha = 0$ . Vê-se, imediatamente, que disso decorre  $\beta = \gamma = 0$ .

③ **Falso.** Tome  $S = \{(0, 0, 0), (1, 0, 0)\}$ .

④ **Falso** Calculemos o determinante da matriz em questão pela expansão de Laplace:

$$\begin{aligned} \det \begin{pmatrix} 1 & x & 0 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix} &= 1 * \det \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} - x * \det \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= -1 - x \end{aligned}$$

Se, portanto,  $x = 1$ , o determinante resulta em  $-2$ , portanto a matriz tem posto 3.

OBS: O  $x$  que faz a matriz ter determinante nulo é  $x = -1$ ; de fato, vê-se que, nesse caso, a primeira linha será exatamente a terceira multiplicada por  $-1$ ; é óbvio, nesse caso, que o posto não é 3.

Alternativamente, poderia ter sido usada a redução da matriz à forma "escada" para chegar à mesma conclusão.

(2010 - Questão 11) Considere as matrizes

$$A = \begin{bmatrix} 1 & a \\ 2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & b \\ b & 1 \end{bmatrix}, \quad C = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}.$$

Julgue as afirmativas:

- ① Para  $a = 1$  e  $b = 2$ , então  $(3A - B^t)^t = \begin{bmatrix} 2 & 1 \\ 4 & -4 \end{bmatrix}$ ;  
 ① Se  $-1$  é autovalor de  $A$ , então  $a = 0$ ;  
 ② Para  $b = 2$ ,  $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  é um autovetor de  $B$ .  
 ③ Se  $a > -1/2$ , então  $A$  é diagonalizável;  
 ④  $C$  é invertível não simétrica.

### Respostas

- ① **Falso.** Usando-se as propriedades da transposição (a saber, para quaisquer matrizes  $A, B$ , de mesma dimensão, para qualquer constante  $c \in \mathbb{R}$ ,  $(cA + B)^t = cA^t + B^t$  e  $(A^t)^t = A$ ):

$$\begin{aligned} (3A - B^t)^t &= 3A^t - (B^t)^t = 3A^t - B \\ &= 3 \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 4 \\ 1 & -4 \end{bmatrix} \end{aligned}$$

O que não corresponde à matriz pedida. (Corresponde à sua transposta!)

- ① **Verdadeiro.** Se  $-1$  é autovalor de  $A$ , existe  $v \in \mathbb{R}^2$ ,  $v \neq \mathbf{0}$ , tal que  $Av = -v$ , isto é,  $(A + I)v = 0$ . Isso só ocorre se a matriz  $A + I$  tem posto  $< 2$ . Mas

$$A + I = \begin{bmatrix} 2 & a \\ 2 & 0 \end{bmatrix}$$

E a matriz acima tem posto  $< 2$  se, e somente se,  $a = 0$ , QED. Recapitulando:  $-1$  ser autovalor de  $A$  implica que a matriz  $A + I$  não tenha posto cheio, logo conclui-se que  $a = 0$ .

- ② **Falso.** Se  $b = 2$ ,

$$Bv = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 + 4 \\ 2 + 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

que claramente não é múltiplo de  $(1, 2)^t$ .

- ③ **Indefinido (gabarito oficial: verdadeiro).** O polinômio característico de  $A$  é, para qualquer  $a \in \mathbb{R}$ :

$$\begin{aligned} \det \begin{pmatrix} 1 - \lambda & a \\ 2 & -(1 + \lambda) \end{pmatrix} &= -(1 - \lambda)(1 + \lambda) - 2a \\ &= -[1^2 - \lambda^2] - 2a = \lambda^2 - (1 + 2a) \end{aligned}$$

As raízes do polinômio serão reais e distintas sempre que  $(1 + 2a) > 0$ , em particular, serão dadas por  $\lambda_1 = 1 + 2a$  e  $\lambda_2 = -(1 + 2a)$ .

Mas a condição  $1 + 2a > 0$  equivale, por sua vez, a  $a < (-1/2)$ . Como nada se disse sobre se as matrizes estão ou não definidas somente para números reais, o caso em que  $1 + 2a < 0$ , i.e.,  $a > -1/2$  fica indefinido. Se as matrizes em questão estiverem definidas sobre o corpo dos números complexos, a matriz será diagonalizável. O gabarito oficial dá este item como verdadeiro.

- ④ **Falso; questão anulada.** Observe que a matriz  $C$  é *ortogonal*. A inversa de  $C$ , portanto, é  $C^t$ . Para dar segurança, pode-se calcular:

$$\begin{aligned} C^t C &= \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \\ &= \begin{bmatrix} (\cos \theta)^2 + (\sin \theta)^2 & \cos \theta \sin \theta - \sin \theta \cos \theta \\ \sin \theta \cos \theta - \cos \theta \sin \theta & (\sin \theta)^2 + (\cos \theta)^2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

Adicionalmente, a matriz em questão não cumpre a definição de simetria ( $C \neq C^T$ ) para quaisquer  $\theta \in \mathbb{R}$  tais que  $-\sin \theta \neq \sin \theta$ , isto é,  $\sin \theta \neq 0$ . Entretanto, para  $\theta = k\pi$ , com  $k \in \mathbb{Z}$ , a matriz é simétrica. Em outras palavras, para “quase todo”  $\theta$ , a matriz não é simétrica, mas esses valores são suficientes para invalidar o enunciado.

(2011 - Questão 2) Considere as retas  $r_1$  e  $r_2$ , no plano, definidas por

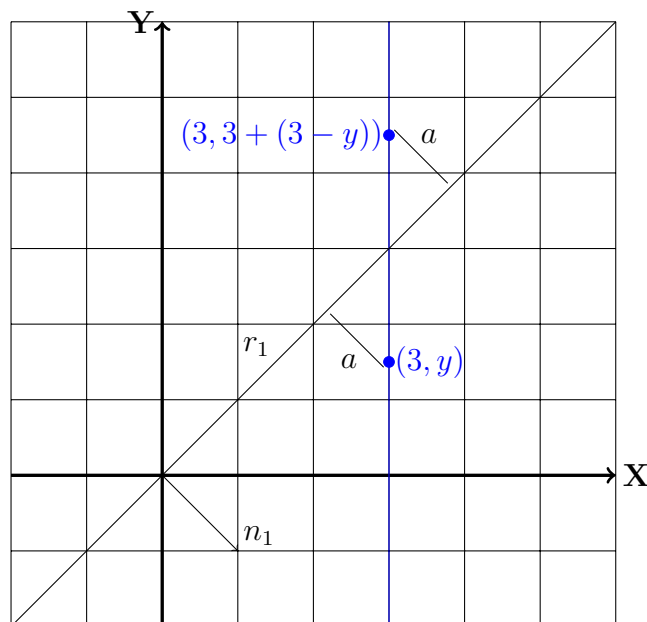
$$\begin{cases} a_1x + b_1y + c_1 = 0 \\ a_2x + b_2y + c_2 = 0 \end{cases}$$

em que  $n_1 = (a_1, b_1)$  e  $n_2 = (a_2, b_2)$  são vetores não nulos ortogonais à  $r_1$  e  $r_2$ , respectivamente. Denotamos por  $d(P, r)$  a distância de um ponto  $P$  à uma reta  $r$  do plano. Julgue as afirmativas:

- ① Se as retas  $r_1$  e  $r_2$  são perpendiculares, então  $a_1a_2 + b_1b_2 = 0$ .
- ① Se  $(1, 1) \in r_1$  e  $r_1$  é paralela à reta dada por  $2x + 3y - 6 = 0$ , então  $(3, 2) \in r_1$ .
- ② Considere em  $r_1$  os valores  $c_1 = 0$  e  $n_1 = (1, -1)$ . Se pontos distintos  $P = (3, y_1)$  e  $Q = (3, y_2)$  são tais que  $d(P, r_1) = d(Q, r_1) = \sqrt{2}$ , então  $y_1 + y_2 = 6$ .
- ③ As retas  $y = x$ ,  $y = 1$  e  $y = -x + 2$  se interceptam formando um triângulo.
- ④ Se  $a_2b_2c_2 \neq 0$  e  $\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$ , então  $r_1$  e  $r_2$  representam a mesma reta.

### Respostas

- ① Verdadeiro. Se  $r_1$  e  $r_2$  são perpendiculares, então um vetor ortogonal a  $r_1$  será perpendicular a um vetor ortogonal a  $r_2$ , logo  $\langle n_1, n_2 \rangle = 0$ , *i.e.*,  $a_1a_2 + b_1b_2 = 0$ .
- ① Falso. Se  $r_1$  é paralela a reta definida por  $2x + 3y - 6 = 0$ , então  $r_1$  é definida por  $2x + 3y + c = 0$ , substituindo o ponto  $(1, 1)$  temos que  $c = -5$ . Agora substituímos o ponto  $(3, 2)$  e temos  $2 * 3 + 3 * 2 - 5 = 7$ , logo  $(3, 2) \notin r_1$ . Uma forma de perceber isso imediatamente é notar que os coeficientes de  $x$  e  $y$  tem o mesmo sinal, logo um aumento em  $x$  tem que ser compensado por uma redução de  $y$  e vice-versa.
- ② Verdadeiro. Aqui uma boa ideia é fazer um desenho



Note que se  $(3, y)$  está abaixo de  $(3, 3)$  e a uma certa distância  $a$  de  $r_1$ , então, pela simetria da figura,  $(3, 3 + (3 - y))$  também está a uma distância  $a$  de  $r_1$ , agora repare que independente de  $a$  vale  $y_1 + y_2 = y + 3 + (3 - y) = 6$ .

Mas também podemos usar a fórmula da distância de um ponto a uma reta

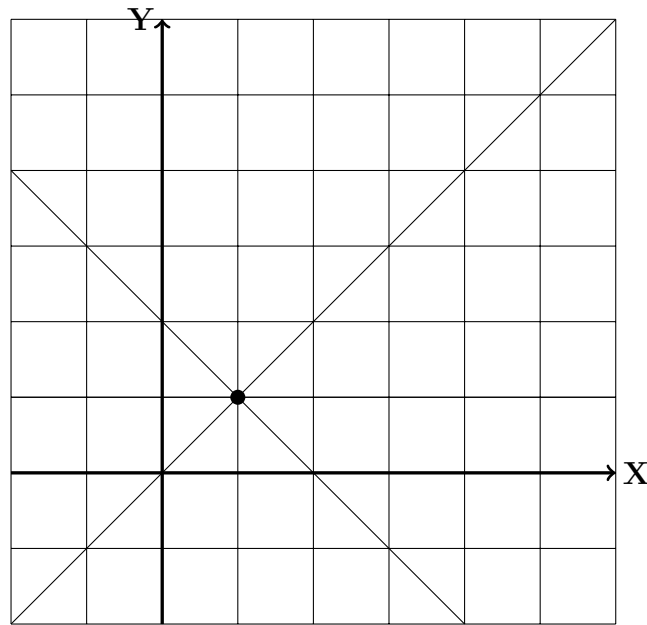
$$d(P, r_1) = \frac{|\langle P, n_1 \rangle + c_1|}{\|n_1\|}$$

Substituindo as informações do problema, temos

$$\frac{|3 - y|}{\sqrt{2}} = \sqrt{2} \quad \Rightarrow \quad y = 5 \quad \text{ou} \quad y = 1$$

Logo  $y_1 + y_2 = 6$ .

- ③ Falso. Só tem dois jeitos de 3 retas no  $\mathbb{R}^2$  não formarem um triângulo, se pelo menos duas das retas forem paralelas — não é esse o caso — ou as 3 passam pelo mesmo ponto — e é o caso, o ponto  $(1, 1)$  pertence às 3 retas.



- ④ Verdadeiro. Se  $\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} = \alpha$ , então

$$(x, y) \text{ satisfaz } a_2x + b_2y + c_2 = 0 \quad \Leftrightarrow \quad 0 = \alpha a_2x + \alpha b_2y + \alpha c_2 = a_1x + b_1y + c_1$$

Logo representam a mesma reta<sup>2</sup>

<sup>2</sup>Note que  $\alpha \neq 0$  pois  $n_1$  não pode ser o vetor nulo.



**(2011 - Questão 6)**

Considere as transformações lineares  $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  e  $L: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  definidas por

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - 2y + 3z \\ 3y - 2z \\ -y + 2z \end{pmatrix} \quad \text{e} \quad L \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Seja  $A$  a matriz de  $T$  relativa à base canônica de  $\mathbb{R}^3$ . Julgue as afirmativas:

- ①  $L$  é sobrejetora.  
 ① Se  $v \in \mathbb{R}^3$  é tal que  $v^t = (-1, -1, 1)$ , então  $\{v\}$  é base para o Núcleo de  $L$ .  
 ②  $A = \begin{bmatrix} 2 & -2 & 3 \\ 0 & 3 & -2 \\ 0 & -1 & 2 \end{bmatrix}$   
 ③  $A$  possui três autovalores distintos e portanto é diagonalizável.  
 ④  $v \in \mathbb{R}^3$  é tal que  $v^t = (1, 1, 1)$ , então  $v$  é autovetor de  $A$  associado ao autovalor 1.

**Respostas**

- ① Falso. Basta verificar se o determinante da matriz associada a  $L$  é diferente de 0.

$$\begin{vmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{vmatrix} = 3 + 0 + 1 - 2 - 2 - 0 = 0$$

Portanto  $L$  não pode ser sobrejetora.

- ① Verdadeiro. Repare que  $L(v) = (-1 + 1, -1 - 1 + 2, -2 - 1 + 3) = (0, 0, 0)$ . Além disso perceba que  $A$  tem posto 2, logo seu núcleo tem dimensão 1 e tem que ser gerado por  $\{v\}$ .  
 ② Verdadeiro. É fácil verificar que

$$\begin{bmatrix} 2 & -2 & 3 \\ 0 & 3 & -2 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{pmatrix} 2x - 2y + 3z \\ 3y - 2z \\ -y + 2z \end{pmatrix} = T \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

- ③ Verdadeiro. Se  $A$  tem três autovalores distintos, então é diagonalizável. Resta checar que  $A$  tem de fato três autovalores distintos.

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2 - \lambda & -2 & 3 \\ 0 & 3 - \lambda & -2 \\ 0 & -1 & 2 - \lambda \end{vmatrix} = (2 - \lambda)^2(3 - \lambda) - 2(2 - \lambda) = (2 - \lambda)(6 - 5\lambda + \lambda^2 - 2) = \\ &= (2 - \lambda)(4 - \lambda)(1 - \lambda) \end{aligned}$$

- ④ Falso.

$$\begin{bmatrix} 2 & -2 & 3 \\ 0 & 3 & -2 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{pmatrix} 2 - 2 + 3 \\ 3 - 2 \\ -1 + 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \neq \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \forall \alpha \in \mathbb{R}$$



(2012 - Questão 1) Sejam  $A$  e  $B$  conjuntos. A diferença entre  $A$  e  $B$  é o conjunto  $A - B = \{x : x \in A \text{ e } x \notin B\}$ .

Julgue as afirmativas:

- ①  $(A \cup B) - C = (A - C) \cap (B - C)$ , quaisquer que sejam os conjuntos  $A, B$  e  $C$ .
- ① Se  $A - B = B - A$ , então  $A = B$ .
- ② Seja  $N$  o conjunto dos inteiros positivos. Se  $A = \{x \in N : x|12\}$  e  $B = \{x \in N : 4|x\}$ , então  $A \cap B$  é um conjunto unitário, em que  $x|y$  significa que existe  $c \in N$ , tal que  $y = cx$ .
- ③ Se  $A = \{x \in R : x - 2x^2 < 0\}$  e  $B = \{x \in R : |x| \leq 3\}$ , então  $A \cap B \subset (0, 3)$ .
- ④ Se  $A = \{(x, y) \in R^2 : |x| + |y| > 3\}$  e  $B = \{(x, y) \in R^2 : |x + y| > 3\}$ , então  $A \supset B$ .

### Respostas

- ① **Verdadeiro.** Observe que, para quaisquer conjuntos  $X$  e  $Y$ , vale que  $X - Y = X \cap \mathcal{C}(Y)$ . Desse modo:

$$\begin{aligned} (A \cup B) - C &= (A \cup B) \cap \mathcal{C}(C) = (A \cap \mathcal{C}(C)) \cup (B \cap \mathcal{C}(C)) \\ &= (A - C) \cap (B - C) \end{aligned}$$

A segunda igualdade decorre da distributividade das operações  $\cup$  e  $\cap$  de conjuntos.

- ① **Falso.** Tome como contra-exemplo dois conjuntos não-vazios disjuntos quaisquer, e.g.,  $(-1, 0)$  e  $(1, 2)$ .
- ② **Falso.** Veja que  $A$  é o conjunto dos inteiros que dividem 12, i.e.,

$$A = \{1, 2, 3, 4, 6, 12\}.$$

O conjunto  $B$ , por sua vez, é o conjunto dos números divisíveis por 4, isto é,  $B = \{4, 8, 12, 16, \dots\}$ . Logo  $A \cap B = \{4, 12\}$ .

A "pegadinha" deste item reside no fato de que qualquer número é sempre um divisor de si mesmo, e o aluno incauto poderia ter esquecido-se de incluir o número 12 no conjunto  $A$ , o que resultaria em uma interseção unitária.

- ③ **Falso.** Veja que  $B = \{x \in R : |x| \leq 3\} = [-3, 3]$ . Já o conjunto  $A$  é tal que  $x \in A$  se, e somente se,  $x(1 - 2x) < 0$ . Para que valha a desigualdade, os números  $x$  e  $1 - 2x$  devem ter sinais opostos, i.e., devem valer:

- $x < 0$  e  $1 - 2x > 0 \Leftrightarrow x < 0$  e  $x < 1/2 \Leftrightarrow x < 0$ , ou
- $x > 0$  e  $1 - 2x < 0 \Leftrightarrow x > 0$  e  $x > 1/2 \Leftrightarrow x > 1/2$

O conjunto em questão, portanto, é  $B = (-\infty, 0) \cup (1/2, \infty)$ . Vê-se que qualquer número em  $[-3, 0)$  pertence tanto a  $A$  quanto a  $B$ , portanto também pertence a  $A \cap B$ , mas não pertence a  $(0, 3)$ .

O aluno atento teria desenhado um gráfico dos dois conjuntos e concluído imediatamente que a afirmação é falsa.

- ④ **Verdadeiro.** Tome  $(x, y) \in B$ . Então  $3 < |x + y|$ . Pela desigualdade triangular,  $|x + y| \leq |x| + |y|$ , logo

$$3 < |x + y| \leq |x| + |y| \Rightarrow |x| + |y| > 3$$

e  $(x, y) \in A$ . Conclui-se que  $B \subset A$ .

(2012 - Questão 3) Julgue as afirmativas:

- ① A equação da reta que passa por  $(\frac{1}{3}, \frac{2}{5})$  e é paralela à reta que passa por  $(0, 3)$  e por  $(5, 0)$  é  $3x + 5y + 3 = 0$ .
- ① As circunferências  $C_1$  de centro  $(0, 0)$  e raio 1 e  $C_2$  de centro em  $(1, 0)$  e raio 2 se interceptam num único ponto.
- ② Os pontos  $(1, 1)$ ,  $(2, 3)$  e  $(a, -8)$  pertencem a mesma reta se e somente se  $a = \frac{7}{2}$ .
- ③ Sejam  $P = (3, -1, 2)$  e  $Q = (4, -2, -1)$ . A equação do plano que passa por  $P$  e é perpendicular ao vetor  $\overrightarrow{PQ}$  é  $x - y - 3z + 2 = 0$ .
- ④ Sejam  $m, k \in \mathbb{R}$ . Se os planos  $2x + ky + 3z - 5 = 0$  e  $mx - 6y - 6z + 2 = 0$  são paralelos, então  $k + m = -1$ .

### Respostas

- ① **Falso.** O jeito mais rápido de responder a esse item é simplesmente substituir o ponto na reta e verificar que não é satisfeita a equação.

Para quem gosta das contas feitas: seja  $r_1$  a reta que passa por  $(0, 3)$  e por  $(5, 0)$ . Ela é certamente paralela ao vetor  $(5, -3) = (0, 5) - (0, 3)$ . Um vetor ortogonal a  $r_1$ , portanto, é  $(3, 5)$ . A equação de qualquer reta paralela a  $r_1$  é dada, assim, por  $3x + 5y + C = 0$ , em que  $C \in \mathbb{R}$ .

Substituindo-se  $(\frac{1}{3}, \frac{2}{5})$  na reta arbitrária, chega-se a  $C = -3$ , portanto a reta paralela a  $r_1$  que passa por  $(\frac{1}{3}, \frac{2}{5})$  tem equação  $3x + 5y - 3 = 0$ .

- ① **Verdadeiro.** Pode-se simplesmente fazer o desenho para concluir o resultado (bastante intuitivo). Rigorosamente, suponha que um par  $(x, y)$  pertença a circunferência de centro na origem e raio 1. Então  $x^2 + y^2 = 1$ . Por outro lado, se  $(x, y)$  também pertence à circunferência de centro  $(1, 0)$  e raio 2, o par  $(x, y)$  também deve satisfazer a

$$4 = (x - 1)^2 + y^2 = x^2 - 2x + 1 + y^2 = (x^2 + y^2) + 1 - 2x = 2 - 2x$$

o que ocorre se, e somente se,  $x = -1$ . O único  $y$  que satisfaz a  $x^2 + y^2 = 1$  com  $x = -1$  é  $y = 0$ . Como  $(-1, 0)$  também está na outra circunferência, só existe um ponto em que as circunferências se interceptam (a saber,  $(-1, 0)$ ).

- ② **Verdadeiro.** Calculemos a reta que passa pelos pontos  $(1, 1)$  e  $(2, 3)$ . Um vetor perpendicular a ela é  $(2, -1)$ , logo a sua equação é  $2x - y + C = 0$ . Substituindo-se  $(1, 1)$ , encontra-se  $C = -1$ . Os pontos  $(x, y)$  pertencentes à reta satisfazem, portanto, a  $2x - y - 1 = 0$ .

Dizer, então, que os pontos  $(a, -8)$ ,  $(1, 1)$  e  $(2, 3)$  pertencem à mesma reta equivale a afirmar que  $(a, -8)$  pertence à reta dada por  $2x - y - 1 = 0$ . Equivalentemente, substituindo-se  $(a, -8)$ :

$$2a - (-8) - 1 = 0 \Leftrightarrow a = 7/2$$

- ③ **Verdadeiro.** Observe que  $\overrightarrow{PQ} = (1, -1, -3)$ . Um vetor perpendicular ao plano, portanto, é  $(1, -1, -3)$ , de onde conclui-se que existe um  $C \in \mathbb{R}$  tal que, para todo ponto  $(x, y, z)$  no plano,  $x - y - 3z + C = 0$ . Substituindo-se  $P$  na equação, obtém-se  $C = 2$ , logo o item é verdadeiro.
- ④ **Verdadeiro.** Os planos  $\Phi = \{(x, y, z) : 2x + ky + 3z = 5\}$  e  $\Psi = \{(x, y, z) : mx - 6y - 6z + 2 = 0\}$  são paralelos, por definição, se os vetores ortogonais a eles

são paralelos. É necessário e suficiente, portanto, que  $(2, k, 3)$  e  $(m, -6, -6)$  sejam paralelos para que os planos também o sejam. Assim, deve existir  $\alpha \in \mathbb{R}$  tal que  $(2, k, 3) = \alpha(m, -6, -6)$ . Equivalentemente:

$$\begin{cases} 2 = \alpha m \\ k = -6\alpha \\ 3 = -6\alpha \end{cases}$$

Conclui-se que, para que sejam paralelos, deve-se ter  $\alpha = -1/2$ , logo  $k = 3$  e  $m = -4$ , logo  $k + m = -1$ .

**(2012 - Questão 4)**

Seja  $A = (a_{ij})$  uma matriz  $n \times n$  com entradas  $(a_{ij}) \in R$ . Julgue as afirmativas:

- ① Existe uma matriz  $B$  de modo que  $BA = 2A$ .
- ② Se  $A^2 + A = I$ , então  $A^{-1} = A + I$ , em que  $I$  é a matriz identidade.
- ③ Se todos os elementos da diagonal principal de  $A$  são nulos, então  $\det(A) = 0$ .
- ④ Seja  $b \in R^n$ . Se  $Ax = b$  possui infinitas soluções, então existe  $c \in R^n$ , tal que  $Ax = c$  admite uma única solução.
- ⑤ Suponha que  $a_{ij} = 0$  quando  $i + j$  for par e  $a_{ij} = 1$  quando  $i + j$  for ímpar. Se  $n \geq 3$ , então  $A$  tem posto  $n$ .

**Respostas**

- ① Verdadeiro.

$$BA = 2A \Rightarrow BA - 2IA = 0 \Rightarrow (B - 2I)A = 0$$

repare que  $B = 2I$  é sempre uma resposta válida para essa equação.

- ② Verdadeiro.

$$A^2 + A = I \Rightarrow A(A + I) = I$$

mas repare que, por definição, a matriz que multiplica  $A$  e resulta na sua identidade é a inversa de  $A$ . Logo

$$A^{-1} = (A + I)$$

- ③ Falso.

Observemos a matriz rotação  $90^\circ$  por exemplo

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

seu determinante é 1.

- ④ Falso.

Se  $\det(A)$  fosse diferente de 0 a inversa de  $A$  estaria bem definida e teríamos

$$A^{-1}Ax = A^{-1}b \Rightarrow x = A^{-1}b$$

logo, o único caso em que  $x$  tem infinitas soluções é quando  $\det(A) = 0$ . Sabemos então que  $Ax = c$  tem infinitas soluções ou não tem nenhuma solução.

- ⑤ Falso.

Repare que qualquer linha da matriz é uma sequência alternada entre zeros e uns de modo que só existem 2 “tipos” de linhas diferentes: as que começam com 0 e as que começam com 1, mas qualquer matriz que só tenha 2 “tipos” de linha terá posto 1 ou 2.

## Referências Bibliográficas

- [1] G. Ávila, *Cálculo*, Livros Técnicos e Científicos S.A., Rio de Janeiro, 1987. Vols. I, II e III.
- [2] Robert G. Bartle, *The elements of real analysis*, 2nd ed., John Wiley & Sons, New York-London-Sydney, 1976. MR0393369 (52 #14179)
- [3] Robert G. Bartle and Donald R. Sherbert, *Introduction to real analysis*, 2nd ed., John Wiley & Sons Inc., New York, 1992. MR1135107 (92i:26002)
- [4] José Boldrini, Sueli Costa, Vera Ribeiro, and Henry Wetzler, *Álgebra Linear*, Editora Harper & Row di Brasil LTDA, São Paulo, 1978.
- [5] A.C. Chiang, *Matemática para Economistas*, McGraw-Hill, São Paulo.
- [6] Djairo G. de Figueiredo, *Funções Reais*, The PanAmerican Union, Washington D.C., 1970 (Portuguese).
- [7] R. Fletcher, *Practical methods of optimization*, 2nd ed., A Wiley-Interscience Publication, John Wiley & Sons Ltd., Chichester, 1987. MR955799 (89j:65050)
- [8] W.A. Granville, P.F. Smith, and W.R. Longley, *Elementos de cálculo diferencial e integral*, Editora Científica, Rio de Janeiro.
- [9] H.L. Guidorizzi, *Um Curso de Cálculo*, Vol. Vols. 1 a 4, Forense-Universitária, Rio de Janeiro. 2<sup>a</sup> ed.
- [10] G. Hadley, *Álgebra Linear*, Forense-Universitária, Rio de Janeiro.
- [11] Roger A. Horn and Charles R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985. MR832183 (87e:15001)
- [12] Elon Lages Lima, *Análise Real, Volume I*, Coleção Matemática Universitária, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 2009 (Portuguese).
- [13] ———, *Álgebra Linear*, Coleção Matemática Universitária, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1996 (Portuguese).
- [14] ———, *Curso de análise. Vol. 1*, Projeto Euclides [Euclid Project], vol. 1, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1976 (Portuguese). MR654861 (83h:26002a)
- [15] ———, *Curso de análise. Vol. 2*, Projeto Euclides [Euclid Project], vol. 13, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1981 (Portuguese). MR654862 (83h:26002b)
- [16] ———, *Espaços métricos*, Projeto Euclides [Euclid Project], vol. 4, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1977 (Portuguese). MR654506 (83d:54001)
- [17] Paul R. Halmos, *Naïve set theory*, Springer-Verlag, New York, 1974. Reprint of the 1960 edition; Undergraduate Texts in Mathematics. MR0453532 (56 #11794)
- [18] S. Kesavan, *Nonlinear functional analysis*, Texts and Readings in Mathematics, vol. 28, Hindustan Book Agency, New Delhi, 2004.
- [19] David G. Luenberger, *Introduction to linear and nonlinear programming*, Addison-Wesley, Reading, MA, 1973. Zbl 0297.90044
- [20] W. Keith Nicholson, *Elementary Linear Algebra with applications*, 2nd ed., PWS-Kent Publishing Company, Boston, 1990.
- [21] Walter Rudin, *Principles of mathematical analysis*, 3rd ed., McGraw-Hill Book Co., New York, 1976. International Series in Pure and Applied Mathematics. MR0385023 (52 #5893)
- [22] *Monkey saddle* — Wikipedia, *The Free Encyclopedia*, Wikipedia (2009).
- [23] Carl Simon and L. Blume, *Mathematics for Economists*, Norton, New York, 1994.
- [24] I. M. Singer and J. A. Thorpe, *Lecture notes on elementary topology and geometry*, Springer-Verlag, New York, 1976. Reprint of the 1967 edition; Undergraduate Texts in Mathematics. MR0413152 (54 #1273)

- [25] Terence Tao, *Analysis. I*, Texts and Readings in Mathematics, vol. 37, Hindustan Book Agency, New Delhi, 2006. MR2195040 (2006g:26002a)
- [26] ———, *Analysis. II*, Texts and Readings in Mathematics, vol. 38, Hindustan Book Agency, New Delhi, 2006. MR2195041 (2006g:26002b)
- [27] S. Viera, *Matemática Financeira*, Atlas, São Paulo.
- [28] Andrew Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. (2) **141** (1995), no. 3, 443–551, DOI 10.2307/2118559. MR1333035 (96d:11071)