

Introdução à Análise Real

Pós-graduação do LNCC ¹

Alexandre L. Madureira

LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA—LNCC, BRASIL

URL: <http://www.lncc.br/~alm>

URL: <http://www.lncc.br/~alm/cursos/analise08LNCC.html>

¹06 de outubro de 2008

PREFÁCIO. Estas notas de aula são relativas ao curso de Análise da pós-graduação em modelagem computacional do LNCC. 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 análise que, espero, sejam úteis. Na verdade, o que eu espero mesmo é apresentar o rigor matemático aos alunos, e mostrar como este deve ser utilizado em conjunto com a intuição matemática. Minha experiência diz que os alunos do LNCC têm a intuição mais desenvolvida que o rigor.

Planejo discutir os seguintes tópicos:

- Os números reais e topologia em R^n
 - Funções; Conjuntos finitos, infinitos, contáveis; Propriedades dos reais;
 - Espaços Vetoriais; Conjuntos abertos e fechados; Vizinhanças; Teorema de Bolzano-Weierstrass;
 - Conjuntos Compactos; Teorema de Heine–Borel;
- Sequências e Convergência;
 - Sequências, Subsequências;
 - Teorema de Bolzano-Weierstrass; Sequências de Cauchy
 - Sequências Contráteis e pontos fixos de contrações; Caracterização de abertos e fechados;
 - Sequências monótonas (em R); limsup, liminf;
- Funções Contínuas
 - Propriedades Locais e Globais
 - Preservação de Compacidade e Continuidade Uniforme
- Sequência de funções
 - Convergência pontual e uniforme; Trocas de limites
 - Equicontinuidade
- Diferenciabilidade
 - Funções de uma variável; Derivadas parciais; Diferenciabilidade
 - Regra da cadeia; Teorema de Taylor;
 - Teorema da função implícita e da função inversa;
 - Aplicações: Minização com restrições de igualdade e desigualdade

A referência básica é o livro *The elements of Real Analysis*, de Robert Bartle [2]. Outras referências importantes são os já clássicos [7, 12], bem como o novo [14]. Para tópicos específicos em uma dimensão, pode-se ler [3, 6, 13]. Finalmente, idéias mais abstratas são apresentadas em [8].

Conteúdo

Capítulo 1. Pré-requisitos	1
1.1. Demonstração por indução e contradição	1
1.2. Funções	2
1.3. Conjuntos finitos, infinitos, enumeráveis	3
1.4. Exercícios	4
Capítulo 2. Os números reais e o \mathbb{R}^n	5
2.1. Os números Reais	5
2.2. Espaços Vetoriais e o \mathbb{R}^n	8
2.3. Conjuntos abertos e fechados em \mathbb{R}^n	13
2.4. Celas encaixantes e o Teorema de Bolzano–Weierstrass	16
2.5. Conjuntos Compactos	18
2.6. Exercícios	20
Capítulo 3. Sequências	23
3.1. Definição e resultados preliminares	23
3.2. Subsequências e Teorema de Bolzano–Weierstrass	28
3.3. Sequências de Cauchy	30
3.4. Resultados Topológicos	32
3.5. Sequências contráteis e o método das aproximações sucessivas	34
3.6. Sequências em \mathbb{R}	36
3.7. Exercícios	40
Capítulo 4. Continuidade e Funções Contínuas	43
4.1. Propriedades locais	43
4.2. Propriedades globais	47
4.3. Funções Uniformemente Contínuas	51
4.4. Exercícios	53
Capítulo 5. Diferenciação	55
5.1. Derivada em uma dimensão	55
5.2. Teorema de Taylor em uma dimensão e Aplicações	62
5.3. Diferenciabilidade para funções de várias variáveis	64
5.4. Matriz Hessiana, Fórmula de Taylor e pontos críticos	71
5.5. Teorema da Função Implícita e da Função Inversa	76
5.6. Minimização com restrições	80
5.7. Exercícios	82
Capítulo 6. Sequência de Funções	85

6.1. Convergência Pontual	85
6.2. Convergência Uniforme	86
6.3. Equicontinuidade	88
6.4. Exercícios	89
Appendix A. Uma introdução não tão formal aos fundamentos da matemática	91
A.1. Argumentação formal	91
Appendix B. Uma introdução não tão formal à teoria de conjuntos	95
B.1. Exercícios	97
Appendix C. Listas de Exercícios	99
Appendix. Index	101
Appendix. Bibliography	105

CAPÍTULO 1

Pré-requisitos

¹ Neste capítulo, exemplificamos duas técnicas de demonstrações, e recordamos definições e notações básicas sobre conjuntos e funções.

Assumiremos aqui que as propriedades básicas de conjuntos são conhecidas. Em particular, são de grande importância os conjuntos

$$\begin{aligned}\mathbb{N} &= \{1, 2, 3, 4, \dots\} \quad (\text{números naturais}), \\ \mathbb{Z} &= \{0, 1, -1, 2, -2, 3, -3, \dots\} \quad (\text{números inteiros}), \\ \mathbb{Q} &= \{m/n : m, n \in \mathbb{Z}, n \neq 0\} \quad (\text{números racionais}),\end{aligned}$$

além é claro do conjunto dos números reais, que denotaremos por \mathbb{R} .

1.1. Demonstração por indução e contradição

Primeiro revemos aqui, através de um exemplo, como é possível demonstrar alguns fatos usando argumentos indutivos.

Considere a afirmativa

$$(1.1.1) \quad \sum_{i=1}^n i = \frac{n}{2}(n+1)$$

para todo $n \in \mathbb{N}$.

Para demonstrar que (1.1.1) vale para todos os inteiros positivos, começamos observando que para $n = 1$, a afirmativa é obviamente verdadeira. Assuma então que (1.1.1) seja verdade para $n = N^*$, i.e,

$$(1.1.2) \quad \sum_{i=1}^{N^*} i = \frac{N^*}{2}(N^* + 1).$$

Para $n = N^* + 1$ temos

$$\sum_{i=1}^{N^*+1} i = N^* + 1 + \sum_{i=1}^{N^*} i.$$

Usamos a hipótese indutiva (1.1.2) obtemos

$$\sum_{i=1}^{N^*+1} i = N^* + 1 + \frac{N^*}{2}(N^* + 1) = \frac{N^* + 1}{2}(N^* + 2),$$

e podemos concluir que (1.1.1) vale para $n = N^* + 1$, e portanto vale para todos os inteiros positivos.

¹Última Atualização: 24/04/2008

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 “provar” que

$$(1.1.3) \quad 2n \text{ é sempre ímpar para todo } n \in \mathbb{N},$$

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 (1.1.3) 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

$$(1.1.4) \quad \emptyset \subseteq A \quad \text{para qualquer conjunto } A.$$

Talvez uma demonstração “direta” não seja tão fácil. Mas suponha que (1.1.4) 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 (1.1.4) (afirmar que (1.1.4) é falso) nos leva a concluir um absurdo, e portanto (1.1.4) só pode ser verdade.

1.2. Funções

Considere A e B dois conjuntos. Uma função 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 $E \subset A$, chamamos de *imagem de E* o conjunto

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

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

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

Se $f(A) = B$ dizemos que f é *sobrejetiva* (ou simplesmente *sobre*). Dizemos que f é *injetiva* (ou *biunívoca* ou *um a um* ou 1-1) quando, dados $a, a' \in D(f)$, se $f(a) = f(a')$ então $a = a'$. Numa forma mais compacta, escrevemos que para todo $a, a' \in D(f)$ temos

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

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

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

Quando esta existir, denotamos a inversa de f por f^{-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 $f^{-1}(\mathbb{R}) = \mathbb{R}$.

1.3. Conjuntos finitos, infinitos, enumeráveis

Um conjunto B é *finito* se é vazio ou se existe uma bijeção entre B e $\{1, 2, \dots, N\}$ para algum $N \in \mathbb{N}$. Caso B não seja finito, o dizemos *infinito*. Se B é finito ou se existe uma bijeção entre B e \mathbb{N} , dizemos que B é *enumerável*.

OBSERVAÇÃO. Existe aqui uma diferença entre os termos usados em inglês no livro do Bartle [2], e suas traduções diretas em português. Seguindo Elon [6], usamos o termo *enumerável* para equivaler ao inglês *countable*. Já as expressões *enumerable* ou *denumerable* são usadas quando existe bijeção com \mathbb{N} , i.e., exclui os conjuntos finitos. Por sua vez, Rudin [12] define os termos de uma terceira forma.

EXEMPLO 1.1. $\{2, 3, 4, 5\}$ é finito, e portanto enumerável.

EXEMPLO 1.2. $P = \{2, 4, 6, \dots\}$ é enumerável pois $\phi : \mathbb{N} \rightarrow P$ definida por $\phi(n) = 2n$ é uma bijeção entre P e \mathbb{N} .

EXEMPLO 1.3. O conjunto \mathbb{Z} é enumerável pois

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\},$$

e $\phi : \mathbb{N} \rightarrow \mathbb{Z}$ dada por $\phi(i) = (-1)^i \lceil i/2 \rceil$ é uma bijeção entre \mathbb{N} e \mathbb{Z} . A função $[\cdot] : \mathbb{R} \rightarrow \mathbb{Z}$ é tal que $[x]$ é a parte inteira de x , i.e., o maior inteiro menor ou igual a x .

EXEMPLO 1.4. \mathbb{Q} é enumerável pela “contagem diagonal”:

$$\begin{array}{cccccc} 0, & & & & & & \\ 1, & -1, & 2, & -2, & 3, & -3, & \dots \\ 1/2, & -1/2, & 2/2, & -2/2, & 3/2, & -3/2, & \dots \\ 1/3, & -1/3, & 2/3, & -2/3, & 3/3, & -3/3, & \dots \\ \vdots & & & & & & \end{array}$$

e podemos contar pois

$$\mathbb{Q} = \left\{ 0, 1, -1, \frac{1}{2}, 2, -\frac{1}{2}, \frac{1}{3}, -2, -\frac{1}{3}, \dots \right\}.$$

EXEMPLO 1.5. O conjunto de números reais \mathbb{R} não é enumerável. Para mostrar isto, usaremos uma demonstração por contradição. Mostraremos na verdade que $I = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$ não é enumerável. Usando a base decimal, todo elemento $x \in I$ pode ser representado por $x = 0, a_1 a_2 a_3 \dots$, onde $a_i \in \{0, \dots, 9\}$. Assuma que I é enumerável. Então existe uma enumeração $x_1, x_2, \dots, x_n, \dots$ dos elementos de I tal que

$$\begin{array}{l} x_1 = 0, a_{11} a_{12} a_{13} \dots, \\ x_2 = 0, a_{21} a_{22} a_{23} \dots, \\ x_3 = 0, a_{31} a_{32} a_{33} \dots, \\ \dots, \end{array}$$

onde $a_{ij} \in \{0, \dots, 9\}$. Seja agora $y = 0, b_1 b_2 b_3 \dots$ onde

$$b_i = \begin{cases} 0 & \text{se } a_{ii} \in \{1, \dots, 9\} \\ 1 & \text{se } a_{ii} = 0. \end{cases}$$

Logo $y \in I$ mas $y \neq x_n$ para todo $n \in \mathbb{N}$. Isto contradiz a afirmação que $x_1, x_2, \dots, x_n, \dots$ é uma enumeração dos elementos de I . Portanto, I não é enumerável.

1.4. Exercícios

EXERCÍCIO 1.1. Mostre por indução que $n < 2^n$ para todo $n \in \mathbb{N}$.

EXERCÍCIO 1.2. Mostre por indução a desigualdade de Bernoulli: se $x > -1$, então $(1 + x)^n \geq 1 + nx$ para todo $n \in \mathbb{N}$.

EXERCÍCIO 1.3. Mostre usando contradição que $\sqrt{2}$ não é racional.

EXERCÍCIO 1.4. Mostre usando contradição que se p_1, \dots, p_n são todos os números primos menores ou iguais a p_n , então $p_1 \times \dots \times p_n + 1$ não é divisível por p_i para nenhum $i \in \{1, \dots, n\}$.

EXERCÍCIO 1.5. Mostre usando contradição que existem infinitos números primos.

EXERCÍCIO 1.6. Mostre que uma função tem inversa se e somente se ela é uma bijeção.

EXERCÍCIO 1.7. Sejam A e B conjuntos enumeráveis. Mostre que o produto cartesiano $A \times B$ é enumerável. Conclua assim que \mathbb{Z} enumerável implica em \mathbb{Q} enumerável.

EXERCÍCIO 1.8. Para $i \in \mathbb{N}$, seja A_i conjunto infinito enumerável. Mostre que o produto cartesiano infinito $\prod_{i=1}^{\infty} A_i$ não é enumerável.

CAPÍTULO 2

Os números reais e o \mathbb{R}^n

¹ Neste capítulo, falaremos sobre números reais. Suporemos aqui que os números reais e as operações neles definidas são bem definidos e “existem”, sem entrar em detalhes sobre a construção deste corpo. A idéia é apenas apresentar propriedades que os reais satisfazem. A seguir, falaremos sobre abertos e fechados nos reais.

2.1. Os números Reais

2.1.1. Propriedades dos Reais. Para discutir uma importante propriedade dos números reais, introduziremos o conceito de cotas. Para tal usaremos o fato de que \mathbb{R} é *ordenado*, i.e., existe uma relação de ordem denotada por $<$ indicando se um elemento é *menor* que outro. Usaremos também os símbolos $>$, \leq , \geq , indicando se um elemento é maior, menor ou igual, maior ou igual, respectivamente.

DEFINIÇÃO 2.1.1. *Considere um conjunto $A \subset \mathbb{R}$. Dizemos que $c^* \in \mathbb{R}$ é cota superior de A se $a \leq c^*$ para todo $a \in A$. Analogamente, dizemos que $c_* \in \mathbb{R}$ é cota inferior de A se $c_* \leq a$ para todo $a \in A$. Se um conjunto tem cota superior dizemos que ele é limitado por cima ou superiormente. Se um conjunto tem cota inferior dizemos que ele é limitado por baixo ou inferiormente. Se um conjunto tem cota superior e inferior, dizemos que ele é limitado.*

Note que nem todos os conjuntos possuem cotas superiores e/ou inferiores. Por exemplo $\mathbb{N} \subset \mathbb{R}$ não possui cota superior, apesar de possuir cota inferior. Segue-se da definição que se um conjunto possui cota superior, então ele possui infinitas cotas superiores:

$$c^* \text{ cota superior de } A \implies c^* + 1 \text{ cota superior de } A.$$

Observação análoga vale para as cotas inferiores.

EXEMPLO 2.1. O conjunto $\mathbb{R}^- = \{x \in \mathbb{R} : x < 0\}$ é limitado superiormente mas não inferiormente. De fato qualquer número não negativo é cota superior de \mathbb{R}^- , pois se $b \geq 0$, então $x \in \mathbb{R}^-$ implica que $x < 0 \leq b$. Por outro lado, nenhum número $a \in \mathbb{R}$ pode ser cota inferior pois sempre existe $y \in \mathbb{R}^-$ tal que $y < a$. Concluímos portanto que \mathbb{R}^- não é limitado.

EXEMPLO 2.2. Usando argumentos como acima, vemos que \mathbb{R} não é limitado nem superiormente nem inferiormente.

EXEMPLO 2.3. Seja $I = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$. Então qualquer número $b \geq 1$ é cota superior de I , e todo número $a \leq 0$ é cota inferior de I . De fato, nestes casos teríamos $a \leq x \leq b$ para todo $x \in I$. Logo, por definição, I é limitado.

¹Última Atualização: 24/04/2008

EXEMPLO 2.4. Note que qualquer número é cota inferior e superior do conjunto vazio.

DEFINIÇÃO 2.1.2. Se um conjunto A é não vazio e limitado superiormente, chamamos de supremo de A ou simplesmente $\sup A$ a menor de suas cotas superiores. Analogamente, se um conjunto A é não vazio e limitado por baixo, chamamos de ínfimo de A ou simplesmente $\inf A$ a maior de suas cotas inferiores.

Logo, se $s^* = \sup A$, então

- (1) $a \leq s^*$ para todo $a \in A$.
- (2) Se existe $v \in \mathbb{R}$ tal que $a \leq v$ para todo $a \in A$, então $s^* \leq v$.

OBSERVAÇÃO. Segue-se da definição a unicidade do supremo e do ínfimo, se estes existirem, ver Exercício 2.4.

O resultado a seguir nos dá uma forma equivalente para determinar o supremo de um conjunto.

LEMA 2.1.3. Seja A não vazio e s^* cota superior de A . Então $s^* = \sup A$ se e somente se para todo $\epsilon > 0$ existir $a_\epsilon \in A$ tal que $s^* - \epsilon < a_\epsilon$.

DEMONSTRAÇÃO. (\Rightarrow) Seja $s^* = \sup A$ e $\epsilon > 0$. Como $s^* - \epsilon < s^*$, então $s^* - \epsilon$ não é cota superior de A . Logo, existe um elemento $a_\epsilon \in A$ tal que $a_\epsilon > s^* - \epsilon$.

(\Leftarrow) Seja s^* cota superior de A . Suponha que para todo ϵ existe $a_\epsilon \in A$ tal que $s^* - \epsilon < a_\epsilon$. Vamos então mostrar que $s^* = \sup A$.

Seja c^* cota superior de A com $c^* \neq s^*$. Se $c^* < s^*$, definimos $\epsilon = s^* - c^*$ e então $\epsilon > 0$ e existe $a_\epsilon \in A$ tal que $a_\epsilon > s^* - \epsilon = c^*$. Isto é uma contradição com o fato de c^* ser cota superior. Logo temos obrigatoriamente $c^* > s^*$, e s^* é a menor das cotas superiores, i.e., $s^* = \sup A$. \square

EXEMPLO 2.5. $I = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$ tem $\sup I = 1$ e $\inf I = 0$. Note que $\sup I \in I$ e $\inf I \in I$.

EXEMPLO 2.6. $U = \{x \in \mathbb{R} : 0 < x < 1\}$ tem $\sup U = 1$ e $\inf U = 0$. Note que neste caso $\sup I \notin U$ e $\inf I \notin U$.

Uma propriedade *fundamental* dos reais, que o distingue por exemplo dos racionais, é dada a seguir.

Propriedade do supremo de \mathbb{R} : Todo conjunto não vazio em \mathbb{R} limitado superiormente tem um supremo em \mathbb{R} .

Da propriedade acima, obtemos o seguinte resultado.

LEMA 2.1.4 (Propriedade arquimediana). Para todo $x \in \mathbb{R}$, existe $n \in \mathbb{N}$ tal que $n > x$.

DEMONSTRAÇÃO. (Por contradição.) Suponha que não existe tal número n . Portanto, x é cota superior de $\mathbb{N} \subset \mathbb{R}$. Pela Propriedade do supremo de \mathbb{R} , então \mathbb{N} tem um supremo s . Logo existe $m \in \mathbb{N}$ tal que $s - 1 < m$. Mas então, $s < m + 1$, uma contradição, pois $m + 1 \in \mathbb{N}$ e s deveria ser cota superior de \mathbb{N} . \square

OBSERVAÇÃO. Densidade de \mathbb{Q} em \mathbb{R} : Se $x, y \in \mathbb{R}$ e $x < y$, então existe $r \in \mathbb{Q}$ tal que $x < r < y$. Da mesma forma, existe $r \in \mathbb{R} \setminus \mathbb{Q}$ tal que $x < r < y$.

2.1.2. Valor absoluto e Intervalos. 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 2.7. Por definição $|5| = 5$, e $|-5| = -(-5) = 5$.

LEMA 2.1.5. 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 2.1.6 (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 2.1.5 temos que $|a + b| \leq |a| + |b|$, como queríamos demonstrar. □

Dentre os mais importantes conjuntos reais estão os intervalos. Sejam $a, b \in \mathbb{R}$. Chamaremos de *intervalo* quaisquer conjuntos dos seguintes tipos:

- (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) $(-\infty, +\infty) = \mathbb{R}$
- (10) \emptyset

A definição de alguns intervalos particulares é imediata usando-se o módulo:

$$(a - d, a + d) = \{x \in \mathbb{R} : |x - a| < d\}, \quad [a - d, a + d] = \{x \in \mathbb{R} : |x - a| \leq d\},$$

Uma importante propriedade dos números reais, intrinsecamente ligada à sua própria definição, é dada por interseções de intervalos encaixantes, noção que discutimos a seguir.

DEFINIÇÃO 2.1.7. Dizemos que uma sequência de intervalos I_n é encaixante se

$$I_1 \supset I_2 \supset I_3 \supset \cdots \supset I_n \supset \cdots$$

Nos dois exemplos abaixo, ilustramos o fato de que interseções de intervalos encaixantes podem ser vazias ou não. Entretanto, quando os intervalos forem fechados, o Teorema dos intervalos encaixantes abaixo garante que estas interseções são sempre não vazias.

EXEMPLO 2.8. Se $I_n = [0, 1/n]$ então $\bigcap_{n=1}^{\infty} I_n = \{0\}$. De fato, $0 \in I_n$ para todo $n \in \mathbb{N}$ e portanto $0 \in \bigcap_{n=1}^{\infty} I_n$. Por outro lado, para $x \in \mathbb{R}$ não nulo a Propriedade arquimediana (Lema 2.1.4) garante a existência de $n \in \mathbb{N}$ tal que $x \notin I_n$. Logo $x \notin \bigcap_{n=1}^{\infty} I_n$.

EXEMPLO 2.9. Usando novamente a Propriedade arquimediana (Lema 2.1.4) temos que se $I_n = (0, 1/n)$ então $\bigcap_{n=1}^{\infty} I_n = \emptyset$.

TEOREMA 2.1.8 (Teorema dos intervalos encaixantes). *Sejam $I_n = [a_n, b_n]$ intervalos fechados, limitados, não vazios e encaixantes. Então existe $\xi \in \mathbb{R}$ tal que $\xi \in \bigcap_{n=1}^{\infty} I_n$. Além disso, se $\inf\{b_n - a_n : n \in \mathbb{N}\} = 0$, então ξ é o único elemento da interseção.*

DEMONSTRAÇÃO. Segue-se das hipóteses que para todo $n \in \mathbb{N}$ temos

$$(2.1.1) \quad a_{n+1} \geq a_n, \quad b_{n+1} \leq b_n, \quad a_n \leq b_n.$$

Temos $b_1 \geq a_n$ para todo n pois $I_n \subset I_1$. Seja $\xi = \sup\{a_n : n \in \mathbb{N}\}$. Logo $\xi \geq a_n$ para todo n . Queremos mostrar agora que $\xi \leq b_n$ para todo n . Suponha o contrário, i.e., que existe $b_k < \xi$ para algum k . Logo $b_k < a_m$ para algum m . Seja $p = \max\{k, m\}$. Então $a_p \geq a_m > b_k \geq b_p$, uma contradição com (2.1.1). Logo $a_n \leq \xi \leq b_n$ para todo $n \in \mathbb{N}$ e portanto $\xi \in I_n$ para todo $n \in \mathbb{N}$.

Supondo agora que $\inf\{b_n - a_n : n \in \mathbb{N}\} = 0$, definimos $\eta = \inf\{b_n : n \in \mathbb{N}\}$. Então $\eta \geq a_n$ para todo $n \in \mathbb{N}$ e $\eta \geq \xi$. Como $0 \leq \eta - \xi \leq b_n - a_n$ para todo $n \in \mathbb{N}$, temos $\eta = \xi$ pois $\inf\{b_n - a_n : n \in \mathbb{N}\} = 0$. Finalmente, seja $p \in \bigcap_{n=1}^{\infty} I_n$. Como $p \geq \xi = \eta$ e $p \leq \eta = \xi$, então $p = \xi = \eta$ é o único ponto em $\bigcap_{n=1}^{\infty} I_n$. \square

2.2. Espaços Vetoriais e o \mathbb{R}^n

O exemplo mais comum e intuitivo de espaço vetorial é o \mathbb{R}^n . 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 2.2.1. *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 2.2.2. 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 2.2.1, 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 2.2.1. \square

EXEMPLO 2.10. O espaço das matrizes $m \times n$ denotado por $\mathbb{R}^{m \times n}$ é um espaço vetorial com a definição usual de soma de matrizes e multiplicação por escalar.

O exemplo usual de espaço vetorial é o \mathbb{R}^n , como definido abaixo.

DEFINIÇÃO 2.2.3. 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.

Para $i \in \{1, \dots, n\}$ seja \mathbf{e}_i onde o vetor com a i -ésima coordenada valendo um e as demais coordenadas com valor 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 $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n$. Definimos então a matriz coluna $\vec{\mathbf{x}} \in \mathbb{R}^{n \times 1}$ dada por

$$(2.2.1) \quad \vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

como sendo as coordenadas de \mathbf{x} na base canônica.

Note que existe uma identificação natural dos pontos em \mathbb{R}^n com suas coordenadas na base canônica. Usaremos neste texto a notação indicada acima. Para cada $\mathbf{x} \in \mathbb{R}^n$, indicaremos por $\vec{\mathbf{x}} \in \mathbb{R}^{n \times 1}$ a matriz coluna das coordenadas na base canônica como em (2.2.1).

EXEMPLO 2.11. 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.

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

DEFINIÇÃO 2.2.4. *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 $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$

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 2.12. 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 2.13. Em \mathbb{R}^2 , a operação

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

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 2.14. Considere o espaço vetorial das funções contínuas em $[0, 1]$, com as operações de multiplicação por escalar e soma como no Exemplo 2.11. Então a operação dada pela integral de Riemann

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

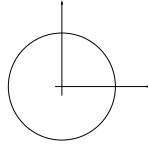
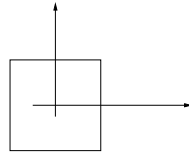
define um produto interno deste espaço.

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 2.2.5. *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*.

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

EXEMPLO 2.15. 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 2.16. 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 2.2.6. *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

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

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 (2.2.2). 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 (2.2.2) 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 que a igualdade $|\mathbf{x} \cdot \mathbf{y}| = \|\mathbf{x}\|\|\mathbf{y}\|$ vale se e somente se $\mathbf{x} = \alpha\mathbf{y}$ para algum $\alpha \in \mathbb{R}$.

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 $\mathcal{L}(V_1, V_2)$ o espaço das aplicações lineares $T : V_1 \rightarrow V_2$ para as quais existe $M \in \mathbb{R}$ tal que

$$\|T\mathbf{x}\|_{V_2} \leq M\|\mathbf{x}\|_{V_1},$$

É possível definir operações canônicas de multiplicação por escalar e soma em $\mathcal{L}(V_1, V_2)$ de tal forma que este seja um espaço vetorial, ver exercício 2.12. Se V_1 for espaço normado com norma $\|\cdot\|_{V_1}$, e V_2 for espaço normado com norma $\|\cdot\|_{V_2}$, é possível induzir uma norma em $\mathcal{L}(V_1, V_2)$, que é chamada *norma dos operadores* (exercício 2.13), dada por

$$(2.2.3) \quad \|T\|_{\mathcal{L}(V_1, V_2)} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|T\mathbf{x}\|_{V_2}}{\|\mathbf{x}\|_{V_1}} \quad \text{para } T \in \mathcal{L}(V_1, V_2).$$

Neste caso, para $\mathbf{y} \in V_1$, sempre vale a desigualdade

$$\|T\mathbf{y}\|_{V_2} \leq \|T\|_{\mathcal{L}(V_1, V_2)}\|\mathbf{y}\|_{V_1}.$$

De fato, para $\mathbf{y} = \mathbf{0}$ vale a igualdade. Para $\mathbf{y} \neq \mathbf{0}$ temos que

$$\|T\mathbf{y}\|_{V_2} = \frac{\|T\mathbf{y}\|_{V_2}}{\|\mathbf{y}\|_{V_1}}\|\mathbf{y}\|_{V_1} \leq \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|T\mathbf{x}\|_{V_2}}{\|\mathbf{x}\|_{V_1}}\|\mathbf{y}\|_{V_1} = \|T\|_{\mathcal{L}(V_1, V_2)}\|\mathbf{y}\|_{V_1}.$$

O exemplo típico de transformação linear é dada por matrizes, da seguinte forma. Seja A uma matriz com n linhas e m colunas, e $T_A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ definida por

$$T_A(\mathbf{x}) = \mathbf{y}, \quad \text{onde } \vec{\mathbf{y}} = A\vec{\mathbf{x}},$$

onde usamos a notação da Definição 2.2.3. Neste caso denotamos a norma de operadores $\|T_A\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)}$ simplesmente por $\|A\|$. Vale portanto a importante desigualdade

$$(2.2.4) \quad \|T_A(\mathbf{x})\|_{\mathbb{R}^n} \leq \|A\|\|\mathbf{x}\|_{\mathbb{R}^m} \quad \text{para todo } \mathbf{x} \in \mathbb{R}^m,$$

onde $\|\cdot\|_{\mathbb{R}^n}$ e $\|\cdot\|_{\mathbb{R}^m}$ são normas em \mathbb{R}^n e \mathbb{R}^m respectivamente.

2.3. Conjuntos abertos e fechados em \mathbb{R}^n

Como já foi comentado, para definirmos conjuntos abertos e fechados no \mathbb{R}^n , utilizaremos o conceito de distância definida por uma norma. No caso, escolhemos a norma definida por

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

É importante ressaltar que esta escolha de norma não implica em nenhuma “escolha de topologia”, pois em espaços de dimensão finita, todas as normas são *equivalentes*, i.e., se $\|\cdot\|$ define uma norma em \mathbb{R}^n , então existem constantes c e C tais que

$$c\|\mathbf{x}\| \leq \|\mathbf{x}\| \leq C\|\mathbf{x}\|,$$

para todo $\mathbf{x} \in \mathbb{R}^n$. As constantes c e C dependem apenas de n (dimensão do espaço).

Para definirmos o que é um conjunto aberto necessitamos das chamadas bolas em \mathbb{R}^n . Dizemos que a *bola aberta de raio r e centro \mathbf{x}* é dada por

$$B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < r\}.$$

De forma similar, chamamos de *bola fechada de raio r e centro \mathbf{x}* , e de *esfera de raio r e centro \mathbf{x}* os conjuntos

$$\{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| \leq r\}, \quad \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| = r\}.$$

EXEMPLO 2.17. Em uma dimensão, para $x \in \mathbb{R}$ temos $B_r(x) = (x - r, x + r)$. A bola fechada de raio r e centro em x é dada por $[x - r, x + r]$, e a esfera de raio r e centro x é simplesmente o conjunto $\{x - r, x + r\}$.

Podemos agora definir conjuntos abertos em \mathbb{R}^n .

DEFINIÇÃO 2.3.1. *Um conjunto $A \subset \mathbb{R}^n$ é aberto em \mathbb{R}^n se para todo $\mathbf{x} \in A$ existe $\epsilon > 0$ tal que $B_\epsilon(\mathbf{x}) \subset A$. Em geral chamaremos conjuntos abertos simplesmente de abertos.*

EXEMPLO 2.18. \emptyset é aberto por “vacuidade”.

EXEMPLO 2.19. \mathbb{R} é aberto nos reais pois para todo $x \in \mathbb{R}$, temos $B_1(x) = (x - 1, x + 1) \subset \mathbb{R}$. Note que tomamos $\epsilon = 1$. Da mesma forma, \mathbb{R}^n também é aberto pois para todo $\mathbf{x} \in \mathbb{R}^n$, tem-se $B_1(\mathbf{x}) \subset \mathbb{R}^n$.

EXEMPLO 2.20. O conjunto $I = [0, 1] \subset \mathbb{R}$ não é aberto. De fato $0 \in I$, e para todo $\epsilon > 0$, a bola $B_\epsilon(0) = (-\epsilon, \epsilon) \not\subset I$, pois, por exemplo, $-\epsilon/2 \in B_\epsilon(0)$ mas $-\epsilon/2 \notin I$.

EXEMPLO 2.21. O conjunto $(0, 1)$ é aberto em \mathbb{R} . De fato para qualquer $x \in (0, 1)$, seja $\epsilon = \min\{x/2, (1 - x)/2\}$. Então $B_\epsilon(x) = (x - \epsilon, x + \epsilon) \subset (0, 1)$. De forma análoga, $B_1(\mathbf{0}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| < 1\}$ é aberto em \mathbb{R}^n .

EXEMPLO 2.22. O subconjunto de \mathbb{R}^2 dado por

$$A = (0, 1) \times \{0\} = \{(x, 0) \in \mathbb{R}^2 : x \in (0, 1)\}$$

não é aberto em \mathbb{R}^2 . De fato, seja $x \in (0, 1)$ e $\mathbf{x} = (x, 0) \in A$. Para todo $\epsilon > 0$ temos que $B_\epsilon(\mathbf{x}) \not\subset A$, pois, por exemplo, $(x, -\epsilon/2) \in B_\epsilon(\mathbf{x})$ mas $(x, -\epsilon/2) \notin A$. Compare com o exemplo 2.21.

LEMA 2.3.2. Duas propriedades fundamentais de conjuntos abertos são

- (1) A união arbitrária de abertos é aberta.
- (2) A interseção finita de abertos é aberta.

DEMONSTRAÇÃO. Para mostrar (1), seja $\{G_\lambda : \lambda \in \Lambda\}$ uma família arbitrária de abertos, e seja $G = \cup_{\lambda \in \Lambda} G_\lambda$ e $\mathbf{x} \in G$. Então $\mathbf{x} \in G_{\lambda_0}$ para algum $\lambda_0 \in \Lambda$. Como G_{λ_0} é aberto, então existe $\epsilon > 0$ tal que $B_\epsilon(\mathbf{x}) \subset G_{\lambda_0}$. Logo $B_\epsilon(\mathbf{x}) \subset \cup_{\lambda \in \Lambda} G_\lambda = G$ e então G é aberto.

Para mostrar (2), sejam G_1, \dots, G_k abertos e $G = \cap_{i=1}^k G_i$. Seja $\mathbf{x} \in G$. Logo $\mathbf{x} \in G_i$ para todo $i \in \mathbb{N}$. Como G_i é aberto, seja $\epsilon_i > 0$ tal que $B_{\epsilon_i}(\mathbf{x}) \subset G_i$. Definindo $\epsilon = \min\{\epsilon_1, \dots, \epsilon_k\}$, temos $\epsilon > 0$ e $B_\epsilon(\mathbf{x}) \subset G_1 \cap \dots \cap G_k = G$. Logo G é aberto. \square

EXEMPLO 2.23. Em uma dimensão, seja $I_n = (0, 1 - 1/n)$ onde $n \in \mathbb{N}$. Então I_n é aberto e $\cup_{n=1}^\infty I_n = (0, 1)$ também o é.

EXEMPLO 2.24. Interseção infinita de abertos pode não ser aberta. Por exemplo, $G_n = (0, 1 + 1/n)$ é aberto em \mathbb{R} , ao contrário de $\cap_{n=1}^\infty G_n = (0, 1]$. Da mesma forma, $B_{1/n}(\mathbf{0})$ é aberto, mas $\cap_{n=1}^\infty B_{1/n}(\mathbf{0}) = \{\mathbf{0}\}$ não é aberto. Qual o passo da demonstração do Lema 2.3.2 que não seria correto para este exemplo?

Um outro importante conceito é o de conjuntos fechados, e temos a seguinte definição.

DEFINIÇÃO 2.3.3. Um conjunto $F \subset \mathbb{R}^n$ é fechado em \mathbb{R}^n se seu complemento

$$\mathcal{C}(F) = \mathbb{R}^n \setminus F = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \notin F\}$$

é aberto.

Para mostrar que um conjunto G é aberto em \mathbb{R}^n , basta mostrar que para todo $\mathbf{x} \in G$ existe $\epsilon > 0$ tal que $B_\epsilon(\mathbf{x}) \subset G$. Para mostrar que F é fechado, basta mostrar que para todo $\mathbf{x} \notin F$ existe $\epsilon > 0$ tal que $B_\epsilon(\mathbf{x}) \cap F = \emptyset$.

EXEMPLO 2.25. $[0, 1]$ é fechado em \mathbb{R} pois $\mathcal{C}([0, 1]) = (-\infty, 0) \cup (1, \infty)$ é aberto em \mathbb{R} .

EXEMPLO 2.26. $(0, 1]$ não é aberto nem fechado em \mathbb{R} .

EXEMPLO 2.27. Os conjuntos \mathbb{R}^n e \emptyset são fechados em \mathbb{R}^n , pois seus complementares $\mathcal{C}(\emptyset) = \mathbb{R}^n$ e $\mathcal{C}(\mathbb{R}^n) = \emptyset$ são abertos em \mathbb{R}^n .

EXEMPLO 2.28. Para todo $\mathbf{x} \in \mathbb{R}^n$ e $r > 0$, as esferas e as bolas fechadas de centro \mathbf{x} e raio r são conjuntos fechados em \mathbb{R}^n .

COROLÁRIO 2.3.4. Como consequência do Lema 2.3.2 temos:

- (1) A interseção arbitrária de fechados é fechada.
- (2) A união finita de fechados é fechada.

DEMONSTRAÇÃO. Utilizamos nas duas demonstrações a regra de *De Morgan* (B.0.9).

- (1) Seja $\{F_\lambda : \lambda \in \Lambda\}$ uma coleção de fechados em \mathbb{R}^n , e seja $F = \cap_{\lambda \in \Lambda} F_\lambda$. Então $\mathcal{C}(F) = \cup_{\lambda \in \Lambda} \mathcal{C}(F_\lambda)$ é uma união de abertos. Logo $\mathcal{C}(F)$ é aberto e, por definição, F é fechado.
- (2) Se F_1, \dots, F_n são fechados em \mathbb{R}^n e $F = F_1 \cup \dots \cup F_n$, então $\mathcal{C}(F) = \mathcal{C}(F_1) \cap \dots \cap \mathcal{C}(F_n)$. Como a interseção finita de abertos é aberta, e $\mathcal{C}(F_i)$ são abertos, então $\mathcal{C}(F)$ é aberto. Logo F é fechado.

\square

EXEMPLO 2.29. $F_n = [1/n, 1]$ é fechado em \mathbb{R} , mas $\cup_{n=1}^\infty F_n = (0, 1]$ não o é.

2.3.1. Outras caracterizações de conjuntos abertos e fechados. Outras noções que podem ser úteis quando precisamos caracterizar conjuntos abertos ou fechados vêm a seguir.

DEFINIÇÃO 2.3.5. *Sejam $\mathbf{x} \in \mathbb{R}^n$, e $A \subset \mathbb{R}^n$. Dizemos então que*

- (1) *uma vizinhança aberta de \mathbf{x} é um conjunto aberto que contenha \mathbf{x} .*
- (2) *\mathbf{x} é ponto interior de A se existe uma vizinhança aberta de \mathbf{x} contida em A .*
- (3) *\mathbf{x} é ponto de fronteira de A se toda vizinhança aberta de \mathbf{x} contém ponto de A e do complementar $\mathcal{C}(A)$.*
- (4) *\mathbf{x} é ponto exterior de A se existe uma vizinhança aberta de \mathbf{x} contida em $\mathcal{C}(A)$.*

Observe que das definições acima, dados um ponto $\mathbf{x} \in \mathbb{R}^n$, e um conjunto $A \subset \mathbb{R}^n$, então \mathbf{x} é ponto interior, exterior, ou de fronteira de A , sendo as opções mutualmente exclusivas.

EXEMPLO 2.30. Seja $U = (0, 1)$. Se $a \in U$, então U é vizinhança aberta de a . De forma análoga, qualquer conjunto aberto é vizinhança aberta de seus pontos.

As seguintes propriedades podem ser usadas para se definir se um conjunto é ou não aberto.

LEMA 2.3.6. Seja $G \subset \mathbb{R}^n$. As afirmativas abaixo são equivalentes.

- (1) G é aberto.
- (2) Todo ponto de G é ponto interior.
- (3) G não contém nenhum de seus pontos de fronteira.

DEMONSTRAÇÃO. ((1) \Rightarrow (2)) Supondo (1), seja $\mathbf{x} \in G$. Como por hipótese G é aberto, temos que G é vizinhança aberta de \mathbf{x} . Logo \mathbf{x} é ponto interior de G . Como \mathbf{x} é arbitrário, obtemos (2).

((2) \Rightarrow (3)) Se todo ponto de G é interior, então nenhum de seus pontos é de fronteira.

((3) \Rightarrow (1)) Suponha que G não contém nenhum de seus pontos de fronteira. Se G é vazio, então é aberto. Suponha então que G é não vazio. Seja $\mathbf{x} \in G$. Como G não contém pontos de fronteira, existe vizinhança aberta U de \mathbf{x} tal que $U \subset G$. Logo G é aberto. \square

COROLÁRIO 2.3.7. Seja $F \subset \mathbb{R}^n$. Então F é fechado se e somente se contém todos os seus pontos de fronteira.

Finalmente concluímos esta seção com o conceito de ponto de acumulação.

DEFINIÇÃO 2.3.8. *Um ponto $\mathbf{x} \in \mathbb{R}^n$ é um ponto de acumulação de $A \subset \mathbb{R}^n$ se toda vizinhança aberta de \mathbf{x} contém pelo menos um ponto de A diferente de \mathbf{x} .*

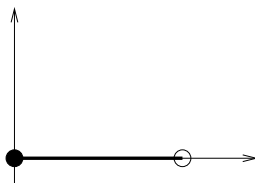
Em uma dimensão, um ponto $x \in \mathbb{R}$ é de acumulação de $A \subset \mathbb{R}$ se para todo $\epsilon > 0$ existir $\xi \neq x$ tal que $\xi \in (x - \epsilon, x + \epsilon) \cap A$.

Note que um ponto pode ser de acumulação de um certo conjunto mesmo sem pertencer a este conjunto. De fato veremos vários exemplos abaixo em que tal situação ocorre.

EXEMPLO 2.31. Se $A = (0, 1)$, então todo ponto em $[0, 1]$ é ponto de acumulação de A .

EXEMPLO 2.32. O conjunto $\mathbb{N} \times \mathbb{N} \subset \mathbb{R}^2$ não tem ponto de acumulação.

EXEMPLO 2.33. O único ponto de acumulação de $\{1, 1/2, 1/3, 1/4, \dots, 1/n, \dots\}$ é o 0.

FIG. 3. Conjunto A .

EXEMPLO 2.34. $([0, 1] \times [0, 1]) \cap \mathbb{Q}^2$ tem como pontos de acumulação o conjunto $[0, 1] \times [0, 1]$.

EXEMPLO 2.35. Seja $A \subset \mathbb{R}$ limitado superiormente e $u = \sup A$. Se $u \notin A$, então u é ponto de acumulação de A , pois para todo $\epsilon > 0$ existe $x \in A$ tal que $x \in (u - \epsilon, u + \epsilon)$.

Uma caracterização útil de fechados utiliza o conceito de pontos de acumulação, como o resultado a seguir indica.

TEOREMA 2.3.9. *Um subconjunto de \mathbb{R}^n é fechado se e somente se contém todos os seus pontos de acumulação.*

DEMONSTRAÇÃO. (\Rightarrow) (Por contradição) Seja F um fechado em \mathbb{R}^n , e \mathbf{x} ponto de acumulação de F . Temos que mostrar que $\mathbf{x} \in F$. De fato, se $\mathbf{x} \notin F$, então $\mathbf{x} \in \mathcal{C}(F)$. Mas como $\mathcal{C}(F)$ é aberto, então existe $\epsilon > 0$ tal que $B_\epsilon(\mathbf{x}) \subset \mathcal{C}(F)$. Logo $B_\epsilon(\mathbf{x}) \cap F = \emptyset$ e \mathbf{x} não é ponto de acumulação de F , uma contradição. Portanto $\mathbf{x} \in F$.

(\Leftarrow) Supomos agora que F contém todos os seus pontos de acumulação. Considere então um ponto $\mathbf{y} \in \mathcal{C}(F)$. Então \mathbf{y} não é ponto de acumulação de F , e portanto existe $\epsilon > 0$ tal que $B_\epsilon(\mathbf{y}) \subset \mathcal{C}(F)$. Logo $\mathcal{C}(F)$ é aberto, e concluímos que F é fechado. \square

EXEMPLO 2.36. Em \mathbb{R}^2 o conjunto

$$A = [0, 1) \times \{0\} = \{(x, 0) \in \mathbb{R}^2 : x \in [0, 1)\}$$

representado na figura 3, não é nem aberto nem fechado. Para mostrar que A não é fechado, considere os pontos $x_n \in A$ dados por $\mathbf{x}_n = (1 - 1/n, 0)$. Então $\|\mathbf{x}_n - (1, 0)\| = 1/n$ e $(1, 0)$ é ponto de acumulação. Como $(1, 0) \notin A$, então A não contém um de seus pontos de acumulação, logo A não é fechado. Para mostrar que A não é aberto, note que *toda* bola de raio ϵ e centro em $(0, 0)$ contém pontos em A e no complementar de A . Compare com o exemplo 2.22.

2.4. Celas encaixantes e o Teorema de Bolzano–Weierstrass

Uma importante e imediata generalização do Teorema dos intervalos encaixantes (Teorema 2.1.8) para o \mathbb{R}^n é descrita a seguir. Antes de mais nada, chamamos de cela fechada ao conjunto dado por

$$[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] = \{\mathbf{x} \in \mathbb{R}^n : a_i \leq x_i \leq b_i, i = 1, \dots, n\},$$

onde $a_i \leq b_i$ para $i = 1, \dots, n$. Dizemos que uma sequência de celas (C_k) é encaixante se

$$C_1 \supset C_2 \supset C_3 \supset \cdots \supset C_k \supset \cdots .$$

Finalmente, dizemos que um conjunto $A \subset \mathbb{R}^n$ é limitado se existe uma constante c tal que para todo $\mathbf{x} \in A$ tem-se $\|\mathbf{x}\| \leq c$.

TEOREMA 2.4.1 (Teorema das celas encaixantes). *Seja (C_k) uma sequência de celas fechadas, limitadas, não vazias e encaixantes. Então existe $\xi \in \mathbb{R}^n$ tal que $\xi \in \bigcap_{i=1}^{\infty} C_i$.*

DEMONSTRAÇÃO. Para $k \in \mathbb{N}$, e $a_{k,i} \leq b_{k,i}$ para $i = 1, \dots, n$, suponha que

$$C_k = \{\mathbf{x} \in \mathbb{R}^n : a_{k,i} \leq x_i \leq b_{k,i}\}.$$

Fixando $i \in \{1, \dots, n\}$ e aplicando o Teorema dos intervalos encaixantes (Teorema 2.1.8) para $I_{k,i} = [a_{k,i}, b_{k,i}]$, temos que existe $\xi_i \in \bigcap_{k=1}^{\infty} I_{k,i}$. portanto $\xi = (\xi_1, \dots, \xi_n) \in \bigcap_{k=1}^{\infty} C_k$. \square

Uma importante aplicação do Teorema das celas encaixantes é na demonstração do resultado a seguir, o Teorema de Bolzano–Weierstrass. Damos a demonstração em uma dimensão, e a demonstração em \mathbb{R}^n é análoga. Uma outra maneira de se mostrar este resultado é baseada na noção de compacidade que discutiremos a seguir, ver o exercício 2.31.

TEOREMA 2.4.2 (Bolzano–Weierstrass no \mathbb{R}^n). *Todo subconjunto de \mathbb{R}^n infinito e limitado tem pelo menos um ponto de acumulação.*

A seguir damos uma idéia da demonstração em uma dimensão, antes de proceder formalmente. Os passos são os seguintes:

Seja $A \subset \mathbb{R}$ infinito e limitado.

- (1) $A \subset I_1 := [a, b]$ para algum $a < b \in \mathbb{R}$, pois A é limitado.
- (2) Seja I_2 um dos conjuntos $[a, (a+b)/2]$ ou $[(a+b)/2, b]$, tal que I_2 contenha infinitos pontos de A . Note que $I_2 \subset I_1$.
- (3) Divida I_2 em dois subconjuntos fechados de mesmo comprimento e defina I_3 como sendo uma das partes tal que contenha infinitos pontos de A . Por definição, $I_3 \subset I_2$.
- (4) Prossiga assim definindo I_4, \dots, I_n tais que $I_n \subset \dots \subset I_2 \subset I_1$, e que I_n seja fechado e contenha infinitos pontos de A .
- (5) Usando Teorema dos intervalos encaixantes, seja $x \in \bigcap_{n=1}^{\infty} I_n$.
- (6) Mostre que x é ponto de acumulação.

DEMONSTRAÇÃO. (do Teorema 2.4.2, versão unidimensional). Seja $A \subset \mathbb{R}$ infinito e limitado. Como A é limitado, existe $I_1 = [a, b] \subset \mathbb{R}$ tal que $A \subset I_1$. Note que $[a, (a+b)/2]$ ou $[(a+b)/2, b]$ ou contém infinitos pontos de A , e chame de I_2 tal intervalo. Da mesma forma, decomponha I_2 em dois subintervalos fechados, e denomine por I_3 um dos subintervalos tal que $I_3 \cap A$ contenha infinitos pontos. Assim procedendo, obtemos uma sequência encaixante $I_n \subset \dots \subset I_2 \subset I_1$. Pelo Teorema dos intervalos encaixantes, existe $\xi \in \bigcap_{n=1}^{\infty} I_n$.

Temos agora que mostrar que ξ é ponto de acumulação. Note que o comprimento de $I_n = b_n - a_n = (b - a)/2^{n-1}$. Dado $\epsilon > 0$, seja $V = (\xi - \epsilon, \xi + \epsilon)$. Seja n tal que $(b - a)/2^{n-1} < \epsilon$. Então $I_n \subset V$, pois se $x \in I_n$, então

$$|x - \xi| < b_n - a_n < \epsilon \implies \xi \in V.$$

Logo V contém infinitos pontos de A , e ξ é ponto de acumulação. \square

2.5. Conjuntos Compactos

Um importante conceito em análise é o de *conjuntos compactos*. Em espaços de dimensão finita, estes conjuntos são na verdade conjuntos fechados limitados, e a noção de compacidade ajuda apenas nas demonstrações, tornando-as mais diretas. Entretanto, em dimensão infinita, nem todo fechado limitado é compacto, e algumas propriedades que continuam valendo para compactos, deixam de valer para fechados limitados.

Antes de definirmos compactos, precisamos introduzir a noção de cobertura aberta.

DEFINIÇÃO 2.5.1. *Seja $A \subset \mathbb{R}^n$. Chamamos $\mathcal{G} = \{G_\alpha\}$ de cobertura aberta de A se para todo α temos G_α conjunto aberto, e $A \subset \cup_\alpha G_\alpha$.*

EXEMPLO 2.37. Como $(0, 1) \subset \cup_{i=1}^\infty (1/i, 1)$, então $\mathcal{G} = \{(1/i, 1)\}_{i=1}^\infty$ é uma cobertura aberta de $(0, 1)$.

EXEMPLO 2.38. Se para $x \in \mathbb{R}$, temos $G_x = (x - 1, x + 1)$, então $\mathcal{G} = \{G_x\}_{x \in \mathbb{R}^n}$ é uma cobertura aberta de \mathbb{R} .

DEFINIÇÃO 2.5.2. *Dizemos que um conjunto $K \subset \mathbb{R}^n$ é compacto se para toda cobertura aberta de K existir uma subcobertura finita de K em \mathcal{G} . Em outras palavras, se existe cobertura aberta $\mathcal{G} = \{G_\alpha\}$ de K tal que $K \subset \cup_\alpha G_\alpha$, então existem $\alpha_1, \alpha_2, \dots, \alpha_J$ tais que $K \subset \cup_{j=1}^J G_{\alpha_j}$.*

Note que para mostrar que um determinado conjunto é compacto precisamos provas que para *toda* cobertura aberta existe subcobertura finita. Para mostrar que não é compacto basta achar uma cobertura que não possui subcobertura finita.

EXEMPLO 2.39. Seja $K = \{x_1, x_2, \dots, x_J\}$ conjunto finito em \mathbb{R} e seja $\mathcal{G} = \{G_\alpha\}$ coleção de conjuntos abertos em \mathbb{R} tais que $K \subset \cup_\alpha G_\alpha$, i.e., \mathcal{G} é uma cobertura aberta de K . Para $j = 1, \dots, J$, seja $G_j \in \mathcal{G}$ tal que $x_j \in G_j$ (tal conjunto sempre existe pois \mathcal{G} é cobertura de K). Então G_1, \dots, G_J geram uma subcobertura finita de K . Logo K é compacto, e concluímos que todo conjunto finito é compacto.

EXEMPLO 2.40. O conjunto $(0, 1)$ não é compacto. De fato $(0, 1) \subset \cup_{j=1}^\infty G_j$, onde $G_j = (1/j, 1)$. Mas se existisse $\{G_{n_1}, \dots, G_{n_p}\}$ tal que $(0, 1) \subset \cup_{i=1}^p G_j$, então $(0, 1) \subset (1/N^*, 1)$, onde $N^* = \max\{n_1, \dots, n_p\} > 0$, um absurdo.

TEOREMA 2.5.3 (Heine–Borel). *Um conjunto em \mathbb{R}^n é compacto se e somente se é fechado e limitado.*

DEMONSTRAÇÃO. (\Rightarrow) Suponha $K \subset \mathbb{R}^n$ conjunto compacto. Então $K \subset \cup_{m=1}^\infty B_m(\mathbf{0})$. Como K é compacto, a cobertura acima possui subcobertura finita e portanto existe M tal que $K \subset B_M(\mathbf{0})$. Logo K é limitado.

Para mostrar que é também fechado, seja $\mathbf{x} \in \mathcal{C}(K)$ e $G_i = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - \mathbf{x}\| > 1/i\}$ para $i \in \mathbb{N}$. Logo G_i é aberto e $\mathbb{R}^n \setminus \{\mathbf{x}\} = \cup_{i=1}^\infty G_i$. Mas como $\mathbf{x} \notin K$, então $K \subset \cup_{i=1}^\infty G_i$. Usando agora que K é compacto, extrairmos uma subcobertura finita e temos $K \subset \cup_{i=1}^{N^*} G_i = G_{N^*}$ para algum $N^* \in \mathbb{N}$. Portanto $K \cap B_{1/N^*}(\mathbf{x}) = \emptyset$ e concluímos que $B_{1/N^*}(\mathbf{x}) \subset \mathcal{C}(K)$. Logo $\mathcal{C}(K)$ é aberto e K é fechado.

(\Leftarrow)(Contradição) Suponha K fechado e limitado. Então existe uma cela

$$C = \{\mathbf{x} \in \mathbb{R}^n : a_i \leq x_i \leq b_i, \text{ para } i = 1, \dots, n\}$$

tal que $K \subset C$. Seja $d = [\sum_{i=1}^n (b_i - a_i)^2]^{1/2}$. A fim de obter uma contradição, suponha que exista um recobrimento aberto $\{G_\alpha\}$ de K que não contenha nenhuma subcobertura finita de K . Seja $c_i = (a_i + b_i)/2$. Então $[a_i, c_i]$ e $[c_i, b_i]$ determinam 2^n celas cuja união é C . Pelo menos uma destas celas contém pontos da parte de K que não tem subcobertura finita. Chame de C_1 esta cela. Subdividindo C_1 desta mesma forma, obtemos uma sequência de celas fechadas $\{C_j\}$ tal que

- (1) $C_1 \supset C_2 \supset C_3 \supset \dots$,
- (2) C_j contém parte de K que não tem subcobertura finita,
- (3) se $\mathbf{x}, \mathbf{y} \in C_j$, então $\|\mathbf{x} - \mathbf{y}\| \leq 2^{-j}d$.

Pelo Teorema das celas encaixantes (Teorema 2.4.1), existe $\boldsymbol{\xi} \in C_j$, para todo $j \in \mathbb{N}$. Como $C_j \cap K$ é infinito, então $\boldsymbol{\xi}$ é ponto de acumulação de K . Mas K fechado implica que $\boldsymbol{\xi} \in K$. Portanto $\boldsymbol{\xi} \in G_\alpha$, para algum α . Como G_α é aberto, então existe r tal que

$$(2.5.1) \quad \|\mathbf{y} - \boldsymbol{\xi}\| \leq r \implies \mathbf{y} \in G_\alpha.$$

Seja $J \in \mathbb{N}$ tal que $2^{-J}d < r$, e \mathbf{y} um ponto arbitrário de C_J . Por (3) acima,

$$\|\boldsymbol{\xi} - \mathbf{y}\| \leq 2^{-n}d < r.$$

Por (2.5.1), concluímos que $\mathbf{y} \in G_\alpha$, e portanto, todo ponto de C_j pertence a G_α . Logo, $C_j \subset G_\alpha$, e G_α é uma cobertura de C_j , uma contradição com (2). \square

Uma outra demonstração que apresentamos abaixo como curiosidade vale no caso unidimensional e pode ser usada para mostrar que um conjunto fechado e limitado em \mathbb{R} é compacto.

TEOREMA 2.5.4. *Um conjunto fechado e limitado em \mathbb{R} é compacto.*

DEMONSTRAÇÃO. Parte (i) Primeiro provamos o resultado para um conjunto do tipo $[-a, a]$, onde $a > 0$, e $\mathcal{G} = \{G_\alpha\}$ cobertura aberta de $[-a, a]$. A seguir mostraremos o caso geral. Seja

$$C_a = \{x \in [-a, a] : [-a, x] \text{ pode ser coberto por finitos abertos de } \mathcal{G}\}.$$

Então C_a é não vazio, pois $-a \in C_a$, e é limitado. Seja $s^* = \sup C_a$. Então $s^* \in [-a, a]$, pois se $s^* > a$ teríamos a como cota superior de C_a menor que o supremo, um absurdo.

Seja então $G_{\bar{\alpha}}$ elemento de \mathcal{G} tal que $s^* \in G_{\bar{\alpha}}$. Sabemos que tal $G_{\bar{\alpha}}$ existe pois \mathcal{G} é cobertura de $[-a, a]$ e $s^* \in [-a, a]$.

Primeiro afirmamos que $s^* \in C_a$, pois caso contrário suponha $\{G_{\alpha_1}, \dots, G_{\alpha_n}\}$ subcobertura finita de C_a . Então teríamos $\{G_{\alpha_1}, \dots, G_{\alpha_n}, G_{\bar{\alpha}}\}$ subcobertura finita de $[-a, s^*]$.

Queremos mostrar agora que $s^* = a$. Supondo $s^* < a$, e como $G_{\bar{\alpha}}$ é aberto então existe ϵ tal que $s^* + \epsilon \in G_{\bar{\alpha}}$, e $s^* + \epsilon < a$, logo $s^* + \epsilon \in C_a$, uma contradição com a definição de supremo.

Parte (ii) Consideramos agora o caso geral, onde $K \subseteq \mathbb{R}^n$ é fechado e limitado, e $\mathcal{G} = \{G_\alpha\}$ é cobertura aberta de K . Como K é fechado, então $\mathcal{C}(K)$ é aberto, e como K é limitado, então existe $a \in \mathbb{R}^n$ tal que $K \subset [-a, a]$. Logo $\{G_\alpha, \mathcal{C}(K)\}$ geram uma cobertura aberta de $[-a, a]$. Pela **Parte (i)**, existe uma subcobertura $\{G_{\alpha_1}, \dots, G_{\alpha_n}, \mathcal{C}(K)\}$ de $[-a, a]$, e portanto também de K pois $K \subset [-a, a]$. Como $K \cap \mathcal{C}(K) = \emptyset$, então $\{G_{\alpha_1}, \dots, G_{\alpha_n}\}$ é uma cobertura finita de K . \square

2.6. Exercícios

EXERCÍCIO 2.1. Prove a afirmativa do exemplo 2.4.

EXERCÍCIO 2.2. Se $A \subset \mathbb{R}$ é um conjunto não vazio e limitado, então $A \subset [\inf S, \sup S]$.

EXERCÍCIO 2.3. Seja $A \subset \mathbb{R}$ e as funções $f : A \rightarrow \mathbb{R}$ e $g : A \rightarrow \mathbb{R}$ sejam tais que os conjuntos $f(A)$ e $g(A)$ sejam limitados superiormente. Defina a função $f + g : A \rightarrow \mathbb{R}$ por $(f + g)(x) = f(x) + g(x)$. Mostre que $\sup(f + g)(A) \leq \sup f(A) + \sup g(A)$. Dê um exemplo em que a desigualdade é estrita.

EXERCÍCIO 2.4. Seja $A \subset \mathbb{R}$ conjunto limitado. Mostre que $\inf A$ e $\sup A$ são únicos.

EXERCÍCIO 2.5. Enuncie e demonstre o resultado análogo ao Lema 2.1.3 no caso do ínfimo.

EXERCÍCIO 2.6 (Densidade dos racionais nos reais). Mostre que dados $x, y \in \mathbb{R}$ com $x < y$, existe $r \in \mathbb{Q}$ tal que $x < r < y$.

EXERCÍCIO 2.7. Demonstre os itens (2) e (3) no Lema 2.1.5.

EXERCÍCIO 2.8. Faça os detalhes do exemplo 2.9.

EXERCÍCIO 2.9. Aponte na demonstração do Teorema 2.1.8 quais o(s) argumento(s) que não é (são) válido(s) se considerarmos uma sequência encaixante de intervalos abertos.

EXERCÍCIO 2.10. Demonstrar os itens (2), (3) e (4) do Lema 2.2.2.

EXERCÍCIO 2.11. Seja V espaço vetorial com norma $\|\cdot\|$ induzida por produto interno. Mostre que vale a *lei do paralelograma*, i.e., para todo $\mathbf{x}, \mathbf{y} \in V$ tem-se

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2).$$

EXERCÍCIO 2.12. 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 2.13. Mostre que (2.2.3) define uma norma.

EXERCÍCIO 2.14. O conjunto $\{\mathbf{x} = (x_1, \dots, x_n) \in B_1(\mathbf{0}) : x_1 \in \mathbb{Q}\}$ é aberto? Prove a sua afirmação.

EXERCÍCIO 2.15. Mostre que $B_1(\mathbf{0})$ é aberto.

EXERCÍCIO 2.16. Um conjunto pode ser aberto em \mathbb{R}^n como na Definição 2.3.1, mas também pode ser *aberto em outro conjunto* (aberto relativo). Dizemos que U é aberto em V se para todo $\mathbf{x} \in U$ existe $\epsilon > 0$ tal que $B_\epsilon(\mathbf{x}) \cap V \subseteq U$. Mostre que todo conjunto do \mathbb{R}^n é aberto nele mesmo.

EXERCÍCIO 2.17. Assim como no exercício 2.16, um conjunto F é fechado em V se $V \setminus F$ (o complementar de F em relação ao V) é aberto em V . Mostre que todo conjunto do \mathbb{R}^n é fechado nele mesmo.

EXERCÍCIO 2.18. Seja $A \subset \mathbb{R}^n$, e denote por interior de A o conjunto A° de pontos interiores de A . Mostre que

$$(1) (A^\circ)^\circ = A^\circ$$

$$(2) (A \cap B)^\circ = A^\circ \cap B^\circ$$

(3) Se $B \subset A$ e B é aberto, então $B \subset A^\circ$ (i.e. A° é o “maior” aberto contido em A)

EXERCÍCIO 2.19. Seja $A \subset \mathbb{R}^n$. Chamamos de *fecho de A* , e denotamos por \bar{A} , a interseção de todos os fechados que contenham A . Mostre que $\mathbf{x} \in \bar{A}$ se e somente se \mathbf{x} é ponto de interior ou de fronteira da A .

EXERCÍCIO 2.20. Demonstre o Corolário 2.3.7.

EXERCÍCIO 2.21. Mostre que um ponto $\mathbf{x} \in \mathbb{R}^n$ é ponto de acumulação de A se e somente se toda vizinhança aberta de \mathbf{x} contém infinitos pontos de A .

EXERCÍCIO 2.22. Mostre que todo ponto de

$$\{1/n : n \in \mathbb{N}\}$$

é ponto de fronteira, e que 0 é o único ponto de acumulação.

EXERCÍCIO 2.23. Sejam $A, B \subset \mathbb{R}^n$, e \mathbf{x} ponto de acumulação de $A \cap B$. Mostre que \mathbf{x} é ponto de acumulação de A e de B .

EXERCÍCIO 2.24. Mostre que se $F \neq \emptyset$ é fechado em \mathbb{R}^n , e $\inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in F\} = 0$, então $\mathbf{x} \in F$.

EXERCÍCIO 2.25. Mostre que se $\mathbf{x} \neq \mathbf{y}$ são pontos em \mathbb{R}^n , então existem vizinhanças abertas U de \mathbf{x} e V de \mathbf{y} tais que $U \cap V = \emptyset$

EXERCÍCIO 2.26. Mostre que se U e V são vizinhanças abertas de $\mathbf{x} \in \mathbb{R}^n$, então $U \cap V$ é vizinhança aberta de \mathbf{x} .

EXERCÍCIO 2.27. Generalize para o \mathbb{R}^n as idéias apresentadas na demonstração unidimensional do Teorema de Bolzano–Weierstrass (Teorema 2.4.2).

EXERCÍCIO 2.28. Para cada um dos conjuntos abaixo, ache, se for possível, uma cobertura de abertos que não contenha subcobertura finita.

(1) \mathbb{R}

(2) $\{1, 1/2, 1/3, 1/4, \dots\}$

(3) $\{0, 1, 1/2, 1/3, 1/4, \dots\}$

EXERCÍCIO 2.29. Mostre sem usar o Teorema de Heine–Borel que a bola aberta $B_1(\mathbf{0}) \subset \mathbb{R}^n$ não é compacta.

EXERCÍCIO 2.30. Mostre, sem usar o Teorema de Heine–Borel, que se K é compacto e $F \subset K$ é fechado, então F é compacto.

EXERCÍCIO 2.31. Sem usar o Teorema de Bolzano–Weierstrass no \mathbb{R}^n (Teorema 2.4.2), mostre que se K é compacto e $A \subset K$ é infinito, então existe pelo menos um ponto de acumulação de A .

EXERCÍCIO 2.32 (Teorema da interseção de Cantor). Suponha que $\{K_j\}$ seja uma coleção de conjuntos não vazios, compactos, com $K_1 \supset K_2 \supset K_3 \supset \dots$. Mostre que $\bigcap_{j=1}^{\infty} K_j$ é não vazio.

CAPÍTULO 3

Sequências

1

3.1. Definição e resultados preliminares

Uma sequência em \mathbb{R}^n é simplesmente uma função de \mathbb{N} em \mathbb{R}^n . Portanto $X : \mathbb{N} \rightarrow \mathbb{R}^n$ indica uma sequência de números reais, que escrevemos também como (\mathbf{x}_k) , ou ainda $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots)$. Para indicar o k -ésimo valor da sequência escrevemos simplesmente \mathbf{x}_k .

EXEMPLO 3.1. $x_k = (-1)^k$ define a sequência $(-1, 1, -1, 1, -1, 1, -1, \dots)$ em \mathbb{R} .

EXEMPLO 3.2. A sequência de Fibonacci é definida recursivamente por $x_1 = 1, x_2 = 1$, e $x_{k+1} = x_k + x_{k-1}$ para $k \geq 2$. Portanto temos $(x_k) = (1, 1, 2, 3, 5, 8, \dots)$.

Podemos realizar com sequências várias das operações que realizamos com números reais, como por exemplo somar, subtrair, etc. Sejam por exemplo (\mathbf{x}_k) e (\mathbf{y}_k) duas sequências em \mathbb{R}^n , e $c \in \mathbb{R}$. Então definimos

$$(\mathbf{x}_k) + (\mathbf{y}_k) = (\mathbf{x}_k + \mathbf{y}_k), \quad (\mathbf{x}_k) - (\mathbf{y}_k) = (\mathbf{x}_k - \mathbf{y}_k), \quad c(\mathbf{x}_k) = (c\mathbf{x}_k).$$

Podemos da mesma forma definir produtos internos de sequências em \mathbb{R}^n por $(\mathbf{x}_k) \cdot (\mathbf{y}_k) = (\mathbf{x}_k \cdot \mathbf{y}_k)$.

EXEMPLO 3.3. Se $x_k = (2, 4, 6, 8, \dots)$ e $(y_k) = (1, 1/2, 1/3, 1/4, \dots)$, então $(x_k) \cdot (y_k) = (2, 2, 2, \dots)$.

A primeira pergunta que surge quando tratamos de sequências é quanto à convergência destas, isto é, se quando k aumenta, os termos x_k 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 k . Em outras palavras queremos determinar o comportamento das sequências no “limite”.

DEFINIÇÃO 3.1.1. Dizemos que $\mathbf{x} \in \mathbb{R}^n$ é limite de uma sequência (\mathbf{x}_k) , se para toda vizinhança aberta U de \mathbf{x} existir $K^* \in \mathbb{N}$ tal que $\mathbf{x}_k \in U$ para todo $k \geq K^*$. Escrevemos neste caso que $\mathbf{x}_k \rightarrow \mathbf{x}$, ou que $\mathbf{x} = \lim \mathbf{x}_k$, ou ainda

$$\mathbf{x} = \lim_{k \rightarrow \infty} \mathbf{x}_k.$$

De forma equivalente, $\mathbf{x}_k \rightarrow \mathbf{x}$ se para todo $\epsilon > 0$, existe $K^* \in \mathbb{N}$ tal que $\mathbf{x}_k \in B_\epsilon(\mathbf{x})$ para todo $k \geq K^*$.

Se uma sequência tem limite, dizemos que ela converge ou que é convergente, e se não tem limite dizemos que ela diverge ou que é divergente.

¹Última Atualização: 25/02/2008

O lema abaixo é consequência da definição de convergência, e portanto na maioria dos exemplos a seguir nos restringimos ao caso unidimensional.

LEMA 3.1.2. Toda sequência (\mathbf{x}_k) em \mathbb{R}^n converge se e somente se a sequência das i -ésimas coordenadas $((x_i)_k)$ converge em \mathbb{R} para $i = 1, \dots, n$.

DEMONSTRAÇÃO. Exercício. □

EXEMPLO 3.4. Se $x_k = 1$, então $\lim x_k = 1$. De fato, dado $\epsilon > 0$, para todo $k \geq 1$ temos $|x_k - 1| = 0 < \epsilon$.

EXEMPLO 3.5. $\lim(1/k) = 0$. De fato, dado $\epsilon > 0$, seja K^* tal que $1/K^* < \epsilon$. Logo, para todo $k \geq K^*$ temos $|1/k - 0| = 1/k \leq 1/K^* < \epsilon$.

Observe que diferentes situações ocorrem nos exemplos acima. Em 3.4 a sequência é constante, e a escolha de K^* independe de ϵ . Já no exemplo 3.5, temos que K^* claramente depende de ϵ .

A seguir, no exemplo 3.6 o objetivo é mostrar que um certo valor x não é o limite da sequência (x_k) . Mostramos então que existe pelo menos um certo $\epsilon > 0$ tal que para todo K^* , conseguimos achar $k \geq K^*$ tal que $|x_k - x| > \epsilon$. Note que o que fazemos é *negar* a convergência.

EXEMPLO 3.6. $(0, 2, 0, 2, 0, 2, 0, 2, \dots)$ não converge para 0. De fato, tome $\epsilon = 1$. Então para todo $K^* \in \mathbb{N}$ temos $2K^* > K^*$ e $x_{2K^*} = 2$. Portanto $|x_{2K^*} - 0| = 2 > \epsilon$.

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 3.1.3 (Unicidade de limite). *Uma sequência pode ter no máximo um limite.*

DEMONSTRAÇÃO. Considere que (\mathbf{x}_k) é uma sequência tal que $\mathbf{x}_k \rightarrow \mathbf{x}$ e $\mathbf{x}_k \rightarrow \mathbf{x}'$, com $\mathbf{x} \neq \mathbf{x}'$. Sejam $\epsilon = \|\mathbf{x} - \mathbf{x}'\|/2 > 0$, e sejam K^* e $K' \in \mathbb{N}$ tais que $\|\mathbf{x}_k - \mathbf{x}\| < \epsilon$ para todo $k \geq K^*$ e $\|\mathbf{x}_k - \mathbf{x}'\| < \epsilon$ para todo $k \geq K'$. Logo, se $k \geq \max\{K^*, K'\}$, então

$$\|\mathbf{x} - \mathbf{x}'\| \leq \|\mathbf{x} - \mathbf{x}_k\| + \|\mathbf{x}_k - \mathbf{x}'\| < 2\epsilon = \|\mathbf{x} - \mathbf{x}'\|.$$

Como um número não pode ser estritamente menor que ele mesmo, temos uma contradição. Portanto $\mathbf{x} = \mathbf{x}'$ e o limite é único. □

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 3.1.4. Seja (\mathbf{x}_k) e (\mathbf{y}_k) tais que $\lim \mathbf{x}_k = \mathbf{x}$ e $\lim \mathbf{y}_k = \mathbf{y}$. Então

- (1) $\lim(\mathbf{x}_k + \mathbf{y}_k) = \mathbf{x} + \mathbf{y}$.
- (2) $\lim(\mathbf{x}_k - \mathbf{y}_k) = \mathbf{x} - \mathbf{y}$.
- (3) $\lim(c\mathbf{x}_k) = c\mathbf{x}$, para $c \in \mathbb{R}$.
- (4) Em \mathbb{R} , temos que $\lim(x_k y_k) = xy$.
- (5) Em \mathbb{R} , temos que se $y_k \neq 0$ para todo k e $y \neq 0$, então $\lim(x_k/y_k) = x/y$.

DEMONSTRAÇÃO. (1) Dado $\epsilon > 0$, seja $K^* \in \mathbb{N}$ tal que $\|\mathbf{x}_k - \mathbf{x}\| < \epsilon/2$ e $\|\mathbf{y}_k - \mathbf{y}\| < \epsilon/2$ para todo $k \geq K^*$. Logo

$$\|\mathbf{x}_k + \mathbf{y}_k - (\mathbf{x} + \mathbf{y})\| \leq \|\mathbf{x}_k - \mathbf{x}\| + \|\mathbf{y}_k - \mathbf{y}\| < \epsilon \quad \text{para todo } k \geq K^*.$$

(2) A demonstração é basicamente a mesma de (1), tomando-se o devido cuidado com os sinais.

(4) Para todo $k \in \mathbb{N}$ temos

$$|x_k y_k - xy| \leq |x_k y_k - x_k y| + |x_k y - xy| = |x_k| |y_k - y| + |y| |x_k - x|.$$

Seja $M \in \mathbb{R}$ tal que $|x_k| < M$ e $|y| < M$. Tal constante M existe pois como (x_k) converge, ela é limitada. Agora, dado $\epsilon > 0$, seja K^* tal que $|y_k - y| < \epsilon/(2M)$ e $|x_k - x| < \epsilon/(2M)$ para todo $k \geq K^*$. Logo,

$$|x_k y_k - xy| \leq M[|y_k - y| + |x_k - x|] < \epsilon,$$

para todo $k \geq K^*$.

Deixamos (3) e (5) como exercícios para o leitor. \square

OBSERVAÇÃO. Os resultados do lema acima continuam válidos para um número finito de somas, produtos, etc.

As vezes, uma sequência se aproxima de algum valor em \mathbb{R}^n de forma mais lenta que alguma outra sequência de reais que converge para 0. É possível assim garantir convergência, como o resultado a seguir nos mostra.

LEMA 3.1.5. Seja (a_k) sequência em \mathbb{R} convergente para 0. Se para (\mathbf{x}_k) sequência em \mathbb{R}^n existir $c > 0$ tal que

$$\|\mathbf{x}_k - \mathbf{x}\| \leq c|a_k| \quad \text{para todo } k \in \mathbb{N},$$

então $\mathbf{x}_k \rightarrow \mathbf{x}$.

DEMONSTRAÇÃO. Como (a_k) converge, dado $\epsilon > 0$, seja $K^* \in \mathbb{N}$ tal que $|a_k| < \epsilon/c$ para todo $k \geq K^*$. Logo

$$\|\mathbf{x}_k - \mathbf{x}\| \leq c|a_k| < \epsilon \quad \text{para todo } k \geq K^*,$$

e $\lim \mathbf{x}_k = \mathbf{x}$. \square

COROLÁRIO 3.1.6. Seja (a_k) sequência em \mathbb{R} convergente para 0. Se para (\mathbf{x}_k) sequência em \mathbb{R}^n existir $c > 0$ e $K^* \in \mathbb{N}$ tal que

$$\|\mathbf{x}_k - \mathbf{x}\| \leq c|a_k| \quad \text{para todo } k \geq K^*,$$

então $\mathbf{x}_k \rightarrow \mathbf{x}$.

EXEMPLO 3.7. Seja $x_k = (2/k) \sin(1/k)$. Então

$$|x_k - 0| \leq \frac{2}{k}.$$

Como $1/k \rightarrow 0$, podemos usar o lema acima para garantir que $\lim[(2/k) \sin(1/k)] = 0$.

EXEMPLO 3.8. $\lim_{k \rightarrow \infty} ((2k + 1)/k) = 2$. De fato,

$$\frac{2k + 1}{k} = (2) + \left(\frac{1}{k}\right).$$

Como $\lim_{k \rightarrow \infty} 2 = 2$ e $\lim_{k \rightarrow \infty} (1/k) = 0$, nós obtemos o resultado.

EXEMPLO 3.9. $\lim_{k \rightarrow \infty} (2k/(k^2 + 1)) = 0$, pois

$$\frac{2k}{k^2 + 1} = \frac{2/k}{1 + 1/k^2}.$$

Como $\lim_{k \rightarrow \infty} (2/k) = 0$ e $\lim_{k \rightarrow \infty} (1 + 1/k^2) = 1 \neq 0$, podemos aplicar o resultado sobre quociente de sequências.

EXEMPLO 3.10. A sequência

$$x_k = \frac{1}{k^2} \sum_{i=1}^k i$$

converge. Primeiro note que

$$(3.1.1) \quad \sum_{i=1}^k i = \frac{k^2 + k}{2}.$$

Para $k = 1$ o resultado (3.1.1) é trivial. Suponha (3.1.1) verdadeiro para $k = k^*$. 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 (3.1.1) é verdadeira. Temos então que

$$x_k = \frac{k^2 + k}{2k^2} = \frac{1}{2} \left(1 + \frac{1}{k}\right) = \frac{1}{2} + \left(\frac{1}{2k}\right).$$

Logo (x_k) é soma de duas sequências convergentes, $(1/2)$ e $(1/2)(1/k)$ e

$$\lim_{k \rightarrow \infty} x_k = \lim_{k \rightarrow \infty} \frac{1}{2} + \lim_{k \rightarrow \infty} \frac{1}{2k} = \frac{1}{2}.$$

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 3.1.7. Dizemos que uma sequência (\mathbf{x}_k) é limitada quando existe um número real C tal que $\|\mathbf{x}_k\| \leq C$ para todo $k \in \mathbb{N}$.

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 norma.

TEOREMA 3.1.8. Toda sequência convergente é limitada

DEMONSTRAÇÃO. Seja (\mathbf{x}_k) sequência convergente e seja \mathbf{x} seu limite. Seja $\epsilon = 1$. Como (\mathbf{x}_k) converge, existe K^* tal que $\|\mathbf{x} - \mathbf{x}_k\| < 1$ para todo $k \geq K^*$. Logo, usando a desigualdade triangular temos

$$\|\mathbf{x}_k\| \leq \|\mathbf{x}_k - \mathbf{x}\| + \|\mathbf{x}\| < 1 + \|\mathbf{x}\| \quad \text{para todo } k \geq K^*.$$

Falta agora limitar os K^* primeiros termos da sequência. Seja então

$$C = \max\{\|\mathbf{x}_1\|, \|\mathbf{x}_2\|, \|\mathbf{x}_3\|, \dots, \|\mathbf{x}_{K^*}\|, 1 + \|\mathbf{x}\|\}.$$

Portanto $\|\mathbf{x}_k\| \leq C$ para todo $k \in \mathbb{N}$. □

EXEMPLO 3.11. A sequência (k) em \mathbb{R} diverge pois não é limitada.

EXEMPLO 3.12. Seja $S_k = 1 + 1/2 + 1/3 + 1/4 + \dots + 1/k$. Mostraremos que (S_k) não é limitada, e portanto divergente. Note que

$$\begin{aligned} x_{2^k} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^{k-1}+1} + \dots + \frac{1}{2^k}\right) \\ &= 1 + \frac{1}{2} + \sum_{i=3}^4 \frac{1}{i} + \sum_{i=5}^8 \frac{1}{i} + \dots + \sum_{i=2^{k-1}+1}^{2^k} \frac{1}{i} > 1 + \frac{1}{2} + \sum_{i=3}^4 \frac{1}{i} + \sum_{i=5}^8 \frac{1}{i} + \dots + \sum_{i=2^{k-1}+1}^{2^k} \frac{1}{i} \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots + \frac{1}{2} = 1 + \frac{k}{2}. \end{aligned}$$

Logo (S_k) não é limitada, e portanto diverge.

Outra forma de ver que a sequência acima diverge é por indução. Quero mostrar que $S_{2^k} \geq 1 + k/2$. Note que $S_2 = 1 + 1/2$. Supondo que $S_{2^{k-1}} \geq 1 + (k-1)/2$ temos

$$S_{2^k} = S_{2^{k-1}} + \frac{1}{2^{k-1}+1} + \dots + \frac{1}{2^k} > 1 + \frac{(k-1)}{2} + \frac{1}{2} > 1 + \frac{k}{2},$$

como queríamos demonstrar. Mais uma vez a conclusão é que (S_k) não é limitada, logo diverge.

EXEMPLO 3.13 (Sequência de Cesàro). Seja (\mathbf{x}_j) sequência convergente em \mathbb{R}^n , e seja $\mathbf{x} \in \mathbb{R}^n$ seu limite. Então a sequência definida por

$$\frac{1}{j}(\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_j)$$

converge e tem \mathbf{x} como seu limite.

Sem perda de generalidade, supomos que (\mathbf{x}_j) converge para zero. Para o caso geral quando (\mathbf{x}_j) converge para \mathbf{x} basta tratar a sequência $(\mathbf{x}_j - \mathbf{x})$.

Seja $\mathbf{S}_j = (\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_j)/j$. Como (\mathbf{x}_j) converge, então é limitada. Seja M tal que $|\mathbf{x}_j| < M$ para todo $j \in \mathbb{N}$. Dado $\epsilon > 0$, seja K^* tal que $M/K^* < \epsilon$ e $|\mathbf{x}_j| < \epsilon$ para todo $j \geq K^*$. Então, temos $\mathbf{S}_j = \check{\mathbf{S}}_j + \hat{\mathbf{S}}_j$, onde

$$\check{\mathbf{S}}_j = \frac{1}{j}(\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_{K^*}), \quad \hat{\mathbf{S}}_j = \frac{1}{j}(\mathbf{x}_{K^*+1} + \mathbf{x}_{K^*+1} + \dots + \mathbf{x}_j).$$

Então (\mathbf{S}_j) é a soma de duas sequências convergentes para zero. De fato para $j \geq (K^*)^2$, temos

$$|\check{\mathbf{S}}_j| \leq \frac{1}{j}(|\mathbf{x}_1| + |\mathbf{x}_2| + \dots + |\mathbf{x}_{K^*}|) \leq \frac{K^*M}{j} \leq \frac{M}{K^*} < \epsilon.$$

Além disso, $|\hat{\mathbf{S}}_j| < \epsilon(j - K^*)/j < \epsilon$. Portanto (\mathbf{S}_j) converge para zero.

Note que sequências convergentes convergem também no sentido de Cesàro. Entretanto o oposto não ocorre. Considere como exemplo $(x_k) = (0, 1, 0, 1, 0, 1, \dots)$. Então S_j como definida acima converge para $1/2$, apesar de (x_k) não convergir.

Outro resultado importante refere-se à convergência das normas de sequências: se uma sequência converge, então a sequência de normas também converge. A recíproca *não* é verdadeira. Basta considerar como contra-exemplo a sequência $((-1)^n)$. Neste caso a sequência diverge mas a sequência de seus valores absolutos converge.

LEMA 3.1.9. Seja (\mathbf{x}_j) convergente. Então $(\|\mathbf{x}_j\|)$ também o é.

DEMONSTRAÇÃO. Exercício. □

3.2. Subsequências e Teorema de Bolzano–Weierstrass

Seja (\mathbf{x}_k) sequência em \mathbb{R}^n e

$$k_1 < k_2 < k_3 < \dots < k_j < \dots$$

sequência de números naturais. Então dizemos que (\mathbf{x}_{k_j}) é uma *subsequência* de (\mathbf{x}_k) .

OBSERVAÇÃO. Para definir subsequências de forma rigorosa, basta supor que $k : \mathbb{N} \rightarrow \mathbb{N}$ é uma função estritamente crescente, i.e., $i > j$ implica em $k(i) > k(j)$. Para facilitar a notação, escrevemos $k(i)$ simplesmente como k_i . Note que sempre $k_i \geq i$.

EXEMPLO 3.14. Se $(x_k) = (1, 1/2, 1/3, 1/4, \dots)$, então $(1, 1/2, 1/4, 1/6, 1/8, \dots)$ e (x_{2k}) são subsequências de (x_k) .

Um primeiro resultado relacionado com subsequências nos diz que se uma sequência converge para um determinado limite, então todas as subsequências convergem e têm o mesmo limite.

LEMA 3.2.1. Se uma sequência (\mathbf{x}_k) converge para \mathbf{x} , então todas as subsequências de (\mathbf{x}_k) são convergentes e têm o mesmo limite \mathbf{x} .

DEMONSTRAÇÃO. Seja (\mathbf{x}_k) sequência convergente, e seja $\mathbf{x} = \lim_{k \rightarrow \infty} \mathbf{x}_k$. Dado $\epsilon > 0$, seja K^* tal que

$$(3.2.1) \quad \|\mathbf{x} - \mathbf{x}_k\| < \epsilon \quad \text{para todo } k \geq K^*.$$

Seja (\mathbf{x}_{k_j}) subsequência de (\mathbf{x}_k) . Como $k_j \geq j$ para todo $j \in \mathbb{N}$, então $j \geq K^*$ implica em $k_j \geq K^*$ e portanto

$$\|\mathbf{x} - \mathbf{x}_{k_j}\| < \epsilon,$$

por (3.2.1). Logo (\mathbf{x}_{k_j}) converge para \mathbf{x} . □

EXEMPLO 3.15. $((-1)^n)$ diverge pois se convergisse para algum $x \in \mathbb{R}$, suas subsequências convergiriam este mesmo valor. Mas

$$\lim_{n \rightarrow \infty} ((-1)^{2n}) = 1, \quad \lim_{n \rightarrow \infty} ((-1)^{2n+1}) = -1.$$

EXEMPLO 3.16. Seja (x_k) sequência convergente para l e tal que $x_{2k} = x_k^2$. Então

$$l = \lim_{n \rightarrow \infty} (x_{2k}) = \lim_{k \rightarrow \infty} x_k \lim_{k \rightarrow \infty} x_k = l^2.$$

Logo $l = 0$ ou $l = 1$. Para concluirmos qual dos dois candidatos a limite é o correto, precisaríamos de mais informações sobre a sequência. Por exemplo, se $x_k = a^k$ para $a < 1$, temos que $l = 0$ pois a sequência é limitada superiormente por $a < 1$. Então $l = 1$ não pode ser limite, e $\lim_{k \rightarrow \infty} (a^k) = 0$. Por outro lado, se $a = 1$ então $l = 1$.

LEMA 3.2.2 (Critérios de divergência). Seja (\mathbf{x}_k) sequência em \mathbb{R}^n . As afirmativas abaixo são equivalentes:

- (1) (\mathbf{x}_k) não converge para $\mathbf{x} \in \mathbb{R}^n$.
- (2) Existe $\epsilon > 0$ tal que para todo $K^* \in \mathbb{N}$, existe $k_j \in \mathbb{N}$, com $k_j > K^*$ e $\|\mathbf{x} - \mathbf{x}_{k_j}\| \geq \epsilon$.
- (3) Existe $\epsilon > 0$ e uma subsequência (\mathbf{x}_{k_j}) de (\mathbf{x}_k) tal que $\|\mathbf{x} - \mathbf{x}_{k_j}\| > \epsilon$ para todo $j \in \mathbb{N}$.

DEMONSTRAÇÃO. (1) \implies (2): Se (\mathbf{x}_k) não converge para \mathbf{x} então existe $\epsilon > 0$ tal que é impossível achar $K^* \in \mathbb{N}$ tal que $\|\mathbf{x} - \mathbf{x}_k\| < \epsilon$ para todo $k > K^*$. Logo, para todo K^* , existe $k_j > K^*$ tal que $\|\mathbf{x} - \mathbf{x}_{k_j}\| \geq \epsilon$.

(2) \implies (3): Seja ϵ como em (2). Para todo $j \in \mathbb{N}$, seja $k_j > j$ tal que $\|\mathbf{x} - \mathbf{x}_{k_j}\| \geq \epsilon$. Portanto a subsequência (\mathbf{x}_{k_j}) satisfaz a propriedade em (3).

(3) \implies (1): Se (\mathbf{x}_k) convergisse para \mathbf{x} teríamos (\mathbf{x}_{k_j}) convergindo para \mathbf{x} , o que contraria a hipótese inicial. Logo (\mathbf{x}_k) não converge para \mathbf{x} . \square

No exemplos abaixo temos uma aplicação imediata do Lema 3.2.2.

EXEMPLO 3.17. Seja (\mathbf{x}_k) sequência em \mathbb{R}^n tal que toda subsequência de (\mathbf{x}_k) contém uma subsequência convergente para \mathbf{x} . Então (\mathbf{x}_k) converge para \mathbf{x} .

Por contradição suponha que (\mathbf{x}_k) não convirja para \mathbf{x} . Portanto existe uma subsequência (\mathbf{x}_{k_j}) e $\epsilon > 0$ tal que

$$(3.2.2) \quad \|\mathbf{x} - \mathbf{x}_{k_j}\| > \epsilon \quad \text{para todo } j \in \mathbb{N}.$$

Mas então, por hipótese, (\mathbf{x}_{k_j}) tem uma subsequência convergindo para \mathbf{x} , uma contradição com (3.2.2).

EXEMPLO 3.18. Sejam (\mathbf{x}_n) e (\mathbf{y}_n) sequências em \mathbb{R}^n e seja (\mathbf{z}_i) a sequência formada por $\mathbf{z}_1 = \mathbf{x}_1, \mathbf{z}_2 = \mathbf{y}_1, \mathbf{z}_3 = \mathbf{x}_2, \mathbf{z}_4 = \mathbf{y}_2, \dots, \mathbf{z}_{2i-1} = \mathbf{x}_i, \mathbf{z}_{2i} = \mathbf{y}_i, \dots$. Então, se $\lim_{i \rightarrow \infty} \mathbf{x}_i = \boldsymbol{\xi}$ e $\lim_{i \rightarrow \infty} \mathbf{y}_i = \boldsymbol{\xi}$, temos que $\lim_{i \rightarrow \infty} \mathbf{z}_i = \boldsymbol{\xi}$.

De fato, suponha que (\mathbf{z}_n) não convirja para $\boldsymbol{\xi}$. Então existe um ϵ , uma subsequência (\mathbf{z}_{n_k}) , e um inteiro N_0 tal que

$$\|\mathbf{z}_{n_k} - \boldsymbol{\xi}\| > \epsilon.$$

para todo $n_k > N_0$. Isto implica que existem infinitos elementos de (\mathbf{z}_n) distando mais que ϵ de $\boldsymbol{\xi}$. Logo existem infinitos elementos de (\mathbf{x}_n) ou de (\mathbf{y}_n) distando mais que ϵ de $\boldsymbol{\xi}$. mas isto contradiz o fato de que $\lim_{n \rightarrow \infty} \mathbf{x}_n = \boldsymbol{\xi}$ e $\lim_{n \rightarrow \infty} \mathbf{y}_n = \boldsymbol{\xi}$.

A noção de subsequência, combinada com o conceito de ponto de acumulação e o Teorema de Bolzano–Weierstrass (Teorema 2.4.2) pode ser aplicada como o exemplo abaixo nos mostra.

EXEMPLO 3.19. Suponha que (\mathbf{x}_k) é uma sequência limitada de elementos *distintos*, e que o conjunto $\{\mathbf{x}_k : k \in \mathbb{N}\}$ tem exatamente um ponto de acumulação. Então (\mathbf{x}_k) é convergente.

De fato, seja \mathbf{x} o ponto de acumulação da sequência. Por absurdo, suponha que (\mathbf{x}_k) não converge para \mathbf{x} . Então existe $\epsilon > 0$ e uma subsequência (\mathbf{x}_{k_j}) tal que

$$(3.2.3) \quad \|\mathbf{x}_{k_j} - \mathbf{x}\| > \epsilon \quad \text{para todo } k \in \mathbb{N}.$$

Mas então o conjunto $\{\mathbf{x}_{k_j} : j \in \mathbb{N}\}$ é infinito pois os \mathbf{x}_{k_j} são distintos e portanto pelo Teorema de Bolzano–Weierstrass ele tem pelo menos um ponto de acumulação, que é diferente de \mathbf{x} , uma contradição com \mathbf{x} ser o único ponto de acumulação de $\{\mathbf{x}_k : k \in \mathbb{N}\}$.

Finalmente mostramos um importante resultado que nos garante convergência de alguma subsequência mesmo quando a sequência original não converge. É o análogo para sequências do Teorema de Bolzano–Weierstrass (Teorema 2.4.2).

TEOREMA 3.2.3 (Bolzano–Weierstrass para sequências). *Toda sequência limitada em \mathbb{R}^n tem pelo menos uma subsequência convergente.*

DEMONSTRAÇÃO. Seja (\mathbf{x}_k) sequência em \mathbb{R}^n e $S = \{\mathbf{x}_k : k \in \mathbb{N}\}$. Então S é finito ou não. Se S for finito, então existe pelo menos um elemento $\boldsymbol{\xi} \in S$ tal que $\boldsymbol{\xi} = \mathbf{x}_{k_1} = \mathbf{x}_{k_2} = \mathbf{x}_{k_3} = \dots$ para algum k_1, k_2, k_3, \dots em \mathbb{N} . Neste caso, a subsequência constante (\mathbf{x}_{k_j}) é convergente.

Se S for infinito, e como este conjunto é limitado por hipótese, então o Teorema de Bolzano–Weierstrass 2.4.2 garante a existência de pelo menos um ponto \mathbf{x} de acumulação de S . Como \mathbf{x} é ponto de acumulação, então para todo $j \in \mathbb{N}$ existe pelo menos um ponto em $S \cap B_{1/j}(\mathbf{x})$, i.e., existe $k_j \in \mathbb{N}$ tal que $\mathbf{x}_{k_j} \in S \cap B_{1/j}(\mathbf{x})$. Então, dado $\epsilon > 0$, para $1/J < \epsilon$ temos

$$\|\mathbf{x} - \mathbf{x}_{k_j}\| < \frac{1}{j} < \frac{1}{J} < \epsilon \quad \text{para todo } j \geq J.$$

Logo, a subsequência (\mathbf{x}_{k_j}) é convergente. □

3.3. Sequências de Cauchy

Um conceito importante tratando-se de sequências é o de sequências de Cauchy. Formalmente, dizemos que uma sequência (\mathbf{x}_k) é *de Cauchy* se para todo $\epsilon > 0$ existe $K^* \in \mathbb{N}$ tal que

$$|\mathbf{x}_k - \mathbf{x}_m| < \epsilon \quad \text{para todo } k, m \geq K^*.$$

Usando os lemas a seguir, mostraremos que uma sequência é convergente se e somente se é de Cauchy.

LEMA 3.3.1. Toda sequência convergente é de Cauchy.

DEMONSTRAÇÃO. Seja (\mathbf{x}_k) sequência convergente, e \mathbf{x} o seu limite. Então, dado $\epsilon > 0$, existe $K^* \in \mathbb{N}$ tal que $|\mathbf{x}_k - \mathbf{x}| < \epsilon/2$ para todo $k \geq K^*$. Portanto,

$$|\mathbf{x}_k - \mathbf{x}_m| \leq |\mathbf{x}_k - \mathbf{x}| + |\mathbf{x} - \mathbf{x}_m| < \epsilon \quad \text{se } k, m \geq K^*.$$

Logo (\mathbf{x}_k) é de Cauchy. □

LEMA 3.3.2. Toda sequência de Cauchy é limitada.

DEMONSTRAÇÃO. Seja (\mathbf{x}_k) sequência de Cauchy. Então, considerando $\epsilon = 1$, temos que existe $K^* \in \mathbb{N}$ tal que $|\mathbf{x}_{K^*} - \mathbf{x}_k| < 1$ para todo $k > K^*$. Logo, para $k > K^*$ temos

$$|\mathbf{x}_k| \leq |\mathbf{x}_k - \mathbf{x}_{K^*}| + |\mathbf{x}_{K^*}| < 1 + |\mathbf{x}_{K^*}|.$$

Definindo $C = \max\{|\mathbf{x}_1|, \dots, |\mathbf{x}_{K^*-1}|, 1 + |\mathbf{x}_{K^*}|\}$, temos imediatamente que $|\mathbf{x}_k| \leq C$ para todo $k \in \mathbb{N}$. Portanto a sequência é limitada. \square

Finalmente podemos enunciar a equivalência entre convergência e o critério de Cauchy.

TEOREMA 3.3.3 (Critério de convergência de Cauchy). *Uma sequência é convergente se e somente se é de Cauchy.*

DEMONSTRAÇÃO. Já vimos no Lema 3.3.1 que se uma sequência é convergente, ela é de Cauchy.

Suponha agora que (\mathbf{x}_k) é sequência de Cauchy. Pelo Lema 3.3.2, a sequência é limitada, e pelo Teorema de Bolzano–Weierstrass (Teorema 3.2.3), existe uma subsequência (\mathbf{x}_{k_j}) convergente. Seja $\mathbf{x} = \lim_{k_j \rightarrow \infty} \mathbf{x}_{k_j}$. Quero mostrar que $\mathbf{x} = \lim_{k \rightarrow \infty} \mathbf{x}_k$. Seja $\epsilon > 0$. Como (\mathbf{x}_k) é de Cauchy, temos que existe $K^* \in \mathbb{N}$ tal que

$$(3.3.1) \quad \|\mathbf{x}_k - \mathbf{x}_m\| \leq \frac{\epsilon}{2} \quad \text{para todo } k, m \geq K^*.$$

Como (\mathbf{x}_{k_j}) é convergente, então existe $m \in \{k_1, k_2, \dots\}$ tal que $m > K^*$, e

$$\|\mathbf{x} - \mathbf{x}_m\| < \frac{\epsilon}{2}.$$

Como $m > K^*$ temos também de (3.3.1) que $\|\mathbf{x}_k - \mathbf{x}_m\| \leq \epsilon/2$ para todo $k \geq K^*$. Finalmente, para todo $k \geq K^*$ temos

$$\|\mathbf{x} - \mathbf{x}_k\| \leq \|\mathbf{x} - \mathbf{x}_m\| + \|\mathbf{x}_m - \mathbf{x}_k\| < \epsilon.$$

Concluimos que (\mathbf{x}_k) converge. \square

EXEMPLO 3.20. Considere $x_1 = 1$, $x_2 = 2$ e $x_j = (x_{j-1} + x_{j-2})/2$ para $j \geq 3$. Então mostraremos que (x_j) converge pois é de Cauchy. Mostramos primeiro que

$$(3.3.2) \quad |x_j - x_{j+1}| = \frac{1}{2^{j-1}}, \quad \text{para } j \in \mathbb{N}.$$

Note que (3.3.2) é válido para $j = 1$. Supondo também válida para $j = k$, i.e., que

$$(3.3.3) \quad |x_k - x_{k+1}| = \frac{1}{2^{k-1}},$$

temos

$$|x_{k+1} - x_{k+2}| = |x_{k+1} - \frac{1}{2}(x_{k+1} + x_k)| = |\frac{1}{2}(x_{k+1} - x_k)| = \frac{1}{2^k},$$

onde usamos (3.3.3) na última igualdade. Concluimos por indução que (3.3.2) é válida.

Tendo (3.3.2) sido demonstrado, basta agora, dado ϵ , tomar K^* tal que $2^{K^*-2}\epsilon > 1$. Neste caso, se $j \geq i \geq K^*$, tem-se

$$\begin{aligned}
 (3.3.4) \quad |x_j - x_i| &\leq |x_j - x_{j-1}| + |x_{j-1} - x_{j-2}| + |x_{j-2} - x_{j-3}| + \cdots + |x_{i+1} - x_i| \\
 &= \frac{1}{2^{j-2}} + \frac{1}{2^{j-3}} + \frac{1}{2^{j-4}} + \cdots + \frac{1}{2^{i-1}} = \frac{1}{2^{i-1}} \left(\frac{1}{2^{j-i-1}} + \frac{1}{2^{j-i-2}} + \frac{1}{2^{j-i-3}} + \cdots + 1 \right) \\
 &= \frac{1}{2^{i-1}} \frac{1 - 1/2^{j-i}}{1 - \frac{1}{2}} \leq \frac{1}{2^{i-2}} < \epsilon,
 \end{aligned}$$

EXEMPLO 3.21. Em geral, se (\mathbf{x}_i) é tal que $\|\mathbf{x}_{i+1} - \mathbf{x}_i\| < c_i$, onde $S_i = \sum_{k=1}^i c_k$ é convergente, então (\mathbf{x}_i) é convergente. De fato, mostramos abaixo que a sequência é de Cauchy, e portanto converge. Note que para $i > j$, temos

$$(3.3.5) \quad \|\mathbf{x}_i - \mathbf{x}_j\| \leq \|\mathbf{x}_i - \mathbf{x}_{i-1}\| + \|\mathbf{x}_{i-1} - \mathbf{x}_{i-2}\| + \cdots + \|\mathbf{x}_{j+1} - \mathbf{x}_j\| \leq c_{i-1} + c_{i-2} + \cdots + c_j = S_{i-1} - S_{j-1}.$$

Como S_i converge, então é de Cauchy. Logo dado $\epsilon > 0$, existe $K^* \in \mathbb{N}$ tal que $i > j > K^*$ implica que $|S_{i-1} - S_{j-1}| < \epsilon$. Logo, por (3.3.5) temos que $i > j > K^*$ implica que $\|\mathbf{x}_i - \mathbf{x}_j\| < \epsilon$ e (\mathbf{x}_i) é de Cauchy.

3.4. Resultados Topológicos

O conceito de sequência é importante também para caracterizar conjuntos quanto à sua topologia. Apresentamos abaixo alguns resultados nesta direção.

Podemos por exemplo usar sequências para caracterizar conjuntos fechados, como o resultado abaixo mostra.

LEMA 3.4.1 (Conjuntos fechados). Seja $F \subset \mathbb{R}^n$. As afirmativas abaixo são equivalentes.

- (1) F é fechado em \mathbb{R}^n .
- (2) Se (\mathbf{x}_k) é sequência convergente, com $\mathbf{x}_k \in F$ para todo $k \in \mathbb{N}$, então $\lim_{k \rightarrow \infty} \mathbf{x}_k \in F$.

DEMONSTRAÇÃO. (1) \Rightarrow (2) (*por contradição*) Suponha F fechado em \mathbb{R}^n , e seja (\mathbf{x}_k) sequência em F com $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}$. Suponha $\mathbf{x} \notin F$. Como $\mathcal{C}(F)$ é aberto, existe aberto V contendo \mathbf{x} tal que $V \cap F = \emptyset$. Logo, para todo $k \in \mathbb{N}$, temos $\mathbf{x}_k \notin V$, uma contradição com $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}$. Portanto $\mathbf{x} \in F$.

(1) \Leftarrow (2) (*por contradição*) Suponha que $\mathcal{C}(F)$ não seja aberto. Então existe $\mathbf{x} \in \mathcal{C}(F)$ tal que para todo $k \in \mathbb{N}$ existe um ponto em $\mathbf{x}_k \in B_{1/k}(\mathbf{x}) \cap F$. Logo (\mathbf{x}_k) é uma sequência em F que converge para \mathbf{x} . Por hipótese, temos que $\mathbf{x} \in F$, uma contradição com $\mathbf{x} \in \mathcal{C}(F)$. Portanto $\mathcal{C}(F)$ é aberto, e F é fechado. \square

A caracterização de fechados dada pelo Lema 3.4.1, é útil na bela aplicação que descrevemos abaixo. Seja $V \subset \mathbb{R}^n$ um *subespaço vetorial* do \mathbb{R}^n , i.e., V é espaço vetorial com as operações “herdadas” do \mathbb{R}^n . Então, dado um ponto $\mathbf{x} \in \mathbb{R}^n$, pode-se perguntar se existe algum ponto em V que minimize a distância entre \mathbf{x} e V , i.e, se existe $\mathbf{x}_* \in V$ tal que

$$(3.4.1) \quad \|\mathbf{x} - \mathbf{x}_*\| = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in V\}.$$

Outra pergunta natural é se \mathbf{x}_* é único.

Supondo que a norma $\|\cdot\|$ é induzida por um produto interno, a resposta é afirmativa para ambas perguntas, existência e unicidade, como nos mostra o resultado abaixo.

LEMA 3.4.2. Seja V subespaço vetorial do \mathbb{R}^n e $\mathbf{x} \in \mathbb{R}^n$. Então existe um único $\mathbf{x}_* \in V$ satisfazendo (3.4.1).

DEMONSTRAÇÃO. Vamos primeiro mostrar a existência. Note que V é não vazio, pois $0 \in V$, e portanto

$$d = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in V\}$$

está bem definido. Para $k \in \mathbb{N}$, seja $\mathbf{x}_k \in V$ tal que $\|\mathbf{x} - \mathbf{x}_k\| < d + 1/k$. Usando a lei do paralelograma, ver exercício 2.11, temos que

$$\|2\mathbf{x} - \mathbf{x}_i - \mathbf{x}_j\|^2 + \|\mathbf{x}_i - \mathbf{x}_j\|^2 = 2\|\mathbf{x} - \mathbf{x}_i\|^2 + 2\|\mathbf{x} - \mathbf{x}_j\|^2,$$

para todo $i, j \in \mathbb{N}$. Mas V é subespaço vetorial, logo $(\mathbf{x}_i + \mathbf{x}_j)/2 \in V$, e portanto,

$$2d \leq 2\|\mathbf{x} - (\mathbf{x}_i + \mathbf{x}_j)/2\| = \|2\mathbf{x} - \mathbf{x}_i - \mathbf{x}_j\|.$$

Temos então que

$$(3.4.2) \quad \|\mathbf{x}_i - \mathbf{x}_j\|^2 = 2\|\mathbf{x} - \mathbf{x}_i\|^2 + 2\|\mathbf{x} - \mathbf{x}_j\|^2 - \|2\mathbf{x} - \mathbf{x}_i - \mathbf{x}_j\|^2 \\ \leq 2\|\mathbf{x} - \mathbf{x}_i\|^2 + 2\|\mathbf{x} - \mathbf{x}_j\|^2 - 4d^2.$$

Mas então, (\mathbf{x}_k) é de Cauchy, pois $\|\mathbf{x} - \mathbf{x}_i\| \rightarrow d$. De fato, dado $\epsilon > 0$, existe K^* tal que para todo $k \geq K^*$ tem-se $\|\mathbf{x} - \mathbf{x}_k\|^2 - d^2 < \epsilon/2$. Logo, por (3.4.2), $\|\mathbf{x}_i - \mathbf{x}_j\|^2 < \epsilon$ se $i, j \geq K^*$.

Seja $\mathbf{x}_* = \lim_{k \rightarrow \infty} \mathbf{x}_k$. Mas V é fechado em \mathbb{R}^n (por quê? ver exercício 3.13), e portanto $\mathbf{x}_* \in V$. Finalmente, para todo $k \in \mathbb{N}$,

$$d \leq \|\mathbf{x} - \mathbf{x}_*\| \leq \|\mathbf{x} - \mathbf{x}_k\| + \|\mathbf{x}_k - \mathbf{x}_*\|.$$

Tomando o limite $k \rightarrow \infty$, temos $\|\mathbf{x} - \mathbf{x}_*\| = d$, como queríamos.

Para mostrar a unicidade, seja $\mathbf{y} \in V$, com $\|\mathbf{x} - \mathbf{y}\| = d$. Então $(\mathbf{y} + \mathbf{x}_*)/2 \in V$, e

$$d^2 \leq \|\mathbf{x} - (\mathbf{y} + \mathbf{x}_*)/2\|^2.$$

Portanto, usando novamente a lei do paralelograma, temos

$$4d^2 + \|\mathbf{y} - \mathbf{x}_*\|^2 \leq \|2\mathbf{x} - \mathbf{y} - \mathbf{x}_*\|^2 + \|\mathbf{y} - \mathbf{x}_*\|^2 = 2\|\mathbf{x} - \mathbf{y}\|^2 + 2\|\mathbf{x} - \mathbf{x}_*\|^2 = 4d^2.$$

Logo $\|\mathbf{y} - \mathbf{x}_*\| = 0$ e $\mathbf{y} = \mathbf{x}_*$. □

OBSERVAÇÃO. Pode-se mostrar também que \mathbf{x}_* é o único vetor de V tal que $\mathbf{x} - \mathbf{x}_*$ é ortogonal a V , i.e.,

$$(\mathbf{x} - \mathbf{x}_*) \cdot \mathbf{y} = 0$$

para todo $\mathbf{y} \in V$.

Também os conceito de fronteira de um conjunto e o de conjunto aberto pode ser dado através de seqüências.

LEMA 3.4.3 (Pontos de fronteira). Um ponto \mathbf{x} é de fronteira de $\Omega \subset \mathbb{R}^n$ se e somente se existe seqüência em Ω e seqüência em $C(\Omega)$, ambas convergentes para \mathbf{x} .

LEMA 3.4.4 (Conjuntos abertos). Seja $\Omega \subset \mathbb{R}^n$. As afirmativas abaixo são equivalentes.

- (1) Ω é aberto em \mathbb{R}^n .
- (2) Seja $\mathbf{x} \in \Omega$ e (\mathbf{x}_k) contida em \mathbb{R}^n com $\mathbf{x}_k \rightarrow \mathbf{x}$. Então existe K^* tal que

$$k \geq K^* \implies \mathbf{x}_k \in \Omega.$$

3.5. Sequências contráteis e o método das aproximações sucessivas

Dizemos que uma sequência (\mathbf{x}_k) é *contrátil* se existem número real $\lambda < 1$ e um natural K^* tais que

$$\|\mathbf{x}_{k+2} - \mathbf{x}_{k+1}\| \leq \lambda \|\mathbf{x}_{k+1} - \mathbf{x}_k\|$$

para todo $k > K^*$.

TEOREMA 3.5.1. *Toda sequência contrátil é convergente*

DEMONSTRAÇÃO. Seja (\mathbf{x}_k) sequência contrátil com constante $\lambda < 1$. Sem perda de generalidade, supomos nesta demonstração que $K^* = 1$, isto é

$$\|\mathbf{x}_{k+2} - \mathbf{x}_{k+1}\| \leq \lambda \|\mathbf{x}_{k+1} - \mathbf{x}_k\|$$

para todo $k \in \mathbb{N}$. Então,

$$\|\mathbf{x}_{k+2} - \mathbf{x}_{k+1}\| \leq \lambda \|\mathbf{x}_{k+1} - \mathbf{x}_k\| \leq \lambda^2 \|\mathbf{x}_k - \mathbf{x}_{k-1}\| \leq \cdots \leq \lambda^k \|\mathbf{x}_2 - \mathbf{x}_1\|.$$

Logo, para $m \in \mathbb{N}$ e $k \geq m$ temos

$$\begin{aligned} \|\mathbf{x}_k - \mathbf{x}_m\| &\leq \|\mathbf{x}_k - \mathbf{x}_{k-1}\| + \|\mathbf{x}_{k-1} - \mathbf{x}_{k-2}\| + \cdots + \|\mathbf{x}_{m+1} - \mathbf{x}_m\| \\ &\leq (\lambda^{k-2} + \lambda^{k-3} + \cdots + \lambda^{m-1}) \|\mathbf{x}_2 - \mathbf{x}_1\| = \lambda^{m-1} (\lambda^{k-m-1} + \lambda^{k-m-2} + \cdots + 1) \|\mathbf{x}_2 - \mathbf{x}_1\| \\ &= \lambda^{m-1} \frac{1 - \lambda^{k-m}}{1 - \lambda} \|\mathbf{x}_2 - \mathbf{x}_1\| \leq \frac{\lambda^{m-1}}{1 - \lambda} \|\mathbf{x}_2 - \mathbf{x}_1\|. \end{aligned}$$

Logo, dado $\epsilon > 0$ se $K^* \in \mathbb{N}$ é tal que

$$\frac{\lambda^{K^*-1}}{1 - \lambda} \|\mathbf{x}_2 - \mathbf{x}_1\| < \epsilon,$$

então $\|\mathbf{x}_k - \mathbf{x}_m\| < \epsilon$ para todo $m \geq K^*$, $k \geq K^*$. Portanto a sequência é de Cauchy e é convergente \square

EXEMPLO 3.22. Seja a sequência definida por

$$x_0 = a > 0, \quad x_{n+1} = 1 + \frac{1}{x_n}.$$

Queremos mostrar que (x_n) é contrátil, e portanto convergente.

Seja $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ dada por $f(x) = 1 + 1/x$. Então a sequência é definida por $x_{n+1} = f(x_n)$, e temos portanto que $x^* = (1 + \sqrt{5})/2$ é a única solução em \mathbb{R}^+ para a equação $x = f(x)$. Usaremos mais tarde o fato de que $x > x^*$ implica em $x^2 > x + 1$. Note ainda que f é tal que

$$(3.5.1) \quad x > y \implies f(x) < f(y),$$

e que se $x, y \in \mathbb{R}^+$ e $c < \min\{x, y\}$, então

$$(3.5.2) \quad |f(x) - f(y)| = \left| \frac{1}{x} - \frac{1}{y} \right| = \frac{|x - y|}{xy} \leq \frac{|y - x|}{c^2}.$$

A fim de utilizar (3.5.2), mostraremos que (x_n) é limitada inferiormente por algum número maior que um.

Temos então três possibilidades: $a = x^*$, $a > x^*$ ou $a < x^*$. Quando $a = x^*$, a série é trivialmente convergente pois temos $x_1 = x_2 = \dots = x^*$. Suponha então que $x_0 = a > x^*$. A análise para $a < x^*$ é similar.

Então $x_1 = f(x_0) < f(x^*) = x^*$. Por indução temos que $x_{2n-2} > x^*$ e $x_{2n-1} < x^*$. De fato, como estas desigualdades são verdadeiras para $n = 1$ e supondo também corretas para $n = k$ temos $x_{2k} = f(x_{2k-1}) > f(x^*) = x^*$ e $x_{2k+1} = f(x_{2k}) < f(x^*) = x^*$, como queríamos demonstrar.

Temos então $x_0 = a$, $x_1 = (a + 1)/a$, e

$$x_2 = 1 + \frac{1}{x_1} = \frac{2a + 1}{a + 1} < \frac{a + a^2}{a + 1} = a = x_0,$$

onde usamos que $a + 1 < a^2$. Da mesma forma, $x_3 = 1 + 1/x_2 > 1 + 1/x_0 = x_1$. Portanto temos que para $n = 1$ vale $x_{2n} < x_{2n-2}$ e $x_{2n+1} > x_{2n-1}$. Supondo estas duas desigualdades para $n = k$ temos

$$x_{2k+2} = 1 + 1/x_{2k+1} < 1 + 1/x_{2k-1} = x_{2k}, \quad x_{2k+3} = 1 + 1/x_{2k+2} > 1 + 1/x_{2k} = x_{2k+1},$$

como queríamos demonstrar.

Concluimos que (x_{2n-1}) é sequência não decrescente, e que $|x_{2n}| > x^* > x_1$ para todo $n \in \mathbb{N}$. Portanto (x_n) é limitada inferiormente por x_1 .

Aplicando agora (3.5.2), temos

$$|x_{k+1} - x_k| = |f(x_k) - f(x_{k-1})| \leq \frac{1}{x_1^2} |x_k - x_{k-1}|.$$

Como $x_1 = 1 + 1/a > 1$, então (x_n) é contrátil e portanto converge.

Para achar o valor limite, basta resolver $x = f(x)$, e temos que $\lim_{n \rightarrow \infty} x_n = x^*$.

Em várias aplicações importantes é necessário achar um *ponto fixo*, i.e., uma solução do tipo $\mathbf{x} = T(\mathbf{x})$, onde $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é dada. É natural perguntar-se se dado algum ponto inicial \mathbf{x}_0 , a sequência gerada por

$$\mathbf{x}_k = T(\mathbf{x}_{k-1}), \quad k \in \mathbb{N},$$

converge para um ponto fixo. Esta forma de determinar pontos fixos é denominada *método das aproximações sucessivas*.

No caso de T ser uma “contração”, (\mathbf{x}_k) será contrátil, e portanto convergente. É exatamente isto que mostraremos a seguir.

DEFINIÇÃO 3.5.2. *Seja $A \subset \mathbb{R}^n$. Dizemos que uma função $T : A \rightarrow \mathbb{R}^n$ é uma contração se existir $\lambda < 1$ tal que*

$$\|T(\mathbf{y}) - T(\mathbf{x})\| \leq \lambda \|\mathbf{y} - \mathbf{x}\|$$

para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$.

Temos então o seguinte resultado.

TEOREMA 3.5.3. *Seja $A \subset \mathbb{R}^n$ fechado, e $T : A \rightarrow A$ uma contração. Então T possui um e somente um ponto fixo em A . Além disto, para qualquer $\mathbf{x}_0 \in A$, a sequência definida por*

$$(3.5.3) \quad \mathbf{x}_k = T(\mathbf{x}_{k-1}), \quad k \in \mathbb{N},$$

converge para o ponto fixo de T em A .

DEMONSTRAÇÃO. Suponha que exista $\lambda < 1$ tal que

$$\|T(\mathbf{y}) - T(\mathbf{x})\| \leq \lambda \|\mathbf{y} - \mathbf{x}\|$$

para todo $\mathbf{x}, \mathbf{y} \in A$.

Mostraremos primeiro a unicidade. Dados dois pontos fixos \mathbf{x} e \mathbf{y} de T em A , temos que

$$\|\mathbf{x} - \mathbf{y}\| = \|T(\mathbf{x}) - T(\mathbf{y})\| \leq \lambda \|\mathbf{x} - \mathbf{y}\|,$$

o que só é possível se $\mathbf{x} = \mathbf{y}$, e portanto o ponto fixo, se existir, é único.

Note que (\mathbf{x}_k) é contrátil pois

$$\|\mathbf{x}_{k+2} - \mathbf{x}_{k+1}\| = \|T(\mathbf{x}_{k+1}) - T(\mathbf{x}_k)\| \leq \lambda \|\mathbf{x}_{k+1} - \mathbf{x}_k\|.$$

Logo (\mathbf{x}_k) converge, e seja \mathbf{x}^* seu limite. Como A é fechado, então $\mathbf{x}^* \in A$. Para mostrar que \mathbf{x}^* é ponto fixo de T , note que para todo $k \in \mathbb{N}$, temos que

$$\begin{aligned} \|x^* - T(x^*)\| &\leq \|\mathbf{x}^* - \mathbf{x}_k\| + \|\mathbf{x}_k - T(x^*)\| = \|\mathbf{x}^* - \mathbf{x}_k\| + \|T(\mathbf{x}_{k-1}) - T(x^*)\| \\ &\leq \|\mathbf{x}^* - \mathbf{x}_k\| + \lambda \|\mathbf{x}_{k-1} - x^*\|. \end{aligned}$$

Tomando o limite $k \rightarrow \infty$ dos dois lados da desigualdade obtemos que $\|x^* - T(x^*)\| = 0$, e portanto $x^* = T(x^*)$. \square

3.6. Sequências em \mathbb{R}

Vários conceitos e propriedades de sequências fazem sentido em \mathbb{R} , mas não em \mathbb{R}^n para n maior que um. Por exemplo, o conceito de monotonicidade, a definição de \limsup , \liminf não se generalizam no \mathbb{R}^n . E propriedades, como por exemplo *o limite de uma sequência positiva é não negativo* também não. Este por sinal é o primeiro resultado que apresentamos a seguir.

Outro resultado importante para se tentar achar um “candidato” a limite nos diz que se temos uma sequência “sanduichadas” entre outras duas sequências convergentes que têm o mesmo limite, então a sequência do meio converge e tem também o mesmo limite.

LEMA 3.6.1. Seja (x_n) convergente com $\lim x_n = x$. Se existe $K^* \in \mathbb{N}$ tal que $x_n \geq 0$ para todo $n \geq K^*$, então $x \geq 0$.

DEMONSTRAÇÃO. (por contradição) Suponha que $x < 0$. Seja então $\epsilon = -x/2 > 0$. Como (x_n) converge para x , seja $\hat{K} \in \mathbb{N}$ tal que $|x_n - x| < \epsilon$ para todo $n \geq \hat{K}$. Seja $\bar{K} = \max\{K^*, \hat{K}\}$. Logo, $x_{\bar{K}} \in (x - \epsilon, x + \epsilon)$, isto é, $x_{\bar{K}} < x + \epsilon = x/2 < 0$. Obtivemos então uma contradição pois $\bar{K} \geq K^*$, e então $x_{\bar{K}}$ não pode ser negativo. \square

OBSERVAÇÃO. Note que o resultado acima não pode ser modificado tal que

$$x_n > 0 \text{ para todo } n \geq K^*, \text{ então } \lim x_n > 0.$$

De fato, considere a sequência $(1/k)$ de números positivos, mas com limite igual a zero.

COROLÁRIO 3.6.2. Se (x_n) e (y_n) são convergentes com $\lim x_n = x$ e $\lim y_n = y$, e se existe $K^* \in \mathbb{N}$ tal que $x_n \geq y_n$ para todo $n > K^*$, 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 3.6.1. \square

LEMA 3.6.3 (sanduíche de sequências). Sejam (x_n) , (y_n) e (z_n) sequências tais que $x_n \leq y_n \leq z_n$ para todo $n > K^*$, para algum $K^* \in \mathbb{N}$. Suponha 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 K^* tal que $|x_n - a| < \epsilon$ e $|z_n - a| < \epsilon$ para todo $n > K^*$. Logo

$$-\epsilon < x_n - a \leq y_n - a \leq z_n - a < \epsilon \implies |y_n - a| < \epsilon$$

para todo $n > K^*$, como queríamos demonstrar. \square

EXEMPLO 3.23. $\lim_{n \rightarrow \infty} ((\sin n)/n) = 0$ pois como $-1 \leq \sin n \leq 1$, então

$$-1/n \leq (\sin n)/n \leq 1/n,$$

e o resultado segue do lema 3.6.3.

LEMA 3.6.4 (teste da razão). Seja (x_n) sequência 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 $L < r < 1$, e seja $\epsilon = r - L > 0$. Portanto existe K^* tal que $x_{n+1}/x_n < L + \epsilon = r$ para todo $n \geq K^*$. Logo,

$$0 < x_{n+1} < x_n r < x_{n-1} r^2 < x_{n-2} r^3 < \dots < x_{K^*} r^{n-K^*+1} \quad \text{para todo } n \geq K^*.$$

Se $c = x_{K^*} r^{-K^*}$, então $0 < x_{n+1} < c r^{n+1}$. O resultado segue do Corolário 3.1.6, pois como $r < 1$, então $\lim_{n \rightarrow \infty} r^n = 0$. \square

COROLÁRIO 3.6.5. 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 $K^* \in \mathbb{N}$ tal que

$$n \geq K^* \implies |x_n| > C.$$

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 K^* tal que

$$n \geq K^* \implies |y_n| < \frac{1}{C}.$$

Portanto para $n \geq K^*$ temos $|x_n| > C$. \square

OBSERVAÇÃO. Observe que no Corolário 3.6.5 acima, (x_n) não é limitada e portanto não converge.

EXEMPLO 3.24. 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 3.25. Note que para $x_n = 1/n$, temos $\lim_{n \rightarrow \infty} x_{n+1}/x_n = 1$ e (x_n) converge. Entretanto, para $y_n = n$, temos $\lim_{n \rightarrow \infty} y_{n+1}/y_n = 1$ mas (y_n) não converge. Portanto o teste não é conclusivo quando o limite da razão entre os termos é um.

3.6.1. 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 3.6.6. Dizemos que uma sequência (x_n) é monótona crescente, ou simplesmente crescente se $x_1 \leq x_2 \leq \dots \leq x_n \leq \dots$. Da mesma forma uma sequência (x_n) é monótona decrescente, ou simplesmente decrescente se $x_1 \geq x_2 \geq \dots \geq x_n \geq \dots$. Finalmente, uma sequência é monótona se for crescente ou decrescente.

EXEMPLO 3.26. $(1, 2, 3, 4, \dots)$ e $(1, 2, 3, 3, 3, 3, \dots)$ são crescentes.

EXEMPLO 3.27. $(1/n)$ é decrescente.

EXEMPLO 3.28. $(-1, 1, -1, 1, -1, \dots)$ não é monótona.

TEOREMA 3.6.7. Uma sequência monótona é convergente se e somente se é limitada.

Além disso, se (x_n) é crescente, então $\lim_{n \rightarrow \infty} (x_n) = \sup\{x_n : n \in \mathbb{N}\}$. Da mesma forma, se (x_n) é decrescente, 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) Suponha (x_n) crescente 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

EXEMPLO 3.29. (a^n) diverge se $a > 1$ pois não é limitada.

EXEMPLO 3.30. (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 3.31. 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 $n = 1$. Suponha 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 3.32. 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. Suponha 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 $|y_n| < 2$. Suponha que $|y_k| < 2$. Logo,

$$|y_{k+1}| = \left| \frac{1}{4}(2y_k + 3) \right| \leq \frac{1}{4}(2|y_k| + 3) < \frac{7}{4} < 2.$$

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 3.33. Seja $0 < a < b$, 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

$$(3.6.1) \quad a_{i+1} > a_i, \quad 0 < a_i < b_i, \quad b_{i+1} < b_i \quad \text{para } i = 0, 1, \dots$$

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 $a_1 = \sqrt{a_0 b_0} > a_0 > 0$. Além disso, $b_1 = (a_0 + b_0)/2 < b_0$ pois $a_0 < b_0$. Portanto (3.6.1) vale para $i = 0$. Suponha que valha também para $i = n$. Então

$$0 < a_n < b_n \implies 0 < (\sqrt{a_n} - \sqrt{b_n})^2 \implies \sqrt{a_n b_n} < \frac{1}{2}(a_n + b_n) \implies a_{n+1} < b_{n+1}.$$

Note também que $a_{n+1} > a_n > 0$. Finalmente,

$$a_{n+2} = \sqrt{a_{n+1} b_{n+1}} > a_{n+1}, \quad b_{n+2} = \frac{a_{n+1} + b_{n+1}}{2} < b_{n+1}.$$

Logo (3.6.1) 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 $A = B$.

3.6.2. Limite superior e inferior. Uma noção importante tratando-se de sequências é a de limites superiores (lim sup) e inferiores (lim inf), que nos dá informações sobre sequências limitadas mesmo quando estas não são convergentes.

Seja (x_n) sequência limitada de reais, e defina

$$V = \{v \in \mathbb{R} : \text{existem finitos } n \in \mathbb{N} \text{ tais que } x_n > v\}.$$

Definimos então

$$\limsup x_n = \inf V.$$

De forma análoga, se

$$W = \{v \in \mathbb{R} : \text{existem finitos } n \in \mathbb{N} \text{ tais que } x_n < v\},$$

definimos

$$\liminf x_n = \sup W.$$

LEMA 3.6.8. Seja (x_n) sequência de reais limitada. Então (x_n) converge para x se e somente se $\limsup x_n = \liminf x_n = x$.

EXEMPLO 3.34. Seja $(x_n) = (-1)^n$. Então $\liminf x_n = -1$ e $\limsup x_n = 1$.

EXEMPLO 3.35. Seja

$$(z_n) = \left((-1)^n + \frac{(-1)^n}{n} \right).$$

Então $\liminf z_n = -1$ e $\limsup z_n = 1$.

3.7. Exercícios

EXERCÍCIO 3.1. Demontre o Lema 3.1.2.

EXERCÍCIO 3.2. Demonstrar o Lema 3.1.9.

EXERCÍCIO 3.3. Seja (\mathbf{x}_k) sequência em \mathbb{R}^n convergente para \mathbf{x} . Mostre que $\{\mathbf{x}_k : k \in \mathbb{N}\} \cup \{\mathbf{x}\}$ é compacto.

EXERCÍCIO 3.4. Mostre que se $k : \mathbb{N} \rightarrow \mathbb{N}$ é uma função estritamente crescente, então $k(i) \geq i$.

EXERCÍCIO 3.5. Dê um exemplo de uma sequência (x_n) em \mathbb{R} tal que toda subsequência convergente de (x_n) convirja para x , mas que (x_n) não seja convergente.

EXERCÍCIO 3.6. Seja (\mathbf{x}_k) sequência em \mathbb{R}^n limitada, e tal que toda subsequência convergente converge para $\mathbf{x} \in \mathbb{R}^n$. Mostre que (\mathbf{x}_k) converge para \mathbf{x} .

EXERCÍCIO 3.7. Seja (\mathbf{x}_k) sequência de Cauchy contendo uma subsequência convergente para \mathbf{x} . Mostre que (\mathbf{x}_k) converge para \mathbf{x} .

EXERCÍCIO 3.8. Sejam (\mathbf{x}_k) e (\mathbf{y}_k) duas sequências de Cauchy em \mathbb{R}^n . Mostre que $\|\mathbf{x}_k - \mathbf{y}_k\|$ converge.

EXERCÍCIO 3.9. Demonstrar o Lema 3.4.3.

EXERCÍCIO 3.10. Demonstrar o Lema 3.4.4.

EXERCÍCIO 3.11. Seja $S \subset \mathbb{R}^n$. Mostre que \mathbf{x} é ponto de acumulação de S se e somente se existe sequência de pontos (\mathbf{x}_j) em $S \setminus \{\mathbf{x}\}$ que converge para \mathbf{x} .

EXERCÍCIO 3.12. Seja (\mathbf{x}_k) sequência convergente de pontos distintos em \mathbb{R}^n , e seja $\mathbf{x} = \lim_{k \rightarrow \infty} \mathbf{x}_k$. Mostre que \mathbf{x} é ponto de acumulação de $\{\mathbf{x}_k : k \in \mathbb{N}\}$.

EXERCÍCIO 3.13. Seja V um subespaço vetorial de \mathbb{R}^n . Mostre que V é fechado em \mathbb{R}^n .

EXERCÍCIO 3.14. Mostre que se $A \subset \mathbb{R}$ e $s = \sup A$, então existe sequência em A convergindo para s .

EXERCÍCIO 3.15. Seja (x_k) sequência monótona em \mathbb{R} , e suponha que (x_k) contenha subsequência convergente. Mostre que (x_k) converge.

EXERCÍCIO 3.16 (Bartle [2]). Seja $x_1 = 1$ e $x_{n+1} = (2+x_n)^{1/2}$. Mostre que x_n é monótona e limitada, e portanto converge. Ache seu limite.

EXERCÍCIO 3.17 (Bartle [2]). Seja $a > 0$ e $x_1 > 0$. Mostre que a sequência dada por $x_{n+1} = (a + x_n)^{1/2}$ converge.

EXERCÍCIO 3.18. Seja (x_k) sequência em \mathbb{R} , limitada. Se $s_i = \sup\{x_j : j \in \mathbb{N}, j \geq i\}$, mostre que

$$\limsup x_k = \inf\{s_i : i \in \mathbb{N}\}.$$

EXERCÍCIO 3.19. Seja (x_k) sequência em \mathbb{R} , limitada, e seja L o conjunto de números reais x tais que existe uma subsequência de (x_k) convergindo para x . Se $L \neq \emptyset$, mostre que $\sup L = \limsup x_k$.

EXERCÍCIO 3.20. Dadas duas sequências limitadas (x_k) e (y_k) , mostre que

$$\limsup(x_k + y_k) \leq \limsup(x_k) + \limsup(y_k).$$

EXERCÍCIO 3.21. Demonstre o Lema 3.6.8.

CAPÍTULO 4

Continuidade e Funções Contínuas

¹ Um dos mais importantes tópicos de análise é o estudo de funções e suas propriedades, em particular a *continuidade*. Seja $\Omega \subseteq \mathbb{R}^m$. Dizemos que uma função $f : \Omega \rightarrow \mathbb{R}^n$ é contínua em $\mathbf{x} \in \Omega$, se para toda vizinhança aberta V de $f(\mathbf{x})$ existir vizinhança aberta U de \mathbf{x} tal que

$$\mathbf{y} \in U \cap \Omega \implies f(\mathbf{y}) \in V.$$

Ver Figura 1. Finalmente, dizemos que f é contínua em $\Omega' \subseteq \Omega$ se f for contínua em todos os pontos de Ω' .

Dividimos o estudo de funções contínuas analisando primeiro propriedades locais, seguido das propriedades globais. A menos que seja explicitamente indicado, neste capítulo utilizaremos a notação acima.

4.1. Propriedades locais

Começamos observando que a *função f é contínua em todo ponto $\mathbf{x} \in \Omega$ que não seja ponto de acumulação de Ω* . De fato, se $\mathbf{x} \in \Omega$ não é ponto de acumulação, existe vizinhança aberta U de \mathbf{x} tal que $\Omega \cap U = \{\mathbf{x}\}$. Logo para toda vizinhança aberta V de $f(\mathbf{x})$, temos que

$$\mathbf{y} \in \Omega \cap U \implies \mathbf{y} = \mathbf{x} \implies f(\mathbf{y}) = f(\mathbf{x}) \in V$$

Logo f é necessariamente contínua em \mathbf{x} .

Abaixo descrevemos outras formas de checar a continuidade de uma função num ponto.

¹Última Atualização: 07/03/2008

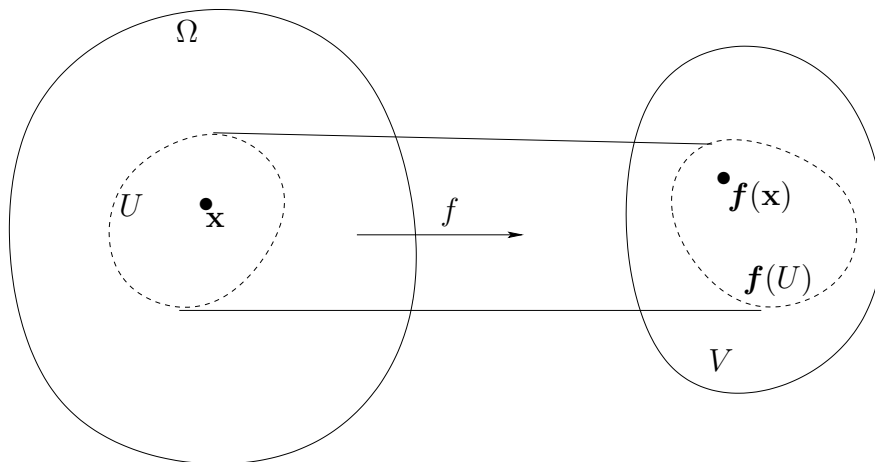


FIG. 1. Continuidade de $f(\mathbf{x})$.

LEMA 4.1.1. Então as afirmativas abaixo são equivalentes.

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

$$\mathbf{y} \in \Omega, \quad \|\mathbf{x} - \mathbf{y}\| < \delta \implies \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| < \epsilon.$$

- (3) Se (\mathbf{x}_k) é sequência em Ω e $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}$, então $\lim_{k \rightarrow \infty} \mathbf{f}(\mathbf{x}_k) = \mathbf{f}(\mathbf{x})$.

Outro resultado importante é o seguinte *critério de descontinuidade*: \mathbf{f} não é contínua em \mathbf{x} se e somente se existe sequência (\mathbf{x}_n) em Ω convergindo para \mathbf{x} mas $(\mathbf{f}(\mathbf{x}_n))$ não convergindo para $\mathbf{f}(\mathbf{x})$.

Uma noção que pode ser útil em algumas ocasiões é a de *limites de funções*. Se \mathbf{x} é ponto de acumulação de Ω , dizemos que \mathbf{p} é o limite de \mathbf{f} em \mathbf{x} se para toda vizinhança aberta V de \mathbf{p} existir vizinhança aberta U de \mathbf{x} tal que

$$\mathbf{y} \in U \cap \Omega, \quad \mathbf{y} \neq \mathbf{x} \implies \mathbf{f}(\mathbf{y}) \in V.$$

Neste caso, escrevemos $\mathbf{p} = \lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})$, e dizemos que \mathbf{f} converge para \mathbf{p} no ponto \mathbf{x} . Uma observação a respeito da definição acima é que só a utilizamos para pontos de acumulação do domínio. Note também que a noção de limite em \mathbf{x} independe do valor de \mathbf{f} em \mathbf{x} . Na verdade, \mathbf{f} não precisa nem estar definida neste ponto.

As seguintes afirmativas são equivalentes:

- (1) $\mathbf{p} = \lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})$
- (2) Para todo $\epsilon > 0$ existe $\delta > 0$ tal que

$$\mathbf{y} \in \Omega \setminus \{\mathbf{x}\}, \quad \|\mathbf{y} - \mathbf{x}\| < \delta \implies \|\mathbf{f}(\mathbf{y}) - \mathbf{p}\| < \epsilon.$$

- (3) Para toda sequência (\mathbf{x}_k) em $\Omega \setminus \{\mathbf{x}\}$, tem-se

$$\mathbf{x}_k \rightarrow \mathbf{x} \implies \mathbf{f}(\mathbf{x}_k) \rightarrow \mathbf{p}.$$

OBSERVAÇÃO. Note algumas diferenças na definição de limite de função e continuidade num ponto \mathbf{x} :

- (1) Para definir limite, a função não precisa estar definida em \mathbf{x} , e mesmo que esteja, o valor não tem importância. Mas faz parte da definição que \mathbf{x} seja ponto de acumulação do domínio da função.
- (2) Na definição de continuidade, a função tem que estar definida em \mathbf{x} , mas este ponto não necessariamente é de acumulação.

Se \mathbf{x} for ponto de acumulação de Ω , então

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

EXEMPLO 4.1. $g(x) = x$ é contínua em \mathbb{R} . De fato, para todo $c \in \mathbb{R}$, temos $\lim_{x \rightarrow c} g(x) = c = g(c)$.

EXEMPLO 4.2. Considere $\mathbf{A} \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, isto é, $\mathbf{A} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é aplicação linear e a norma $\|\mathbf{A}\|$ está bem definida (ver página 12). Seja $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ 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} é contínua.

De fato, dado $\epsilon > 0$, seja $\delta = \epsilon/\|\mathbf{A}\|$. Logo, para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ tem-se

$$\|\mathbf{x} - \mathbf{y}\| < \delta \implies \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| = \|\mathbf{A}(\mathbf{x} - \mathbf{y})\| \leq \|\mathbf{A}\| \|\mathbf{x} - \mathbf{y}\| < \|\mathbf{A}\| \delta = \epsilon.$$

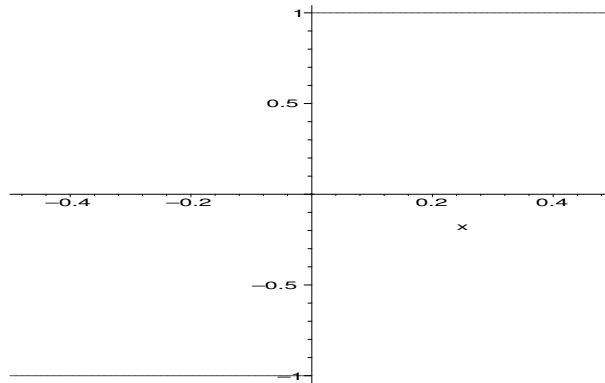


FIG. 2. Gráfico de $\text{sgn}(x)$, que é descontínua em $x = 0$.

EXEMPLO 4.3. Seja $A \subseteq \mathbb{R}^m$ conjunto fechado, e $\mathbf{f} : A \rightarrow \mathbb{R}^n$ contínua em A , e seja (\mathbf{x}_k) sequência de Cauchy em A . Então $(\mathbf{f}(\mathbf{x}_k))$ é sequência de Cauchy em \mathbb{R}^n .

Realmente, como (\mathbf{x}_k) é de Cauchy, então converge. Seja \mathbf{x} seu limite. Como A é fechado, então $\mathbf{x} \in A$. Posto que \mathbf{f} é contínua em A , e portanto em \mathbf{x} , então $\mathbf{f}(\mathbf{x}_k)$ converge para $\mathbf{f}(\mathbf{x})$. Logo, $\mathbf{f}(\mathbf{x}_k)$ é convergente. Como toda sequência convergente é de Cauchy, temos que $\mathbf{f}(\mathbf{x}_k)$ é de Cauchy.

EXEMPLO 4.4. Seja

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

como na figura 2.

Tomando-se as sequências $(-1/n)$ e $(1/n)$, ambas convergindo para $c = 0$ mas nunca atingindo este valor, tem-se $(\text{sgn}(-1/n)) = -1$ e $(\text{sgn}(1/n)) = 1$. Então esta função não tem limite em $c = 0$, pois se o limite existe, este tem que ser único. Portanto, a função $\text{sgn}(x)$ não é contínua no zero, já que não existe $\lim_{x \rightarrow 0} \text{sgn}(x)$.

EXEMPLO 4.5. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por

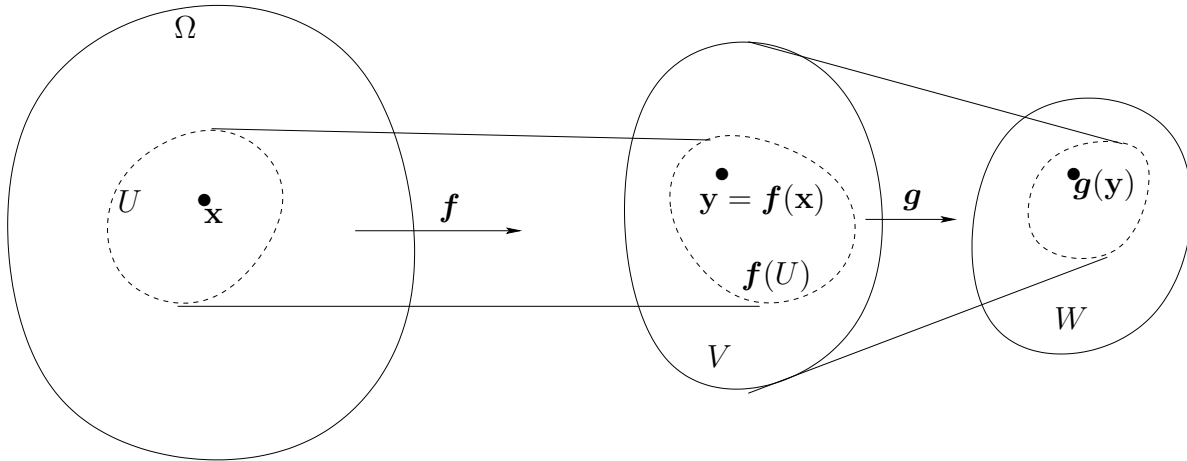
$$f(x) = \begin{cases} 1 & \text{se } x \in \mathbb{Q}, \\ 0 & \text{caso contrário,} \end{cases}$$

é descontínua para todo $x \in \mathbb{R}$. Para mostrar isto, suponha $x \in \mathbb{Q}$, 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. Seja $\mathbf{x} \notin \Omega$ ponto de acumulação de Ω . Se existir $\lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})$, então definimos $\mathbf{f}(\mathbf{x})$ como sendo este limite, e \mathbf{f} será contínua em \mathbf{x} .

EXEMPLO 4.6. Considere a função

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

FIG. 3. Continuidade em $x = 0$.

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 4.7. 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.

4.1.1. Composição de funções. Em geral, se \mathbf{f} e \mathbf{g} são contínuas, então $\mathbf{f} + \mathbf{g}$, $\mathbf{f} - \mathbf{g}$, e no caso unidimensional, $f g$ também o são. Da mesma forma, se $h : \Omega \rightarrow \mathbb{R}$ é tal que $h(\mathbf{x}) \neq 0$ para todo \mathbf{x} do domínio, então \mathbf{f}/h é contínua. O próximo resultado garante que a *composição* de funções contínuas também é contínua. Denotamos a composição de uma função \mathbf{f} com \mathbf{g} por $\mathbf{g} \circ \mathbf{f}$, i.e., $\mathbf{g} \circ \mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$.

TEOREMA 4.1.2. *Sejam $\Omega \subseteq \mathbb{R}^m$, $R \subseteq \mathbb{R}^n$, e $\mathbf{f} : \Omega \rightarrow R$ e $\mathbf{g} : R \rightarrow \mathbb{R}^l$. Suponha \mathbf{f} contínua em $\mathbf{x} \in \Omega$ e \mathbf{g} contínua em $\mathbf{f}(\mathbf{x}) \in R$. Então a composição $\mathbf{g} \circ \mathbf{f} : \Omega \rightarrow \mathbb{R}^l$ é contínua em \mathbf{x} .*

DEMONSTRAÇÃO. Seja $\mathbf{y} = \mathbf{f}(\mathbf{x})$ e W vizinhança aberta de $\mathbf{g}(\mathbf{y})$. Como \mathbf{g} é contínua em \mathbf{y} , então existe vizinhança aberta V de \mathbf{y} tal que

$$(4.1.1) \quad \mathbf{y}' \in V \cap R \implies \mathbf{g}(\mathbf{y}') \in W.$$

Como \mathbf{f} é contínua em \mathbf{x} , então existe vizinhança aberta U de \mathbf{x} tal que

$$\mathbf{x}' \in U \cap \Omega \implies \mathbf{f}(\mathbf{x}') \in V.$$

Logo

$$\mathbf{x}' \in U \cap \Omega \implies \mathbf{f}(\mathbf{x}') \in V \implies \mathbf{f}(\mathbf{x}') \in V \cap R \implies \mathbf{g}(\mathbf{f}(\mathbf{x}')) \in W,$$

pois $\mathbf{f}(\mathbf{x}') \in R$, já que está imagem de \mathbf{f} . Na última implicação usamos ainda (4.1.1). Logo $\mathbf{g} \circ \mathbf{f}$ é contínua em \mathbf{x} . \square

EXEMPLO 4.8. A função $g(\mathbf{x}) = \|\mathbf{x}\|$ é contínua em \mathbb{R}^m . Realmente, como

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

se (\mathbf{x}_n) converge para \mathbf{x} então

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

Portanto, se f é contínua em \mathbf{x} , então $h(\mathbf{x}) = \|f(\mathbf{x})\|$ também o é, pois $h = g \circ f$ é composição de funções contínuas.

OBSERVAÇÃO. Note que não podemos concluir a continuidade de $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$, mesmo que $\|\mathbf{f}\|$ seja contínua. Por exemplo se $f : \mathbb{R} \rightarrow \mathbb{R}$ é tal que

$$f(x) = \begin{cases} -1 & \text{se } x \geq 0, \\ 1 & \text{se } x < 0, \end{cases}$$

então $|f|$ é contínua mesmo sendo f descontínua.

4.2. Propriedades globais

Algumas propriedades de funções contínuas não estão restritas a apenas um ponto, mas sim a todo o domínio. Como exemplos citamos preservação de compacidade, e a continuidade uniforme.

Antes de prosseguirmos com as propriedades e suas aplicações, temos o seguinte resultado que caracteriza funções contínuas em todo domínio.

TEOREMA 4.2.1 (Continuidade Global). *Seja $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$. Então as afirmativas abaixo são equivalentes:*

- (1) \mathbf{f} é contínua em Ω
- (2) Se $V \subseteq \mathbb{R}^n$ é aberto, então existe aberto U tal que $U \cap \Omega = \mathbf{f}^{-1}(V)$
- (3) Se $H \subseteq \mathbb{R}^n$ é fechado, então existe fechado F tal que $F \cap \Omega = \mathbf{f}^{-1}(H)$

DEMONSTRAÇÃO. (1) \implies (2): Seja \mathbf{f} contínua em Ω e $V \subseteq \mathbb{R}^n$ aberto. Seja $\mathbf{x} \in \mathbf{f}^{-1}(V)$. Como é \mathbf{f} contínua, existe aberto $U_{\mathbf{x}}$ contendo \mathbf{x} tal que

$$\mathbf{y} \in U_{\mathbf{x}} \cap \Omega \implies \mathbf{f}(\mathbf{y}) \in V.$$

Logo $U_{\mathbf{x}} \cap \Omega \subseteq \mathbf{f}^{-1}(V)$. Seja $U = \cup_{\mathbf{x} \in \mathbf{f}^{-1}(V)} U_{\mathbf{x}}$. Então U é aberto pois é união de abertos, e $U \cap \Omega = \mathbf{f}^{-1}(V)$.

(2) \implies (1): Seja $\mathbf{x} \in \Omega$ e V vizinhança aberta de $\mathbf{f}(\mathbf{x})$. Por hipótese existe um aberto U tal que $U \cap \Omega = \mathbf{f}^{-1}(V)$. Mas como $\mathbf{f}(\mathbf{x}) \in V$, então $\mathbf{x} \in U$ e portanto U é vizinhança aberta de \mathbf{x} . Além disto, para todo $\mathbf{y} \in U \cap \Omega$ tem-se $\mathbf{f}(\mathbf{y}) \in V$.

(2) \implies (3): Seja $H \subseteq \mathbb{R}^n$ fechado. Então como $\mathcal{C}(H)$ é aberto, temos por hipótese que existe aberto U tal que $U \cap \Omega = \mathbf{f}^{-1}(\mathcal{C}(H))$. Seja $F = \mathcal{C}(U)$. Então

$$\mathbf{x} \in F \cap \Omega \implies \mathbf{f}(\mathbf{x}) \notin \mathcal{C}(H) \implies \mathbf{f}(\mathbf{x}) \in H \implies F \cap \Omega \subseteq \mathbf{f}^{-1}(H).$$

Por outro lado,

$$\mathbf{x} \in \mathbf{f}^{-1}(H) \implies \mathbf{x} \notin U \cap \Omega \text{ e } \mathbf{x} \in \Omega \implies \mathbf{x} \in F \cap \Omega \implies \mathbf{f}^{-1}(H) \subseteq F \cap \Omega.$$

Logo $\mathbf{f}^{-1}(H) = F \cap \Omega$.

(3) \implies (2): semelhante ao caso anterior. □

OBSERVAÇÃO. Observe que pelas definições de abertos e fechados relativos dadas nos Exercícios 2.16, 2.17, o Teorema da Continuidade Global (Teorema 4.2.1) diz simplesmente que uma função $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ é contínua em Ω se e somente se *imagens inversas de abertos são abertos em Ω* e se e somente se *imagens inversas de fechados são fechados em Ω* .

OBSERVAÇÃO. Note que U aberto e \mathbf{f} contínua *não* implica em $\mathbf{f}(U)$ aberto. Da mesma forma, F fechado *não* implica em $\mathbf{f}(F)$ fechado. Como exemplo tome $f(x) = x^2$ e $U = (-1, 1)$ implica em $f(U) = [0, 1)$. E se $F = [1, +\infty)$, que é fechado, com $g(x) = 1/x$, então $g(F) = (0, 1]$.

4.2.1. Funções Contínuas em Conjuntos Compactos. Um resultado com várias aplicações vem a seguir e garante que a compacidade é uma propriedade preservada por funções contínuas.

TEOREMA 4.2.2 (Preservação de compacidade). *Se K é compacto, e $\mathbf{f} : K \rightarrow \mathbb{R}^n$ é contínua, então $\mathbf{f}(K)$ é compacto.*

DEMONSTRAÇÃO. Seja $\mathcal{G} = \{G_\alpha\}$ cobertura aberta para $\mathbf{f}(K)$, i.e., $\mathbf{f}(K) \subseteq \cup_\alpha G_\alpha$. Logo $K \subseteq \cup_\alpha \mathbf{f}^{-1}(G_\alpha)$. Por \mathbf{f} ser contínua, pelo Teorema 4.2.1, para todo α existe H_α aberto tal que $\mathbf{f}^{-1}(G_\alpha) = H_\alpha \cap K$. Portanto $\{H_\alpha\}$ é uma cobertura aberta de K . Como K é compacto, então existe $\{H_{\alpha_1}, \dots, H_{\alpha_J}\}$ subcobertura finita. Logo,

$$K \subseteq \cup_{j=1}^J H_{\alpha_j} \cap K = \cup_{j=1}^J \mathbf{f}^{-1}(G_{\alpha_j}),$$

e então $\mathbf{f}(K) \subseteq \cup_{j=1}^J G_{\alpha_j}$. Portanto, achamos uma subcobertura aberta finita para $\mathbf{f}(K)$, e concluímos que $\mathbf{f}(K)$ é compacto. \square

Uma aplicação imediata do resultado acima é a existência de máximos e mínimos de funções contínuas definidas em compactos. Em particular, estas funções são limitadas.

DEFINIÇÃO 4.2.3. Dizemos que $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ é limitada em Ω se existe $M \in \mathbb{R}$ tal que $\|\mathbf{f}(\mathbf{x})\| \leq M$ para todo $\mathbf{x} \in \Omega$.

EXEMPLO 4.9. $\sin x$ é limitada em \mathbb{R} pois $|\sin x| \leq 1$ para todo $x \in \mathbb{R}$.

EXEMPLO 4.10. A função $1/x$ 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.

O Teorema 4.2.2 garante que imagens de compactos são conjuntos compactos, portanto pelo Teorema de Heine–Borel (Teorema 2.5.3) fechados e limitados. O resultado abaixo é consequência imediata deste fato.

TEOREMA 4.2.4. *Seja K compacto, e $\mathbf{f} : K \rightarrow \mathbb{R}^n$ contínua em K . Então \mathbf{f} é limitada em K .*

Uma demonstração alternativa do Teorema 4.2.4 que dispensa o uso de noções de compacidade vem a seguir.

DEMONSTRAÇÃO. (alternativa do Teorema 4.2.4; *por contradição*) Suponha K fechado e limitado e \mathbf{f} não limitada. Então para todo $n \in \mathbb{N}$ existe $\mathbf{x}_n \in K$ tal que $\mathbf{f}(\mathbf{x}_n) > n$. Como K é fechado e limitado, então, pelo Teorema de Bolzano–Weierstrass, (\mathbf{x}_n) possui subsequência (\mathbf{x}_{n_k}) convergente. Seja $\mathbf{x} = \lim_{n_k \rightarrow \infty} \mathbf{x}_{n_k}$. Como K é fechado, então $\mathbf{x} \in K$. Mas como \mathbf{f} é contínua, então \mathbf{f} tem limite em \mathbf{x} , e portanto é localmente limitada, uma contradição com a construção de (\mathbf{x}_n) . \square

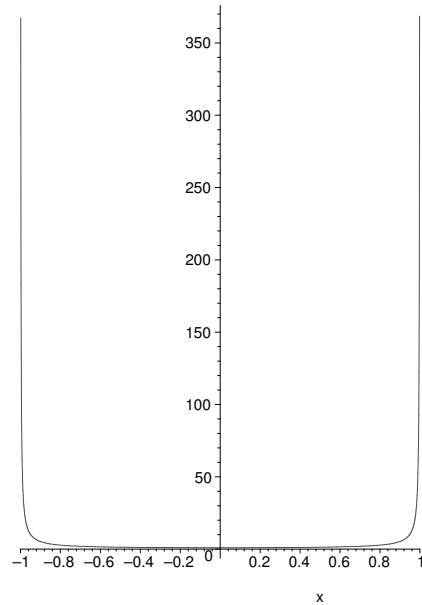


FIG. 4. Gráfico de $1/(1-x^2)$, que não é limitada em $(-1, 1)$.

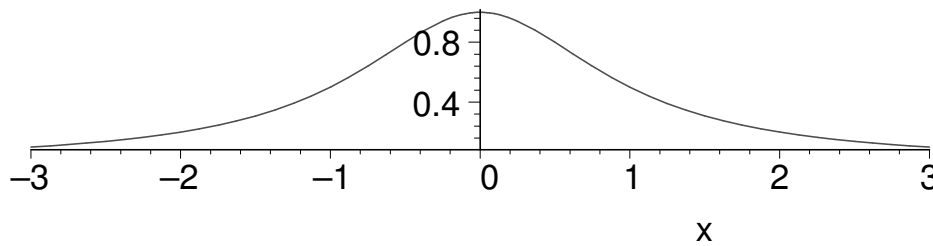


FIG. 5. Gráfico de $1/(1+x^2)$, que tem seu máximo mas não o seu mínimo em \mathbb{R} .

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

OBSERVAÇÃO. Se uma função f como acima definida tem seus valores máximo e mínimo em Ω , então f é limitada em Ω .

EXEMPLO 4.11. A função $f : (-1, 1) \rightarrow \mathbb{R}$ dada por $f(x) = 1/(1-x^2)$ (Figura 4) não é limitada em $(-1, 1)$, mas é limitada em $[-1/2, 1/2]$ por exemplo.

EXEMPLO 4.12. $f(x) = x$ é contínua e limitada em $(-1, 1)$, mas não toma valor máximo nem mínimo em $(-1, 1)$. Entretanto f tem seus valores máximo e mínimo em $[-1, 1]$.

EXEMPLO 4.13. A função $h(x) = 1/(1+x^2)$ (Figura 5) é limitada em \mathbb{R} , tem seu valor máximo em $x^* = 0$, mas não tem 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 $[-1, 1]$.

O resultado a seguir mais uma vez é consequência do Teorema 4.2.2.

TEOREMA 4.2.5 (Pontos Extremos). *Seja K compacto e $f : K \rightarrow \mathbb{R}$ contínua em K . Então f tem pelo menos um ponto de máximo e um de mínimo em K .*

DEMONSTRAÇÃO. Como K é compacto, então o Teorema 4.2.2 garante que $f(K)$ também é compacto. Logo $f(K)$ é limitado e portanto tem supremo, e $f(K)$ é fechado, e portanto o supremo pertence a $f(K)$. Logo existe $x^* \in K$ tal que $f(x^*) = \sup f(K)$.

Mesmo tipo de argumento assegura que existe ponto de mínimo em K . \square

A seguinte demonstração dispensa o uso direto de compacidade.

DEMONSTRAÇÃO. (alternativa do Teorema 4.2.5) Demonstraremos somente que existe um ponto de máximo para f . O caso de valor mínimo é análogo. Como K é fechado limitado, então $f(K)$ é limitado. Seja $s^* = \sup f(K)$. Seja \mathbf{x}_n tal que $f(\mathbf{x}_n) > s^* - 1/n$. Mas pelo Teorema de Bolzano–Weierstrass, K limitado implica em existência de uma subsequência (\mathbf{x}_{n_k}) convergente. Seja \mathbf{x}^* o limite de tal subsequência. Como K é fechado, então $\mathbf{x}^* \in K$. Como f é contínua, então $f(\mathbf{x}^*) = \lim_{n_k \rightarrow \infty} f(\mathbf{x}_{n_k})$. Finalmente, usamos que

$$s^* - \frac{1}{n_k} \leq f(\mathbf{x}_{n_k}) \leq s^*,$$

e pelo Lema do sanduíche de seqüências 3.6.3, temos que $f(\mathbf{x}^*) = \lim_{n_k \rightarrow \infty} f(\mathbf{x}_{n_k}) = s^*$. \square

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

TEOREMA 4.2.6 (Teorema do Valor Intermediário). *Sejam $a < b$ e suponha $f : [a, b] \rightarrow \mathbb{R}$ contínua. Se existe $d \in \mathbb{R}$ tal que $f(a) < d < f(b)$, então existe $c \in (a, b)$ tal que $f(c) = d$.*

DEMONSTRAÇÃO. Seja

$$I = \{x \in [a, b] : f(x) < d\} = f^{-1}((-\infty, d)).$$

Logo I é não vazio pois $a \in I$, e definimos $c = \sup I$. Então $c \in [a, b]$, pois b é cota superior de I . Pelo Teorema da Continuidade Global (Teorema 4.2.1), existe aberto U tal que

$$I = U \cap [a, b] = U \cap [a, b),$$

pois $b \notin I$. Logo, para todo $x \in I$, existe $\epsilon > 0$ tal que $x + \epsilon \in I$. Portanto $c \notin I$, i.e., $f(c) \geq d$. Seja então $x_n \in I$ tal que $x_n \rightarrow c$, ver Exercício 3.14. Por continuidade de f , temos $f(c) = \lim_{n \rightarrow \infty} f(x_n)$. Como $f(x_n) < d$, então $f(c) \leq d$. Portanto $f(c) = d$. \square

COROLÁRIO 4.2.7 (Teorema do ponto fixo em uma dimensão). *Seja $f : [0, 1] \rightarrow [0, 1]$ contínua. Então f tem um ponto fixo, i.e., existe $x \in [0, 1]$ tal que $f(x) = x$.*

DEMONSTRAÇÃO. seja $d : [0, 1] \rightarrow \mathbb{R}$ dada por $d(x) = f(x) - x$. Portanto d é contínua. Nosso objetivo é achar raiz para d em $[0, 1]$. Se $d(0) = 0$ ou $d(1) = 0$, então nada mais há a fazer. Suponha que nem 0 nem 1 sejam raízes de d . Logo $d(0) = f(0) > 0$ e $d(1) = f(1) - 1 < 0$ pois $f(x) \in [0, 1]$. Aplicando o Teorema do Valor Intermediário (Teorema 4.2.6), temos que existe $x \in (0, 1)$ tal que $d(x) = 0$, como queríamos demonstrar. \square

Concluimos esta parte com uma importante consequência dos resultados anteriores.

TEOREMA 4.2.8. *Seja I intervalo fechado limitado e $f : I \rightarrow \mathbb{R}$ função contínua. Então $f(I)$ é intervalo fechado limitado.*

4.3. Funções Uniformemente Contínuas

Considere $g(x) = 1/x$, para $x \in (0, 1)$. Seja $c \in (0, 1)$. Então

$$g(c) - g(x) = \frac{1}{c} - \frac{1}{x} = \frac{x - c}{cx}.$$

Para mostrarmos que g é contínua em c , seja $\epsilon > 0$. Sem perda de generalidade, podemos supor que $\epsilon < 1$, e portanto $\epsilon c < 1$. Seja $\delta = c^2\epsilon/2$. Então

$$|x - c| < \delta \implies c < x + \delta = x + \frac{c^2\epsilon}{2} < x + \frac{c}{2} \implies \frac{c}{2} < x.$$

Logo

$$|x - c| < \delta \implies |g(c) - g(x)| = \frac{|x - c|}{cx} < \frac{\delta}{cx} = \frac{c^2\epsilon}{2cx} = \frac{c\epsilon}{2x} < \epsilon$$

onde usamos que $c/2 < x$ na última desigualdade. Mostramos então, usando ϵ s e δ s que $1/x$ é contínua em todo ponto diferente de zero. O objetivo principal do cálculo acima é ressaltar que a escolha de δ não é uniforme em relação ao ponto c , i.e., δ depende de c .

Em outros casos, a escolha de δ independe do ponto em questão. Por exemplo, para $f(x) = x$, dado $\epsilon > 0$, tomando $\delta = \epsilon$ temos

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

Outro caso já foi visto no exemplo 4.2. Dizemos que estas funções são uniformemente contínuas.

DEFINIÇÃO 4.3.1. *Seja $\Omega \subseteq \mathbb{R}^m$ e $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$. Dizemos que \mathbf{f} é uniformemente contínua em Ω se para todo $\epsilon > 0$, existir δ tal que para todo $\mathbf{x}, \mathbf{y} \in \Omega$ tem-se*

$$\|\mathbf{x} - \mathbf{y}\| < \delta \implies \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| < \epsilon.$$

Note que a definição de continuidade uniforme só faz sentido no domínio ou subdomínio da função, e não pontualmente como na definição de continuidade. Uma forma equivalente de se definir uma função uniformemente contínua, é exigir que dado $\epsilon > 0$ exista δ tal que para todo $\mathbf{x} \in \Omega$ tem-se

$$\mathbf{y} \in B_\delta(\mathbf{x}) \cap \Omega \implies \mathbf{f}(\mathbf{y}) \in B_\epsilon(\mathbf{f}(\mathbf{x})).$$

Além disto, pode-se usar o seguinte resultado abaixo para se mostrar que uma função *não* é uniformemente contínua.

LEMA 4.3.2. *Seja $\Omega \subseteq \mathbb{R}^m$ e $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$. Então as afirmativas abaixo são equivalentes.*

- (1) \mathbf{f} não é uniformemente contínua em Ω .
- (2) Existe $\epsilon > 0$ tal que para todo $\delta > 0$ existem pontos $\mathbf{x}, \mathbf{y} \in \Omega$ tais que $\|\mathbf{x} - \mathbf{y}\| < \delta$ mas $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| > \epsilon$.
- (3) Existe $\epsilon > 0$ e duas seqüências (\mathbf{x}_k) e (\mathbf{y}_k) em Ω tais que $\lim_{k \rightarrow \infty} (\mathbf{x}_k - \mathbf{y}_k) = 0$ e $\|\mathbf{f}(\mathbf{x}_k) - \mathbf{f}(\mathbf{y}_k)\| > \epsilon$ para todo $k \in \mathbb{N}$.

EXEMPLO 4.14. O resultado acima pode ser usado por exemplo para mostrar que $f(x) = 1/x$ não é uniformemente contínua em \mathbb{R}^+ . Considere as seqüências $(1/k)$ e $(1/(k+1))$. Então $\lim_{k \rightarrow \infty} (1/k - 1/(k+1)) = 0$ mas $f(1/k) - f(1/(k+1)) = 1$ para todo $k \in \mathbb{N}$.

Uma interessante propriedade da continuidade uniforme é dada abaixo, e tem aplicação na extensão de funções, ver exercício 4.16. Seja $\Omega \subseteq \mathbb{R}^m$ e suponha que $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ é uniformemente contínua. Então (\mathbf{x}_i) ser sequência de Cauchy implica que $(\mathbf{f}(\mathbf{x}_i))$ também é sequência de Cauchy.

De fato, seja $\epsilon > 0$. Como \mathbf{f} é uniformemente contínua, então existe δ tal que

$$(4.3.1) \quad \|\mathbf{x} - \mathbf{y}\| < \delta \implies \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| < \epsilon,$$

para todo $\mathbf{x}, \mathbf{y} \in \Omega$. Como (\mathbf{x}_i) é sequência de Cauchy, então existe N_0 tal que se

$$(4.3.2) \quad i, j > N_0 \implies \|\mathbf{x}_i - \mathbf{x}_j\| < \delta.$$

Combinando (4.3.1) e (4.3.2), temos então que

$$i, j > N_0 \implies \|\mathbf{f}(\mathbf{x}_i) - \mathbf{f}(\mathbf{x}_j)\| < \epsilon.$$

Note que isto nos dá um outro critério para determinar quando uma função não é uniformemente contínua. Por exemplo, para o caso considerado no exemplo 4.14, temos que $x_k = 1/k$ é de Cauchy mas $f(x_k) = k$ não é de Cauchy. Logo f não é uniformemente contínua em \mathbb{R}^+ .

Apesar de parecer difícil conferir se uma dada função é ou não uniformemente contínua, o (supreendente?) resultado abaixo garante que *todas* as funções contínuas em conjuntos compactos são uniformemente contínuas.

TEOREMA 4.3.3 (Continuidade Uniforme em compactos). *Seja $K \subseteq \mathbb{R}^m$ conjunto compacto, e $\mathbf{f} : K \rightarrow \mathbb{R}^n$ contínua em K . Então \mathbf{f} é uniformemente contínua em K .*

DEMONSTRAÇÃO. Seja $\epsilon > 0$. Entao, para todo $\mathbf{x} \in K$, existe $\delta(\mathbf{x}) > 0$ tal que

$$(4.3.3) \quad \mathbf{y} \in B_{\delta(\mathbf{x})}(\mathbf{x}) \cap K \implies \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| < \epsilon/2.$$

Seja a cobertura aberta de K gerada por $\{B_{\frac{1}{2}\delta(\mathbf{x})}(\mathbf{x})\}_{\mathbf{x} \in K}$. Como K é compacto, então existe $\{\mathbf{x}_1, \dots, \mathbf{x}_J\}$ tal que $\{B_{\frac{1}{2}\delta(\mathbf{x}_i)}(\mathbf{x}_i)\}_{i=1}^J$ é uma subcobertura de K . Seja

$$\delta = \frac{1}{2} \min\{\delta(\mathbf{x}_1), \dots, \delta(\mathbf{x}_J)\}.$$

Sejam $\mathbf{x}, \mathbf{y} \in K$ tais que $\|\mathbf{x} - \mathbf{y}\| < \delta$. Então existe índice $j \in \{1, \dots, J\}$ tal que $\mathbf{x} \in B_{\frac{1}{2}\delta(\mathbf{x}_j)}(\mathbf{x}_j)$, i.e., $\|\mathbf{x} - \mathbf{x}_j\| < \delta(\mathbf{x}_j)/2$. Portanto, usando (4.3.3) temos que $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_j)\| < \epsilon/2$. Da mesma forma,

$$\|\mathbf{y} - \mathbf{x}_j\| \leq \|\mathbf{y} - \mathbf{x}\| + \|\mathbf{x} - \mathbf{x}_j\| < \delta + \frac{1}{2}\delta(\mathbf{x}_j) \leq \delta(\mathbf{x}_j),$$

e então $\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}_j)\| < \epsilon/2$. Concluimos que

$$\|\mathbf{x} - \mathbf{y}\| < \delta \implies \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_j)\| + \|\mathbf{f}(\mathbf{x}_j) - \mathbf{f}(\mathbf{y})\| < \epsilon,$$

e portanto \mathbf{f} é uniformemente contínua. \square

Abaixo apresentamos uma demonstração alternativa do Teorema 4.3.3, que não usa argumentos de compacidade.

DEMONSTRAÇÃO. (alternativa do Teorema 4.3.3; *por contradição*) Suponha que f não seja uniformemente contínua. Como K é compacto, então é fechado e limitado. Então, pelo Lema 4.3.2, existe $\epsilon > 0$ e existem sequências (\mathbf{x}_n) e (\mathbf{y}_n) em K tais que $\|\mathbf{x}_n - \mathbf{y}_n\| < 1/n$ e $\|\mathbf{f}(\mathbf{x}_n) - \mathbf{f}(\mathbf{y}_n)\| > \epsilon$. Como K é fechado, pelo Teorema de Bolzano–Weierstrass, existe

subsequência (\mathbf{x}_{n_k}) convergente. Seja $\mathbf{z} = \lim_{n_k \rightarrow \infty} (\mathbf{x}_{n_k})$. Como K é fechado, então $\mathbf{z} \in K$. Note que (\mathbf{y}_{n_k}) também converge para \mathbf{z} pois

$$(\mathbf{y}_{n_k} - \mathbf{z}) = (\mathbf{y}_{n_k} - \mathbf{x}_{n_k}) + (\mathbf{x}_{n_k} - \mathbf{z}).$$

Como \mathbf{f} é contínua em \mathbf{z} , então $\mathbf{f}(\mathbf{z}) = \lim_{n_k \rightarrow \infty} \mathbf{f}(\mathbf{x}_{n_k})$, e $\mathbf{f}(\mathbf{z}) = \lim_{n_k \rightarrow \infty} \mathbf{f}(\mathbf{y}_{n_k})$, uma contradição com $\|\mathbf{f}(\mathbf{x}_n) - \mathbf{f}(\mathbf{y}_n)\| > \epsilon$. Logo \mathbf{f} é uniformemente contínua. \square

Outra importante situação em que temos continuidade uniforme, mesmo com domínios não compactos, é quando a função é de Lipschitz. Seja $\Omega \subseteq \mathbb{R}^m$ e $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$. Dizemos que \mathbf{f} é de Lipschitz se existe $M \in \mathbb{R}$ tal que

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq M\|\mathbf{x} - \mathbf{y}\|$$

para todo $\mathbf{x}, \mathbf{y} \in \Omega$.

TEOREMA 4.3.4. *Se $\Omega \subseteq \mathbb{R}^m$ e $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$, e \mathbf{f} é de Lipschitz, então \mathbf{f} é uniformemente contínua em Ω .*

DEMONSTRAÇÃO. Seja $M \in \mathbb{R}$ tal que

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq M\|\mathbf{x} - \mathbf{y}\|$$

para todo $\mathbf{x}, \mathbf{y} \in \Omega$. Dado $\epsilon > 0$, seja $\delta = \epsilon/M$. Então se $\mathbf{x}, \mathbf{y} \in \Omega$ e $\|\mathbf{x} - \mathbf{y}\| < \delta$, temos que

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq M\|\mathbf{x} - \mathbf{y}\| \leq M\delta = \epsilon.$$

o que mostra que \mathbf{f} é uniformemente contínua em Ω . \square

Nem toda função uniformemente contínua é de Lipschitz, como o exemplo abaixo mostra.

EXEMPLO 4.15. Seja $g : [0, 1] \rightarrow \mathbb{R}$, tal que $g(x) = \sqrt{x}$. Como $[0, 1]$ é compacto, e g é contínua, então g é uniformemente contínua em $[0, 1]$. Entretanto note que se g fosse de Lipschitz, nós teríamos a existência de $M \in \mathbb{R}$ tal que

$$\sqrt{x} = |g(x) - g(0)| \leq k|x - 0| = Mx \implies \frac{1}{\sqrt{x}} \leq M \quad \text{para todo } x > 0,$$

um absurdo. Logo g não é de Lipschitz apesar de ser uniformemente contínua em seu domínio.

4.4. Exercícios

EXERCÍCIO 4.1. 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, $[2] = 2$, $[2.5] = 2$, $[-2.5] = -3$.

EXERCÍCIO 4.2. Demonstre o Lema 4.1.1.

EXERCÍCIO 4.3. Seja $f : \mathbb{R}^m \rightarrow \mathbb{R}$ contínua em $\mathbf{x} \in \mathbb{R}^m$, e $f(\mathbf{x}) > 0$. Mostre que existe uma vizinhança aberta de \mathbf{x} tal que f seja estritamente positiva.

EXERCÍCIO 4.4. Seja $f : \mathbb{R}^m \rightarrow \mathbb{R}$ contínua. Mostre que o conjunto $\{\mathbf{x} \in \mathbb{R}^m : f(\mathbf{x}) = 0\}$ é fechado em \mathbb{R}^m .

EXERCÍCIO 4.5. Sejam $f, g : \mathbb{R}^m \rightarrow \mathbb{R}$ funções contínuas. Mostre que o conjunto $\{\mathbf{x} \in \mathbb{R}^m : f(\mathbf{x}) > g(\mathbf{x})\}$ é aberto em \mathbb{R}^m .

EXERCÍCIO 4.6. Mostre que toda contração é uma função contínua.

EXERCÍCIO 4.7. Dê exemplos de

- (1) Um conjunto F fechado em \mathbb{R} e uma função $f : F \rightarrow \mathbb{R}$ contínua tais que $f(F)$ não seja compacto.
- (2) Um conjunto A aberto em \mathbb{R} e uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ tais que $f^{-1}(A)$ não seja aberto em \mathbb{R} .
- (3) Um conjunto $\Omega \subseteq \mathbb{R}$, um conjunto A aberto em \mathbb{R} e uma função contínua $f : \Omega \rightarrow \mathbb{R}$ tais que $f^{-1}(A)$ não seja aberto em \mathbb{R} .

EXERCÍCIO 4.8. Seja $f : \Omega \rightarrow \mathbb{R}$, onde $\Omega \subset \mathbb{R}^m$ é aberto. Mostre que f é contínua em Ω se e somente se $f^{-1}((\alpha, +\infty))$ e $f^{-1}((-\infty, \alpha))$ são abertos para todo $\alpha \in \mathbb{R}$.

EXERCÍCIO 4.9. Seja $f : [0, 1] \rightarrow \mathbb{R}$ contínua tal que $f(0) < 0$ e $f(1) > 0$. Mostre que se $s = \sup\{x \in [0, 1] : f(x) < 0\}$, então $f(s) = 0$.

EXERCÍCIO 4.10. Seja $\Omega \subseteq \mathbb{R}^m$ conjunto limitado. Dê exemplo de uma função $f : \Omega \rightarrow \mathbb{R}$ contínua e limitada que não atinja seu máximo.

EXERCÍCIO 4.11 (Equivalência de normas no \mathbb{R}^n). Dizemos que duas normas $\|\cdot\|$ e $\|\|\cdot\|\|$ de uma espaço vetorial V são *equivalentes* se existem constantes c_1 e c_2 tais que

$$c_1\|\|\mathbf{v}\|\| \leq \|\mathbf{v}\| \leq c_2\|\|\mathbf{v}\|\|$$

para todo $v \in V$. Mostre que no \mathbb{R}^n todas as normas são equivalentes.

(Dica: mostre que todas as normas são equivalentes à norma euclidiana, i.e., considere $\|\cdot\|$ como sendo a norma euclidiana. Para tal, comece mostrando que existe constante c_1 tal que $c_1\|\|\mathbf{v}\|\| \leq \|\mathbf{v}\|$ para todo $v \in V$. Para obter a desigualdade inversa, mostre que $\|\|\cdot\|\|$ define uma função contínua em \mathbb{R}^n . Conclua então usando o Teorema dos pontos extremos (Teorema 4.2.5) de forma apropriada.)

EXERCÍCIO 4.12. Mostre que não é possível generalizar o Teorema do ponto fixo (Teorema 4.2.7) para o intervalo $(0, 1]$.

EXERCÍCIO 4.13. Mostre que se $f : \mathbb{R}^m \rightarrow \mathbb{R}$ e $g : \mathbb{R}^m \rightarrow \mathbb{R}$ são uniformemente contínuas, então $f + g$ é uniformemente contínua. Mostre que, mesmo que f seja limitada, a função fg não é necessariamente uniformemente contínua.

EXERCÍCIO 4.14. Mostre que o produto de duas funções uniformemente contínuas e limitadas é função uniformemente contínua.

EXERCÍCIO 4.15. Sejam $a < b$ números reais, e $f : [a, b] \rightarrow \mathbb{R}$ contínua. Mostre que dado $\epsilon > 0$, existem $a = a_0 < a_1 < \dots < a_n = b$ tais que se $x, y \in [a_{i-1}, a_i]$ para algum $i \in \{1, \dots, n\}$, então $|f(x) - f(y)| < \epsilon$.

EXERCÍCIO 4.16. Suponha $f : (0, 1] \rightarrow \mathbb{R}$ uniformemente contínua em $(0, 1]$. Mostre que podemos definir $f(0)$ tal que f seja uniformemente contínua em $[0, 1]$.

EXERCÍCIO 4.17. Suponha $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ uniformemente contínua em Ω . Mostre que podemos definir $\bar{\mathbf{f}} : \bar{\Omega} \rightarrow \mathbb{R}^n$ tal que $\bar{\mathbf{f}}$ seja contínua em $\bar{\Omega}$, e $\bar{\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ para todo $\mathbf{x} \in \Omega$. Neste caso dizemos que $\bar{\mathbf{f}}$ é uma *extensão contínua* de \mathbf{f} .

EXERCÍCIO 4.18. Seja $B \subseteq \mathbb{R}^m$ limitado, e $\mathbf{f} : B \rightarrow \mathbb{R}^n$ uniformemente contínua. Mostre que \mathbf{f} é limitada em B . Mostre que esta conclusão não é necessariamente verdadeira se B não for limitado.

CAPÍTULO 5

Diferenciação

¹ Neste capítulo vemos a noção de diferenciabilidade e suas aplicações. Começaremos com o caso unidimensional, onde veremos algumas propriedades e aplicações particulares.

5.1. Derivada em uma dimensão

Seja $f : I \rightarrow \mathbb{R}$, onde I é um intervalo. 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 $L = f'(c)$.

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} .

Existem outras formas de se definir a diferenciabilidade. De fato as afirmativas abaixo são equivalentes:

- (1) $f : I \rightarrow \mathbb{R}$ é diferenciável em $c \in I$, com derivada L .
- (2) O limite abaixo existe e é igual a L :

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

- (3) Para toda sequência (x_k) em $I \setminus c$ convergindo para c tem-se

$$\lim_{k \rightarrow \infty} \frac{f(x_k) - f(c)}{x_k - c} = L$$

- (4) Existe uma função r tal que

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

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

$$(5.1.1) \quad f(c + h) = f(c) + Lh + r(h) \quad \text{com} \quad \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0.$$

Podemos também entender L como a aplicação linear (neste caso dada por um número) que torna (5.1.1) possível. Esta interpretação induz de forma natural a generalização da noção de derivada para o caso multidimensional.

OBSERVAÇÃO. Seja $f : D \rightarrow \mathbb{R}$ e $c \in D$, onde $D \subseteq \mathbb{R}$. Mesmo que D não seja um intervalo, é possível definir a derivada de f em c , desde que c seja de acumulação [6, 13].

A seguir temos dois exemplos de funções diferenciáveis.

¹Última Atualização: 25/04/2008

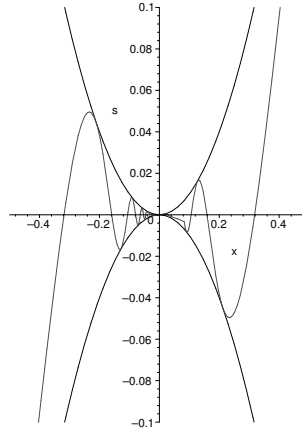


FIG. 1. Gráfico de $f(x)$, que é diferenciável, mas a derivada não é contínua.

EXEMPLO 5.1. Se $f(x) = x^2$, então para $c \in \mathbb{R}$ tem-se

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

EXEMPLO 5.2. Seja

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

mostrada na Figura 1. Para $x \neq 0$ temos $f'(x) = 2x \sin 1/x - \cos 1/x$. Observe que não existe o limite $\lim_{x \rightarrow 0} f'(x)$. 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.

Diferenciabilidade implica em continuidade, como nos mostra o resultado a seguir.

TEOREMA 5.1.1. *Se $f : I \rightarrow \mathbb{R}$, onde I é um intervalo, 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 $f(x) = |x|$,

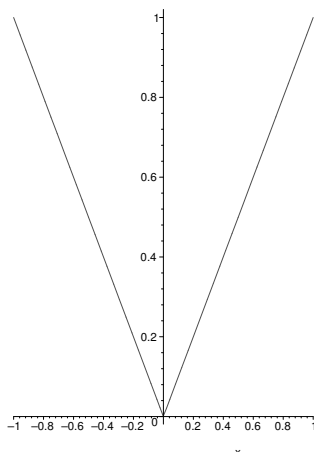


FIG. 2. Gráfico de $f(x) = |x|$, função contínua mas não diferenciável.

representada na Figura 2. Então f é contínua em \mathbb{R} mas não é diferenciável em zero pois para $x \neq 0$ temos

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

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

Sejam f e g funções de $I \rightarrow \mathbb{R}$, onde I é um intervalo, 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 5.3. Pela regra acima temos que se $f(x) = x^n$, para $n \in \mathbb{N}$, então f é diferenciável e $f'(c) = nx^{n-1}$.

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

$$(5.1.2) \quad y \in (x - \delta, x + \delta) \cap I \implies f(y) \leq f(x).$$

Se a desigualdade em (5.1.2) for estrita, chamamos o ponto de *máximo estrito local*. Definição análoga serve para mínimo local e mínimo estrito local. Chamamos um ponto de máximo ou mínimo (estrito) local de ponto extremo (estrito) local.

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

TEOREMA 5.1.2 (Ponto extremo interior). *Seja $f : I \rightarrow \mathbb{R}$, onde I é um intervalo, e c ponto interior de I e extremo local de f . Se f é diferenciável em c , então $f'(c) = 0$.*

DEMONSTRAÇÃO. Sem perda de generalidade, suponha que c seja ponto interior 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 aberta 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 $f'(c) = 0$. \square

Note que se a derivada de uma função se anula num determinado ponto, não se pode concluir que este seja um ponto extremo. Como exemplo temos $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = x^3$, que tem derivada zero em $x = 0$ mas este não é ponto de máximo nem mínimo local. Ver Figura 3.

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*, ilustrados na Figura 4.

TEOREMA 5.1.3 (Teorema de Rolle). *Seja $a < b \in \mathbb{R}$ e $f : [a, b] \rightarrow \mathbb{R}$ contínua em $[a, b]$ e diferenciável em (a, b) . Suponha ainda que $f(a) = f(b) = 0$. Então existe $c \in (a, b)$ tal que $f'(c) = 0$.*

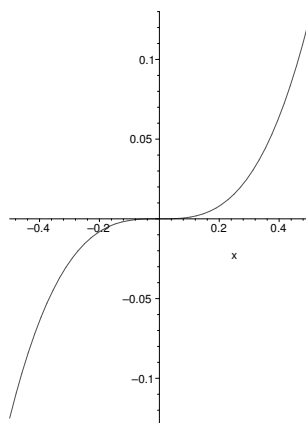


FIG. 3. Gráfico de $f(x) = x^3$, que tem derivada zero em $x = 0$, mas este não é ponto extremo.

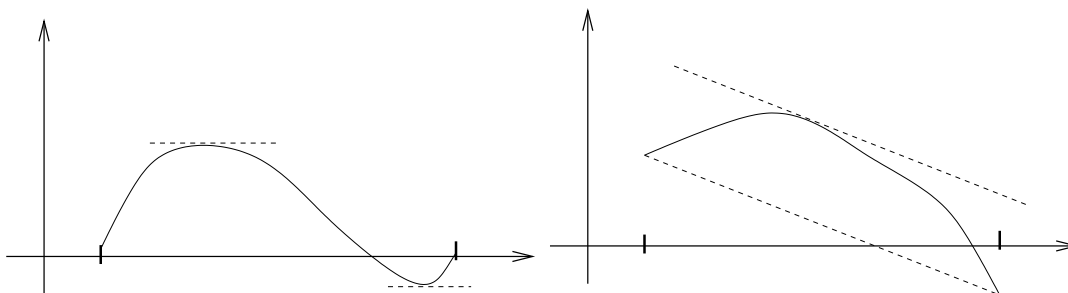


FIG. 4. Os Teoremas de Rolle e do Valor Médio versam sobre a existência de derivada com valor determinado pelos pontos extremos.

DEMONSTRAÇÃO. Se f é identicamente nula em $[a, b]$, então o resultado é verdadeiro. Caso contrário, então f tem algum valor positivo ou negativo em (a, b) . Sem perda de generalidade, suponha que f tem algum valor positivo. Como $[a, b]$ é compacto, então f atinge seu máximo em algum $c \in (a, b)$. Mas pelo Teorema do ponto extremo interior (Teorema 5.1.2), $f'(c) = 0$, como queríamos demonstrar. \square

TEOREMA 5.1.4 (Teorema do Valor Médio). *Seja $a < b \in \mathbb{R}$ e $f : [a, b] \rightarrow \mathbb{R}$ contínua em $[a, b]$ 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 ϕ também o é no mesmo intervalo. Logo, pelo Teorema de Rolle 5.1.3 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 5.1.5. Suponha 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 $x \in (a, b)$, então f é constante em $[a, b]$.

DEMONSTRAÇÃO. Seja $a < x \leq b$. Pelo Teorema do Valor Médio 5.1.4, 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 $[a, b]$. \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. Dizemos que uma função $f : I \rightarrow \mathbb{R}$ é *crescente* no intervalo I se para $x, y \in I$ com $y > x$ tem-se $f(y) \geq f(x)$. Dizemos ainda que $f : I \rightarrow \mathbb{R}$ é *estritamente crescente* em I se para $x, y \in I$ com $y > x$ tem-se $f(y) > f(x)$. Definições análogas valem para funções *decrecentes* e *estritamente decrecentes*.

LEMA 5.1.6. Seja I 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. (\Rightarrow) Suponha f crescente. Para $x, c \in I$,

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

Portanto

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

(\Leftarrow) Suponha $f'(x) \geq 0$ para todo $x \in I$. Sejam $x_1, x_2 \in I$ com $x_1 < x_2$. Usando o teorema do valor médio 5.1.4, existe $c \in (x_1, x_2)$ tal que $f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$. \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 $f(x) = x^3$ (Figura 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

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

é diferenciável em zero com $g'(0) = 1$, mas não é crescente em nenhuma vizinhança do zero, ver Figura 5. De fato, considere as sequências $x_k = 1/(2k + 1/2)$ e $y_k = 1/(2k)$. Então $x_k < y_k$ mas $g(x_k) > g(y_k)$. Para comprovar esta última desigualdade note que

$$\begin{aligned} 2k \left(2k + \frac{1}{2}\right)^2 [g(x_k) - g(y_k)] &= 2k \left(2k + \frac{1}{2}\right)^2 \left[\frac{1}{2k + \frac{1}{2}} + \frac{2}{(2k + \frac{1}{2})^2} - \frac{1}{2k} \right] \\ &= 2k \left(2k + \frac{1}{2}\right) + 4k - \left(2k + \frac{1}{2}\right)^2 = 4k^2 + k + 4k - (4k^2 + 2k + 1/4) = 3k - 1/4 > 0 \end{aligned}$$

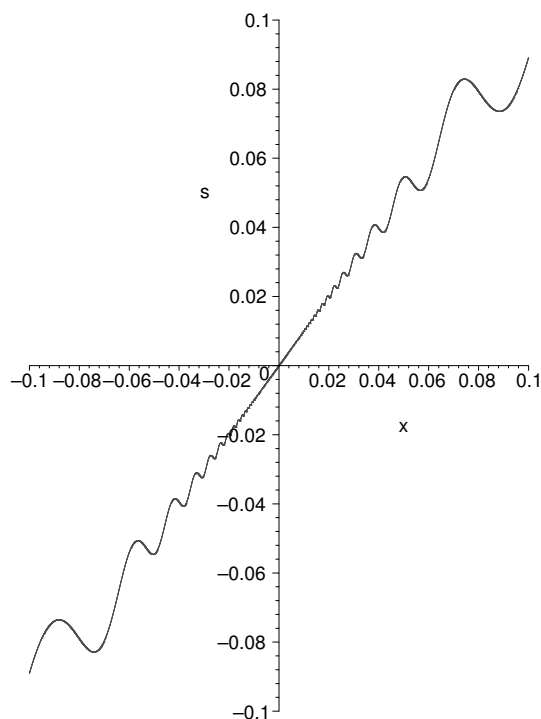


FIG. 5. Gráfico de $g(x)$, que tem $g'(0) = 1$ mas não é localmente crescente.

para todo k inteiro positivo.

Outras aplicações do Teorema do Valor Médio seguem nos exemplos abaixo.

EXEMPLO 5.4. Seja $f(x) = \exp(x)$. Então $f'(x) = \exp(x)$. Queremos mostrar que

$$(5.1.3) \quad \exp(x) > 1 + x \text{ para todo } x \neq 0.$$

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 (5.1.3) vale.

EXEMPLO 5.5 (Ponto Fixo). Seja I intervalo fechado e $f : I \rightarrow I$ diferenciável em I tal que $|f'(x)| < c$ para todo $x \in I$, onde $c < 1$. Então a sequência definida por x_0 e $x_i = f(x_{i-1})$ para $i \in \mathbb{N}$ converge, e $x^* = \lim_{n \rightarrow \infty} x_n$ é ponto fixo, i.e, $f(x^*) = x^*$. Além disto, este ponto fixo é único.

De fato, note que para f é uma contração pois

$$|f(y) - f(x)| \leq f'(\xi)|y - x| \leq c|y - x|,$$

onde ξ é um ponto entre y e x , e como I é intervalo, então $\xi \in I$. Logo pelo Teorema 3.5.3, o ponto fixo é único, e é o limite da sequência (x_i) acima, pois esta é a gerada pelo método das aproximações sucessivas, e portanto converge.

5.2. Teorema de Taylor em uma dimensão 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 por um polinômio uma função que pode ser complicada. Suponha que $f : I \rightarrow \mathbb{R}$ onde $I \subseteq \mathbb{R}$ tenha $k \geq 0$ derivadas num ponto $x_0 \in I$. Defina

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

onde escrevemos $g^{(j)}(c)$ para indicamos a j -ésima deriva de g num ponto c .

Note que com a definição acima, temos $f^{(j)}(x_0) = P_k^{(j)}(x_0)$ para $j = 1, \dots, k$. Chamamos P_k de polinômio de Taylor de ordem k 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 5.2.1 (Taylor). *Seja $k \geq 0$ e $I = [a, b]$, com $a < b$. Seja $f : I \rightarrow \mathbb{R}$ função k vezes diferenciável em I com $f^{(k)}$ contínua em I e tal $f^{(k+1)}$ exista em (a, b) . Se $x_0, x \in I$ então existe $\xi \in (x_0, x) \cup (x, x_0)$ tal que*

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

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

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

Logo

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

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

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

temos $G(x_0) = G(x) = 0$. Pelo Teorema de Rolle (Teorema 5.1.3) existe $\xi \in (x_0, x)$ tal que

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

Portanto

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

□

EXEMPLO 5.6. Seja $f : I \rightarrow \mathbb{R}$, onde $I = [a, b] \subseteq \mathbb{R}$, com $a < b$. Suponha que f e suas derivadas $f', f'', \dots, f^{(k+1)}$ existam e sejam contínuas em I . Se $f^{(k+1)}(x) = 0$ para todo $x \in I$ e $f(x_0) = f'(x_0) = \dots = f^{(k)}(x_0) = 0$ para algum $x_0 \in I$, então $f(x) = 0$ para todo $x \in I$. De fato, pelo Teorema de Taylor unidimensional (Teorema 5.2.1), dado $x \in I$, existe ξ entre x e x_0 tal que

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

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

Uma aplicação da série de Taylor refere-se à caracterização de extremos locais.

TEOREMA 5.2.2. *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 x_0 é ponto de mínimo estrito local.*
- (2) *Se k é par e $f^{(k)}(x_0) < 0$, então x_0 é ponto de máximo estrito local.*
- (3) *Se k é ímpar, então x_0 não é extremo local.*

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

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

Supondo 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 k 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 $(x - x_0)^k/k!$ 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

5.3. Diferenciabilidade para funções de várias variáveis

A noção de diferenciabilidade e de derivada em dimensões maiores simplesmente generaliza de forma natural a derivada unidimensional. Seja $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$, onde $\Omega \subseteq \mathbb{R}^m$ é aberto e $\mathbf{x} \in \Omega$. Dizemos que \mathbf{f} é diferenciável em \mathbf{x} se existe uma aplicação linear $\mathbf{L} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ tal que

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{L}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0.$$

Chamamos \mathbf{L} de derivada de \mathbf{f} em \mathbf{x} , e que também denotamos por $D\mathbf{f}(\mathbf{x})$ ou $\mathbf{f}'(\mathbf{x})$. No caso escrevemos

$$\mathbf{L}(\mathbf{h}) = D\mathbf{f}(\mathbf{x})(\mathbf{h}) = \mathbf{f}'(\mathbf{x})(\mathbf{h}).$$

Suporemos neste texto a convenção que \mathbf{h} é sempre suficientemente pequeno de tal forma que $\mathbf{x} + \mathbf{h} \in \Omega$.

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

$$(5.3.1) \quad \mathbf{f}(\mathbf{x} + \mathbf{h}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{h}) + \mathbf{r}(\mathbf{h}) \quad \text{com} \quad \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{r}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0.$$

Note que pela identidade acima, temos imediatamente que diferenciabilidade implica em continuidade.

A derivada de uma função num determinado ponto, se existe, é única. De fato, se \mathbf{L}_1 e \mathbf{L}_2 são duas derivadas de \mathbf{f} em \mathbf{x} , então substituindo $\mathbf{h} = t\boldsymbol{\xi}$, com $\|\boldsymbol{\xi}\| = 1$ em (5.3.1) concluímos que existem funções \mathbf{r}_1 e \mathbf{r}_2 tais que

$$\begin{aligned} \mathbf{f}(\mathbf{x} + t\boldsymbol{\xi}) &= \mathbf{f}(\mathbf{x}) + t\mathbf{L}_1(\boldsymbol{\xi}) + \mathbf{r}_1(t\boldsymbol{\xi}), & \mathbf{f}(\mathbf{x} + t\boldsymbol{\xi}) &= \mathbf{f}(\mathbf{x}) + t\mathbf{L}_2(\boldsymbol{\xi}) + \mathbf{r}_2(t\boldsymbol{\xi}), \\ \lim_{t \rightarrow 0} \frac{\|\mathbf{r}_1(t\boldsymbol{\xi})\|}{\|t\boldsymbol{\xi}\|} &= \lim_{t \rightarrow 0} \frac{\|\mathbf{r}_2(t\boldsymbol{\xi})\|}{\|t\boldsymbol{\xi}\|} = 0, \end{aligned}$$

Logo concluímos que

$$\|(\mathbf{L}_1 - \mathbf{L}_2)(\boldsymbol{\xi})\| = \frac{\|\mathbf{r}_2(t\boldsymbol{\xi}) - \mathbf{r}_1(t\boldsymbol{\xi})\|}{t} \leq \frac{\|\mathbf{r}_2(t\boldsymbol{\xi})\|}{t} + \frac{\|\mathbf{r}_1(t\boldsymbol{\xi})\|}{t}.$$

Tomando o limite quando $t \rightarrow 0$ em ambos os lados da equação concluímos que $(\mathbf{L}_1 - \mathbf{L}_2)(\boldsymbol{\xi}) = 0$, para todo $\boldsymbol{\xi} \in \mathbb{R}^m$ com norma unitária. Mas isto só é possível se $\mathbf{L}_1 = \mathbf{L}_2$, como queríamos demonstrar.

EXEMPLO 5.7. Podemos usar o resultado de unicidade acima descrito para encontrar derivadas em casos simples. Como exemplo 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 $\mathbf{x}, \mathbf{h} \in \mathbb{R}^m$, e para mostrar tal fato vemos que se $\mathbf{f}'(\mathbf{x}) = \mathbf{A}$ então

$$\mathbf{r}(\mathbf{h}) = \mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})(\mathbf{h}) = \mathbf{A}(\mathbf{x} + \mathbf{h}) + \mathbf{c} - (\mathbf{A}\mathbf{x} + \mathbf{c}) - \mathbf{A}\mathbf{h} = \mathbf{0},$$

e portanto $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \|\mathbf{r}(\mathbf{h})\|/\|\mathbf{h}\| = 0$. A unicidade da derivada garante que $\mathbf{f}'(\mathbf{x})(\mathbf{h}) = \mathbf{A}(\mathbf{h})$. Note que neste caso, a derivada $\mathbf{f}'(\mathbf{x})$ é na verdade *independente* de \mathbf{x} .

EXEMPLO 5.8. 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

$$\vec{\mathbf{f}}(\mathbf{x}) = A\vec{\mathbf{x}} + \vec{\mathbf{c}} \quad \text{i.e., } f_i(\mathbf{x}) = \sum_{j=1}^m A_{ij}x_j + c_j \quad \text{para } i = 1, \dots, n.$$

Então, para $\mathbf{h} = (h_1, \dots, h_m)$ tem-se $\mathbf{f}'(\mathbf{x})(\mathbf{h}) = \mathbf{y}$ onde $\vec{\mathbf{y}} = A\vec{\mathbf{h}}$, i.e.,

$$y_i = \sum_{j=1}^m A_{ij}h_j.$$

Compare com o exemplo 5.7.

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 $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$, onde $\Omega \subseteq \mathbb{R}^m$ é aberto. Dado $\mathbf{x} \in \Omega$, seja $\epsilon > 0$ tal que $\mathbf{x} + t\mathbf{u} \in \Omega$ para todo $t \in (0, \epsilon)$. Finalmente, seja $\phi : [0, \epsilon) \rightarrow \mathbb{R}^n$, para $\phi = (\phi_1, \dots, \phi_n)$ dada por $\phi(t) = \mathbf{f}(\mathbf{x} + t\mathbf{u})$. Então, definimos a *derivada direcional de \mathbf{f} em \mathbf{x} na direção \mathbf{u}* como $\phi'(0)$, quando esta existir. Note que neste caso, $\phi'(0)$ define a aplicação linear de \mathbb{R} em \mathbb{R}^n dada por $t \mapsto (t\phi'_1(0), \dots, t\phi'_n(0))$.

Noutra forma de definir, a derivada direcional é dada por $\mathbf{L}_{\mathbf{u}} \in \mathbb{R}^m$ tal que

$$\lim_{t \rightarrow 0} \frac{\mathbf{f}(\mathbf{x} + t\mathbf{u}) - \mathbf{f}(\mathbf{x})}{t} - \mathbf{L}_{\mathbf{u}} = 0.$$

Escrevemos neste caso $\mathbf{D}_{\mathbf{u}}\mathbf{f}(\mathbf{x}) = \mathbf{L}_{\mathbf{u}}$.

No caso em que $\mathbf{u} = \mathbf{e}_i$, então temos a *derivada parcial* em relação à i -ésima coordenada e escrevemos

$$\mathbf{D}_{\mathbf{e}_i}\mathbf{f}(\mathbf{x}) = \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{x}).$$

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 5.9. 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 $ab \neq 0$, 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 5.3.1. *Seja $\Omega \subseteq \mathbb{R}^m$ aberto e $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ diferenciável em $\mathbf{x} \in \Omega$. Seja $\mathbf{u} \in \mathbb{R}^m$ com $\|\mathbf{u}\| = 1$. Então existe a derivada direcional $D_{\mathbf{u}}\mathbf{f}(\mathbf{x})$, e esta é dada por*

$$D_{\mathbf{u}}\mathbf{f}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})(\mathbf{u}).$$

DEMONSTRAÇÃO. Como \mathbf{f} é diferenciável em \mathbf{x} , então para todo $\epsilon > 0$ existe $\delta > 0$ tal que $B_{\delta}(\mathbf{x}) \in \Omega$ e

$$\mathbf{h} \in \mathbb{R}^m, 0 < \|\mathbf{h}\| < \delta \implies \frac{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})(\mathbf{h})\|}{\|\mathbf{h}\|} < \epsilon.$$

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 $\mathbf{f}'(\mathbf{x})(\mathbf{u})$. \square

O teorema acima é importante porque podemos calcular $\mathbf{f}'(\mathbf{x})$ tomando-se derivadas nas direções das coordenadas. De fato, considerando-se $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))$, temos que

$$D_{\mathbf{e}_i}\mathbf{f}(\mathbf{x}) = \left(\frac{\partial f_1}{\partial x_i}(\mathbf{x}), \frac{\partial f_2}{\partial x_i}(\mathbf{x}), \dots, \frac{\partial f_n}{\partial x_i}(\mathbf{x}) \right).$$

Usando agora a linearidade de $\mathbf{f}'(\mathbf{x})$ e que $\mathbf{f}'(\mathbf{x})(\mathbf{e}_i) = D_{\mathbf{e}_i}\mathbf{f}(\mathbf{x})$ obtemos

$$\mathbf{f}'(\mathbf{x})(\mathbf{y}) = \sum_{i=1}^m y_i \mathbf{f}'(\mathbf{x})(\mathbf{e}_i) = \sum_{i=1}^m y_i D_{\mathbf{e}_i}\mathbf{f}(\mathbf{x}) = \sum_{i=1}^m y_i \left(\frac{\partial f_1}{\partial x_i}(\mathbf{x}), \frac{\partial f_2}{\partial x_i}(\mathbf{x}), \dots, \frac{\partial f_n}{\partial x_i}(\mathbf{x}) \right).$$

Em termos matriciais, definindo a matriz $[\mathbf{f}'(\mathbf{x})] \in \mathbb{R}^{n \times m}$ dada por

$$[\mathbf{f}'(\mathbf{x})] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_m} \end{bmatrix},$$

temos que se $\boldsymbol{\xi} = \mathbf{f}'(\mathbf{x})(\mathbf{y})$ então $\vec{\boldsymbol{\xi}} = [\mathbf{f}'(\mathbf{x})]\vec{\mathbf{y}}$. A matriz $[\mathbf{f}'(\mathbf{x})]$ também é chamada de *matriz jacobiana* de f no ponto \mathbf{x} .

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

$$(5.3.2) \quad D_{(a,b)}f(0, 0) = \frac{a^2}{b}.$$

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 (5.3.2). 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 5.3.2. *Seja $f : \Omega \rightarrow \mathbb{R}$, onde $\Omega \subseteq \mathbb{R}^m$ é aberto. Se $\mathbf{x} \in \Omega$ e $\partial f / \partial x_i$ existir e for contínua numa vizinhança aberta de \mathbf{x} para $i = 1, \dots, m$, então f é diferenciável em \mathbf{x} .*

DEMONSTRAÇÃO. Dado $\epsilon > 0$, seja δ tal que

$$\|\mathbf{y} - \mathbf{x}\| < \delta \implies \left\| \frac{\partial f}{\partial x_i}(\mathbf{y}) - \frac{\partial f}{\partial x_i}(\mathbf{x}) \right\| < \frac{\epsilon}{\sqrt{m}}.$$

Dados $\mathbf{x} = (x_1, x_2, \dots, x_m)$ e $\mathbf{y} = (y_1, y_2, \dots, y_m)$, sejam

$$\begin{aligned} \mathbf{z}^0 &= \mathbf{y}, & \mathbf{z}^1 &= (x_1, y_2, y_3, \dots, y_m), & \mathbf{z}^2 &= (x_1, x_2, y_3, \dots, y_m), \\ & & \dots & & \mathbf{z}^{m-1} &= (x_1, x_2, \dots, x_{m-1}, y_m), & \mathbf{z}^m &= \mathbf{x}. \end{aligned}$$

Temos então que $\|\mathbf{y} - \mathbf{x}\| < \delta$ implica em $\|\mathbf{z}^i - \mathbf{x}\| < \delta$, para todo i . Note que

$$f(\mathbf{y}) - f(\mathbf{x}) = f(\mathbf{z}^0) - f(\mathbf{z}^1) + f(\mathbf{z}^1) - f(\mathbf{z}^2) + \dots + f(\mathbf{z}^{m-1}) - f(\mathbf{z}^m).$$

Pelo Teorema do valor médio (Teorema 5.1.4), existe $\hat{\mathbf{z}}^i$ no segmento determinado por \mathbf{z}^{i-1} e \mathbf{z}^i tal que

$$f(\mathbf{z}^i) - f(\mathbf{z}^{i-1}) = (y_i - x_i) \frac{\partial f}{\partial x_i}(\hat{\mathbf{z}}^i).$$

Logo,

(5.3.3)

$$\begin{aligned} \left| f(\mathbf{y}) - f(\mathbf{x}) - \sum_{i=1}^m \frac{\partial f}{\partial x_i}(\mathbf{x})(y_i - x_i) \right| &\leq \sum_{i=1}^m \left| \frac{\partial f}{\partial x_i}(\hat{\mathbf{z}}^i) - \frac{\partial f}{\partial x_i}(\mathbf{x}) \right| |y_i - x_i| \leq \frac{\epsilon}{\sqrt{m}} \sum_{i=1}^m |y_i - x_i| \\ &\leq \epsilon \|\mathbf{y} - \mathbf{x}\|, \end{aligned}$$

onde usamos a desigualdade de Cauchy–Schwartz para obter a última desigualdade. Portanto de (5.3.3) concluímos que f é diferenciável em \mathbf{x} . \square

COROLÁRIO 5.3.3. *Seja $f : \Omega \rightarrow \mathbb{R}^n$, onde $\Omega \subseteq \mathbb{R}^m$ é aberto. Se $\mathbf{x} \in \Omega$ e $\partial \mathbf{f} / \partial x_i$ existir e for contínua numa vizinhança aberta de \mathbf{x} para $i = 1, \dots, m$, então 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 5.3.4 (Regra da Cadeia). *Sejam $\Omega \subseteq \mathbb{R}^l$ e $R \subseteq \mathbb{R}^m$ conjuntos abertos. Sejam $\mathbf{f} : \Omega \rightarrow R$ e $\mathbf{g} : R \rightarrow \mathbb{R}^n$. Se \mathbf{f} é diferenciável em $\mathbf{x} \in \Omega$ e \mathbf{g} é diferenciável em $\mathbf{f}(\mathbf{x})$, então $\mathbf{g} \circ \mathbf{f}$ é diferenciável em \mathbf{x} e*

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}) = \mathbf{g}'(\mathbf{f}(\mathbf{x})) \circ \mathbf{f}'(\mathbf{x}).$$

DEMONSTRAÇÃO. Seja $\mathbf{y} = \mathbf{f}(\mathbf{x})$. Note que para \mathbf{h} tal que $\mathbf{x} + \mathbf{h} \in \Omega$ e \mathbf{k} tal que $\mathbf{y} + \mathbf{k} \in R$, temos

$$\begin{aligned} \mathbf{f}(\mathbf{x} + \mathbf{h}) &= \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{h}) + \mathbf{r}(\mathbf{h}) \quad \text{com} \quad \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{r}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0, \\ \mathbf{g}(\mathbf{y} + \mathbf{k}) &= \mathbf{g}(\mathbf{y}) + \mathbf{g}'(\mathbf{y})(\mathbf{k}) + \mathbf{p}(\mathbf{k}) \quad \text{com} \quad \lim_{\mathbf{k} \rightarrow \mathbf{0}} \frac{\|\mathbf{p}(\mathbf{k})\|}{\|\mathbf{k}\|} = 0, \end{aligned}$$

e em particular, para $\epsilon > 0$ fixado, existe $\delta > 0$ tal que

$$(5.3.4) \quad \|\mathbf{k}\| < \delta \implies \frac{\|\mathbf{p}(\mathbf{k})\|}{\|\mathbf{k}\|} < \epsilon.$$

Definindo $\mathbf{k} = \mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})(\mathbf{h}) + \mathbf{r}(\mathbf{h})$, temos

$$\begin{aligned} \mathbf{g} \circ \mathbf{f}(\mathbf{x} + \mathbf{h}) &= \mathbf{g}(\mathbf{f}(\mathbf{x} + \mathbf{h})) = \mathbf{g}(\mathbf{y} + \mathbf{k}) = \mathbf{g}(\mathbf{y}) + \mathbf{g}'(\mathbf{y})(\mathbf{k}) + \mathbf{p}(\mathbf{k}) \\ &= \mathbf{g}(\mathbf{y}) + \mathbf{g}'(\mathbf{y})[\mathbf{f}'(\mathbf{x})(\mathbf{h}) + \mathbf{r}(\mathbf{h})] + \mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})) = \mathbf{g}(\mathbf{y}) + \mathbf{g}'(\mathbf{y})\mathbf{f}'(\mathbf{x})(\mathbf{h}) + \mathbf{q}(\mathbf{h}) \end{aligned}$$

onde $\mathbf{q}(\mathbf{h}) = \mathbf{g}'(\mathbf{y})\mathbf{r}(\mathbf{h}) + \mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}))$. Finalmente,

$$(5.3.5) \quad \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{q}(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{g}'(\mathbf{y}) \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{r}(\mathbf{h})}{\|\mathbf{h}\|} + \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}))}{\|\mathbf{h}\|} = \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}))}{\|\mathbf{h}\|}.$$

Temos que mostrar que o limite acima se anula. Seja $\mathbf{h} \in \mathbb{R}^l \setminus \{\mathbf{0}\}$ tal que $\mathbf{x} + \mathbf{h} \in \Omega$. Duas situações mutuamente exclusivas se apresentam.

- (1) Se \mathbf{h} é tal que $\mathbf{f}(\mathbf{x} + \mathbf{h}) = \mathbf{f}(\mathbf{x})$, então $\mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})) = \mathbf{0}$.
- (2) Caso contrário, temos

$$\begin{aligned} \frac{\|\mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}))\|}{\|\mathbf{h}\|} &= \frac{\|\mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}))\|}{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})\|} \frac{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})\|}{\|\mathbf{h}\|} \\ &\leq \frac{\|\mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}))\|}{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})\|} \left(\|\mathbf{f}'(\mathbf{x})\| + \frac{\|\mathbf{r}(\mathbf{h})\|}{\|\mathbf{h}\|} \right) \leq c \frac{\|\mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}))\|}{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})\|}. \end{aligned}$$

onde c é constante tal que $\|\mathbf{f}'(\mathbf{x})\| + \|\mathbf{r}(\mathbf{h})\|/\|\mathbf{h}\| < c$. Portanto, da continuidade de \mathbf{f} e por (5.3.4), dado $\epsilon > 0$, existem $\delta, \delta' > 0$ tais que

$$\|\mathbf{h}\| < \delta \implies \|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})\| < \delta' \implies \frac{\|\mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}))\|}{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})\|} < \frac{\epsilon}{c}.$$

Dos casos (1) e (2) acima, concluímos que dado $\epsilon > 0$ existe $\delta > 0$ tal que

$$0 < \|\mathbf{h}\| < \delta'' \implies \frac{\|\mathbf{p}(\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}))\|}{\|\mathbf{h}\|} < \epsilon.$$

i.e.,

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

Portanto, concluímos por (5.3.5) que

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{q}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0,$$

donde obtemos o resultado. □

EXEMPLO 5.10. 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 $\mathbf{f}'(\mathbf{x})\mathbf{g}'(\mathbf{y}) = \mathbf{I}$.

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 5.1.4) quando restringimos uma função de várias variáveis a um segmento de reta.

TEOREMA 5.3.5. *Seja $f : \Omega \rightarrow \mathbb{R}$, diferenciável em Ω , onde $\Omega \subseteq \mathbb{R}^m$ é aberto. Sejam $\mathbf{x}, \mathbf{y} \in \Omega$ e seja $S = \{\mathbf{x} + t(\mathbf{y} - \mathbf{x}) : t \in [0, 1]\}$ o segmento de reta unindo estes pontos. Se $S \subseteq \Omega$, então existe $\boldsymbol{\xi} \in S$ tal que*

$$f(\mathbf{y}) - f(\mathbf{x}) = \mathbf{f}'(\boldsymbol{\xi})(\mathbf{y} - \mathbf{x}).$$

DEMONSTRAÇÃO. Este resultado segue-se de uma aplicação do teorema do valor médio unidimensional (Teorema 5.1.4) 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 ξ .

Existe entretanto o seguinte resultado para funções em \mathbb{R}^n .

TEOREMA 5.3.6. *Seja $f : \Omega \rightarrow \mathbb{R}^n$, diferenciável em Ω , onde $\Omega \subseteq \mathbb{R}^m$ é aberto. Sejam $\mathbf{x}_0, \mathbf{x}_1 \in \Omega$ e seja S o segmento de reta unindo estes pontos. Se $S \subseteq \Omega$, então existe $\boldsymbol{\xi} \in S$ tal que*

$$\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)\| \leq \|\mathbf{f}'(\boldsymbol{\xi})(\mathbf{x}_1 - \mathbf{x}_0)\|.$$

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 5.3.5, existe $\boldsymbol{\xi} \in S$ tal que $\phi(\mathbf{x}_1) - \phi(\mathbf{x}_0) = \phi'(\boldsymbol{\xi})(\mathbf{x}_1 - \mathbf{x}_0)$, i.e.,

$$\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)\|^2 = [\mathbf{f}'(\boldsymbol{\xi})(\mathbf{x}_1 - \mathbf{x}_0)] \cdot (\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)) \leq \|\mathbf{f}'(\boldsymbol{\xi})(\mathbf{x}_1 - \mathbf{x}_0)\| \|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_0)\|.$$

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. \square

EXEMPLO 5.11. Seja $\mathbf{f} : B_1(\mathbf{0}) \rightarrow \mathbb{R}^m$ diferenciável e com derivada limitada em $B_1(\mathbf{0})$. Se (\mathbf{x}_n) é sequência de Cauchy em $B_1(\mathbf{0})$, então $(\mathbf{f}(\mathbf{x}_n))$ é sequência de Cauchy em \mathbb{R}^m .

Para mostrar este fato, como \mathbf{f} tem derivada limitada, seja c constante tal que $\|\mathbf{f}'(\mathbf{x})\| < c$ para todo $\mathbf{x} \in B_1(\mathbf{0})$. Dado $\epsilon > 0$, como (\mathbf{x}_i) é sequência de Cauchy em $B_1(\mathbf{0})$, então existe N tal que

$$i, j > N \implies \|\mathbf{x}_i - \mathbf{x}_j\| < \frac{\epsilon}{c}.$$

Pelo Teorema do valor Médio, temos para todo \mathbf{x}_i e \mathbf{x}_j que existe $\boldsymbol{\xi}_{i,j} \in B_1(\mathbf{0})$ tal que

$$\|\mathbf{f}(\mathbf{x}_i) - \mathbf{f}(\mathbf{x}_j)\| = \|\mathbf{f}'(\boldsymbol{\xi}_{i,j})(\mathbf{x}_i - \mathbf{x}_j)\|.$$

Logo,

$$\|\mathbf{f}(\mathbf{x}_i) - \mathbf{f}(\mathbf{x}_j)\| \leq \|\mathbf{f}'(\boldsymbol{\xi}_{i,j})\| \|\mathbf{x}_i - \mathbf{x}_j\| \leq c \|\mathbf{x}_i - \mathbf{x}_j\|,$$

e portanto

$$i, j > N \implies \|\mathbf{f}(\mathbf{x}_i) - \mathbf{f}(\mathbf{x}_j)\| \leq c \|\mathbf{x}_i - \mathbf{x}_j\| < \epsilon,$$

e $(\mathbf{f}(\mathbf{x}_i))$ é sequência de Cauchy.

Encontramos na demonstração do resultado abaixo uma outra aplicação da regra da cadeia, desta vez para funções de \mathbb{R} em \mathbb{R} .

TEOREMA 5.3.7 (Derivada da Função Inversa). *Seja I 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. Se $y \in J \setminus \{d\}$, então $g(y) \neq c$. Logo, se $f'(c) \neq 0$,

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

onde usamos a continuidade de g no último passo. Concluimos que 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 5.12. 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 $f'(0) = 0$.

5.4. 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} com certa capacidade de aproximar a função f no ponto \mathbf{x} . 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 $f''(\mathbf{x}) : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ tal que

$$f''(\mathbf{x})(\mathbf{y}, \mathbf{z}) = \sum_{i,j=1}^m \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} y_i z_j, \quad \text{onde } \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right),$$

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

$$f''(\mathbf{x})(\mathbf{y}, \mathbf{z}) = (\vec{\mathbf{y}})^t H(\mathbf{x}) \vec{\mathbf{z}}.$$

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 [14], e à *Schwartz* em [2, 7]

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 5.4.1 (Taylor). *Seja $\Omega \subseteq \mathbb{R}^m$ aberto e $f : \Omega \rightarrow \mathbb{R}$ duas vezes diferenciável em Ω , com derivadas contínuas. Para $\mathbf{x} \in \Omega$, e $\mathbf{h} \in \mathbb{R}^m$ tais que $\mathbf{x} + t\mathbf{h} \in \Omega$ para todo $t \in [0, 1]$, existe $\hat{t} \in (0, 1)$ tal que para $\boldsymbol{\xi} = \mathbf{x} + \hat{t}\mathbf{h}$ tem-se*

$$(5.4.1) \quad f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{h}) + \frac{1}{2}f''(\boldsymbol{\xi})(\mathbf{h}, \mathbf{h}).$$

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 5.2.1), obtemos que existe $\hat{t} \in (0, 1)$ tal que

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2}\phi''(\hat{t}).$$

Usando a definição de ϕ obtemos o resultado diretamente. \square

Assim como em uma dimensão, usaremos o Teorema de Taylor para estudarmos pontos extremos de uma função. Dizemos que $f : \Omega \rightarrow \mathbb{R}$, onde $\Omega \subseteq \mathbb{R}^m$, tem um *máximo local* em $\mathbf{x} \in \Omega$ se existe $\delta > 0$ tal que

$$(5.4.2) \quad \mathbf{y} \in B_\delta(\mathbf{x}) \cap \Omega \implies f(\mathbf{y}) \leq f(\mathbf{x}).$$

Dizemos que \mathbf{x} é *máximo estrito local* se valer a desigualdade estrita em (5.4.2). 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 5.1.2, 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 5.4.2 (Ponto extremo interior). *Seja $f : \Omega \rightarrow \mathbb{R}$, onde $\Omega \subseteq \mathbb{R}^m$ é aberto, e $\mathbf{x} \in \Omega$ 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 5.3.1 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 5.1.2 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 5.4.1), 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

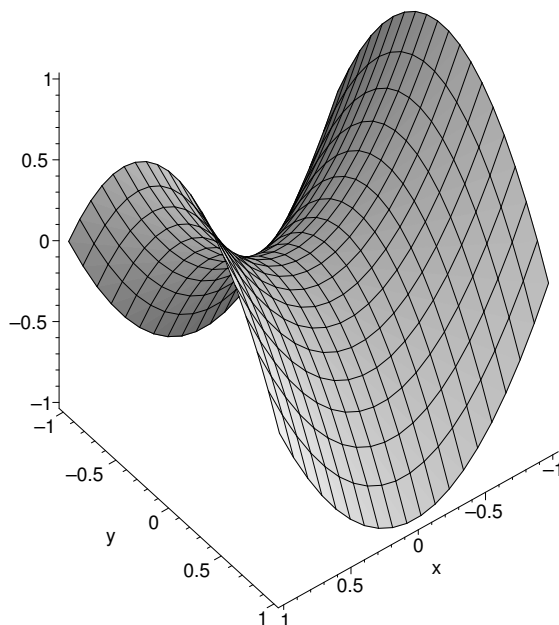


FIG. 6. Gráfico de $x^2 - y^2$, que tem ponto de sela em $(0, 0)$.

que $f''(\xi_s) \rightarrow f''(\mathbf{x})$. Como \mathbf{x} é mínimo local, então $f(\mathbf{x} + s\mathbf{u}) - f(\mathbf{x}) \geq 0$ para todo s suficientemente pequeno. Portanto $f''(\xi_s)(\mathbf{u}, \mathbf{u}) \geq 0$, 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 $f(x) = x^3$, Figura 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 6.

O resultado a seguir dá 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}, ξ em \mathbb{R}^m tais que $[f''(\mathbf{x})(\mathbf{h}, \mathbf{h})][f''(\mathbf{x})(\xi, \xi)] < 0$. Aí então \mathbf{x} é ponto de sela.

TEOREMA 5.4.3. *Seja $\Omega \subseteq \mathbb{R}^m$ aberto e $f : \Omega \rightarrow \mathbb{R}$ duas vezes diferenciável, com derivadas contínuas, e $\mathbf{x} \in \Omega$ 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 $\mathbf{y} \in \Omega \setminus \{\mathbf{x}\}$ satisfazendo as condições do Teorema 5.4.1, 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

$$(5.4.3) \quad f(\mathbf{y}) - f(\mathbf{x}) = \frac{1}{2} f''(\boldsymbol{\xi})(\mathbf{y} - \mathbf{x}, \mathbf{y} - \mathbf{x}).$$

Portanto \mathbf{x} é mínimo estrito local pois a expressão do lado direito de (5.4.3) 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 [4] apresenta este e vários outros resultados relacionados ao tema.

EXEMPLO 5.13. Seja $F : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $A \in \mathbb{R}^{n \times n}$ é simétrica positiva definida, $\vec{\mathbf{b}} \in \mathbb{R}^{n \times 1}$, 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 $[F'(\mathbf{x}_*)] \in \mathbb{R}^{n \times 1}$ é dada por

$$[F'(\mathbf{x}_*)] = (\mathbf{x}_*)^t A + \mathbf{b}^t,$$

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 .

EXEMPLO 5.14. [Método de Newton] O exemplo 5.13 acima, conjuntamente com a fórmula 5.4.1 motiva a seguinte linha de raciocínio. Seja $f : \Omega \rightarrow \mathbb{R}$ uma função suficientemente suave (tres vezes diferenciável, com a terceira derivada contínua) e $\mathbf{x}_* \in \Omega$ ponto de mínimo local com hessiana positiva definida. Então, para $\mathbf{x} \in \Omega$ e \mathbf{h} suficientemente pequeno temos que

$$g(\mathbf{h}) \stackrel{\text{def}}{=} f(\mathbf{x}) + \vec{\mathbf{b}}^t \vec{\mathbf{h}} + \frac{1}{2} \vec{\mathbf{h}}^t H \vec{\mathbf{h}} \approx f(\mathbf{x} + \mathbf{h}),$$

onde \mathbf{b} e H são as matrizes jacobiana e hessiana de f em \mathbf{x} . Note que não temos uma igualdade na expressão acima pois a hessiana é calculada em \mathbf{x} , e não num ponto entre \mathbf{x} e $\mathbf{x} + \mathbf{h}$, como nos diz o Teorema de Taylor 5.4.1. Então \mathbf{h} dado por $\vec{\mathbf{h}} = -H^{-1} \vec{\mathbf{b}}$ é mínimo local de g .

Temos então o seguinte esquema iterativo. Seja $\mathbf{x}_0 \in \Omega$ próximo o suficiente de \mathbf{x}_* , e dado $\mathbf{x}_k \in \Omega$ seja $\mathbf{x}_{k+1} \in \Omega$ definido por

$$\mathbf{x}_{k+1} = \mathbf{x}_k - H^{-1} \mathbf{f}'(\mathbf{x}_k).$$

Alguns fatos podem ser demonstrados:

- (1) a sequência (\mathbf{x}_k) está de fato bem-definida em Ω .
- (2) $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}_*$, e existe constante c tal que

$$\|\mathbf{x}_* - \mathbf{x}_{k+1}\| \leq c \|\mathbf{x}_* - \mathbf{x}_k\|^2,$$

i.e., a convergência é quadrática.

O método acima descrito é o *Método de Newton* para localizar pontos mínimos locais. Ver mais detalhes em [11, 1].

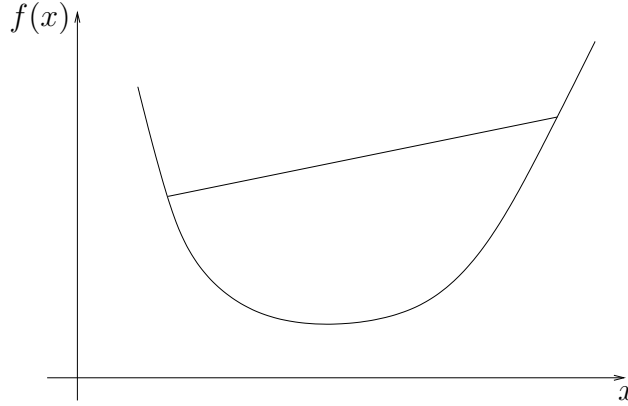


FIG. 7. Função convexa.

Uma segunda aplicação do Teorema 5.4.1 diz respeito às funções convexas definidas em conjuntos 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 Ω 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 7.

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 [10].

TEOREMA 5.4.4. *Seja $\Omega \subseteq \mathbb{R}^m$ conjunto aberto e convexo e $f : \Omega \rightarrow \mathbb{R}$ 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 $\mathbf{x} \in \Omega$.

DEMONSTRAÇÃO. (\Leftarrow) Suponha que $f''(\mathbf{x})$ seja semi-definida positiva em Ω . Seja S o segmento de reta unindo \mathbf{x} e $\mathbf{y} \in \Omega$, 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

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}f''(\boldsymbol{\xi}_1)(\mathbf{x} - \mathbf{x}_0, \mathbf{x} - \mathbf{x}_0),$$

$$f(\mathbf{y}) = f(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)(\mathbf{y} - \mathbf{x}_0) + \frac{1}{2}f''(\boldsymbol{\xi}_2)(\mathbf{y} - \mathbf{x}_0, \mathbf{y} - \mathbf{x}_0).$$

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,

$$f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y})$$

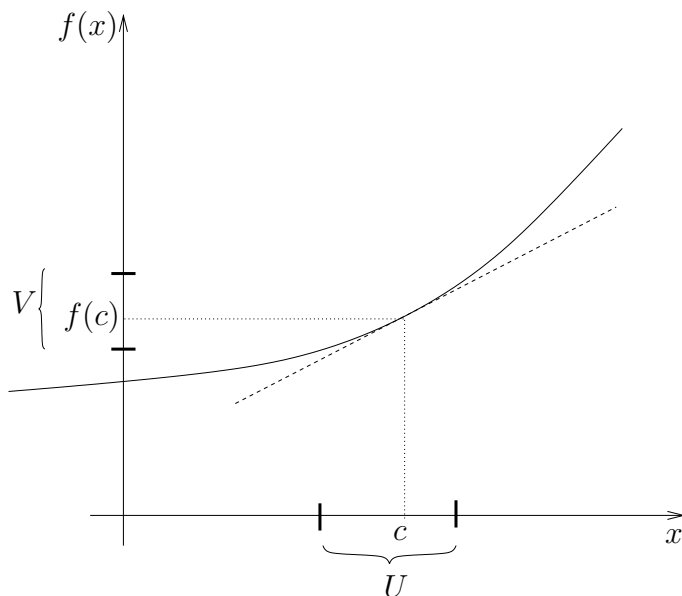


FIG. 8. Teorema da função inversa.

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 $\mathbf{f}'(\mathbf{x})(\mathbf{x} - \mathbf{y}) \leq f(\mathbf{y}) - f(\mathbf{x})$. Seja $s = \|\mathbf{x} - \mathbf{y}\|$ e $\mathbf{h} = (\mathbf{y} - \mathbf{x})/s \in B_1(\mathbf{0})$. Usando agora a fórmula de Taylor obtemos que existe $\hat{s} \in (0, c)$ tal que

$$\frac{1}{2}f''(\mathbf{x} + \hat{s}\mathbf{h})(s\mathbf{h}, s\mathbf{h}) = f(\mathbf{y}) - f(\mathbf{x}) - \mathbf{f}'(\mathbf{x})(\mathbf{x} - \mathbf{y}) \geq 0.$$

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

OBSERVAÇÃO. Note que no processo de demonstração do Teorema 5.4.4, mostramos também que uma função f ser convexa implica em $\mathbf{f}'(\mathbf{x})(\mathbf{x} - \mathbf{y}) \leq f(\mathbf{y}) - f(\mathbf{x})$ para todo \mathbf{x}, \mathbf{y} .

5.5. Teorema da Função Implícita e da Função Inversa

5.5.1. 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 Ω . A regularidade exigida nos garante o seguinte resultado.

LEMA 5.5.1. Seja $\Omega \subseteq \mathbb{R}^m$ aberto, e $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$ de classe $\mathcal{C}^1(\Omega)$. Suponha que $D = \mathbf{f}'(\hat{\mathbf{x}})$ seja invertível, onde $\hat{\mathbf{x}} \in \Omega$. Então $\mathbf{f}'(\mathbf{x})$ é invertível com inversa contínua, na vizinhança aberta de $\hat{\mathbf{x}}$ dada por

$$(5.5.1) \quad U = \left\{ \mathbf{x} \in \Omega : \|D - \mathbf{f}'(\mathbf{x})\| < \frac{1}{2\|D^{-1}\|} \right\}.$$

DEMONSTRAÇÃO. Como $\mathbf{f} \in \mathcal{C}^1(\Omega)$, então U é de fato vizinhança aberta de $\hat{\mathbf{x}}$. Se \mathbf{f}' fosse não invertível para algum $\mathbf{x} \in U$, existiria $\boldsymbol{\xi} \in \mathbb{R}^m$ não nulo tal que $\mathbf{f}'(\mathbf{x})\boldsymbol{\xi} = 0$ (por que?). Mas então

$$\|D\boldsymbol{\xi}\| = \|D\boldsymbol{\xi} - \mathbf{f}'(\mathbf{x})\boldsymbol{\xi}\| < \frac{1}{2\|D^{-1}\|}\|\boldsymbol{\xi}\|.$$

Finalmente,

$$\|\boldsymbol{\xi}\| = \|D^{-1}D\boldsymbol{\xi}\| \leq \|D^{-1}\|\|D\boldsymbol{\xi}\| < \|D^{-1}\|\frac{1}{2\|D^{-1}\|}\|\boldsymbol{\xi}\| = \frac{1}{2}\|\boldsymbol{\xi}\|,$$

uma contradição. Logo tal \mathbf{x} não existe e \mathbf{f}' é invertível em U .

Para mostrar que a inversa de \mathbf{f}' é uma função contínua em $\mathbf{x}_0 \in U$, seja $D_0 = \mathbf{f}'(\mathbf{x}_0)$ e $A(\mathbf{x}) = \mathbf{f}'(\mathbf{x})$. Usando o fato de que para cada $\mathbf{x} \in U$, o operador linear $A(\mathbf{x}) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é uma sobrejeção, temos que

$$(5.5.2) \quad \|D_0^{-1} - A^{-1}(\mathbf{x})\| = \sup_{\boldsymbol{\eta} \in \mathbb{R}^m, \boldsymbol{\eta} \neq \mathbf{0}} \frac{\|D_0^{-1}\boldsymbol{\eta} - A^{-1}(\mathbf{x})\boldsymbol{\eta}\|}{\|\boldsymbol{\eta}\|} = \sup_{\boldsymbol{\xi} \in \mathbb{R}^m, \boldsymbol{\xi} \neq \mathbf{0}} \frac{\|D_0^{-1}A(\mathbf{x})\boldsymbol{\xi} - \boldsymbol{\xi}\|}{\|A(\mathbf{x})\boldsymbol{\xi}\|} \\ \leq \|D_0^{-1}A(\mathbf{x}) - I\| \sup_{\boldsymbol{\xi} \in \mathbb{R}^m, \boldsymbol{\xi} \neq \mathbf{0}} \frac{\|\boldsymbol{\xi}\|}{\|A(\mathbf{x})\boldsymbol{\xi}\|}$$

Note primeiro que

$$(5.5.3) \quad \mathbf{x} \rightarrow \mathbf{x}_0 \implies \|D_0^{-1}A(\mathbf{x}) - I\| \rightarrow 0.$$

Portanto, basta mostrar que existe uma constante M tal que

$$(5.5.4) \quad \sup_{\boldsymbol{\xi} \in \mathbb{R}^m, \boldsymbol{\xi} \neq \mathbf{0}} \frac{\|\boldsymbol{\xi}\|}{\|A(\mathbf{x})\boldsymbol{\xi}\|} \leq M,$$

pois por (5.5.2), (5.5.3), (5.5.4), temos que $A^{-1}(\mathbf{x}) \rightarrow D_0^{-1}$ se $\mathbf{x} \rightarrow \mathbf{x}_0$. Portanto A^{-1} é contínua em \mathbf{x}_0 . Como \mathbf{x}_0 é arbitrário, o resultado vale para todo ponto de U .

Para mostrar (5.5.4), note que para cada $\boldsymbol{\xi}$ não nulo fixo temos

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|\boldsymbol{\xi}\|}{\|A(\mathbf{x})\boldsymbol{\xi}\|} = \frac{\|\boldsymbol{\xi}\|}{\|D_0\boldsymbol{\xi}\|} \leq \sup_{\boldsymbol{\xi} \in \mathbb{R}^m, \boldsymbol{\xi} \neq \mathbf{0}} \frac{\|\boldsymbol{\xi}\|}{\|D_0\boldsymbol{\xi}\|} = \sup_{\boldsymbol{\zeta} \in \mathbb{R}^m, \boldsymbol{\zeta} \neq \mathbf{0}} \frac{\|D_0^{-1}\boldsymbol{\zeta}\|}{\|\boldsymbol{\zeta}\|} = \|D_0^{-1}\|.$$

Portanto, existe uma vizinhança aberta de \mathbf{x}_0 e uma constante M tal que para todo \mathbf{x} nesta vizinhança, e todo $\boldsymbol{\xi} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$,

$$\frac{\|\boldsymbol{\xi}\|}{\|A(\mathbf{x})\boldsymbol{\xi}\|} \leq M.$$

Tomando o supremo em $\boldsymbol{\xi}$ temos (5.5.4). □

TEOREMA 5.5.2 (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 dada a vizinhança aberta U de $\hat{\mathbf{x}}$ definida por (5.5.1), temos 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 $\mathbf{g}'(\hat{\mathbf{y}}) = [\mathbf{f}'(\hat{\mathbf{x}})]^{-1}$.

DEMONSTRAÇÃO. (Rudin) Para mostrar (1) começamos definindo

$$(5.5.5) \quad \lambda = \frac{1}{2\|D^{-1}\|}.$$

Como \mathbf{f}' é contínua, então U é de fato vizinhança aberta de $\hat{\mathbf{x}}$. Definindo para $\mathbf{y} \in \mathbb{R}^m$ a função $\phi : \Omega \rightarrow \mathbb{R}^m$ dada por

$$(5.5.6) \quad \phi(\mathbf{x}) = \mathbf{x} + D^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x})),$$

temos que $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ se e só se \mathbf{x} é ponto fixo de ϕ .

Mas $\phi'(\mathbf{x}) = I - D^{-1}\mathbf{f}'(\mathbf{x}) = D^{-1}(D - \mathbf{f}'(\mathbf{x}))$, e portanto $\|\phi'(\mathbf{x})\| \leq 1/2$ em U . Logo, se $\mathbf{y} \in U$, temos pelo Teorema 5.3.6 que

$$(5.5.7) \quad \|\phi(\mathbf{x}) - \phi(\mathbf{y})\| \leq \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|,$$

e ϕ é contração. Portanto tem no máximo um ponto fixo, e \mathbf{f} é injetiva em U (observe que a existência de ponto fixo não está garantida pois ϕ está definida em Ω somente. Por que?).

Para mostrar que V é aberto, seja $\mathbf{y}_0 \in V$. Então $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ para algum $\mathbf{x}_0 \in U$. Considere $\rho > 0$ tal que $\overline{B_\rho(\mathbf{x}_0)} \subseteq U$. Mostraremos que $B_{\rho\lambda}(\mathbf{y}_0) \subseteq V$, para concluir que V é aberto. Para $\mathbf{y} \in B_{\rho\lambda}(\mathbf{y}_0)$,

$$\|\phi(\mathbf{x}_0) - \mathbf{x}_0\| = \|D^{-1}(\mathbf{y} - \mathbf{y}_0)\| < \|D^{-1}\|\rho\lambda = \frac{\rho}{2}.$$

Mas então para $\mathbf{x} \in \overline{B_\rho(\mathbf{x}_0)}$,

$$\|\phi(\mathbf{x}) - \mathbf{x}_0\| \leq \|\phi(\mathbf{x}) - \phi(\mathbf{x}_0)\| + \|\phi(\mathbf{x}_0) - \mathbf{x}_0\| < \frac{1}{2}\|\mathbf{x} - \mathbf{x}_0\| + \frac{\rho}{2} \leq \rho,$$

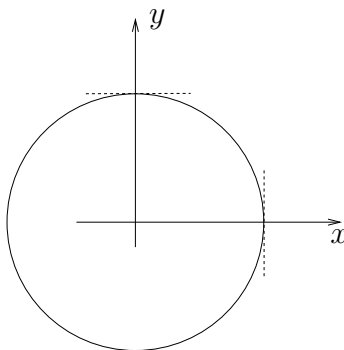
e $\phi(\mathbf{x}) \in B_\rho(\mathbf{x}_0)$. Portanto ϕ é contração de $\overline{B_\rho(\mathbf{x}_0)}$ em $\overline{B_\rho(\mathbf{x}_0)}$, e pelo Teorema 3.5.3 tem então um único ponto fixo $\mathbf{x} \in \overline{B_\rho(\mathbf{x}_0)} \subseteq U$. Logo $\mathbf{f}(\mathbf{x}) = \mathbf{y}$, e $\mathbf{y} \in \mathbf{f}(\overline{B_\rho(\mathbf{x}_0)}) \subseteq \mathbf{f}(U) = V$, como queríamos demonstrar.

Mostramos agora (2). Por (1), temos que \mathbf{f} é invertível em U , e seja $\mathbf{g} : V \rightarrow U$ sua inversa. Para $\mathbf{y}, \mathbf{y} + \mathbf{k} \in V$, existem $\mathbf{x} = \mathbf{g}(\mathbf{y})$, $\mathbf{x} + \mathbf{h} = \mathbf{g}(\mathbf{y} + \mathbf{k})$. Considerando ϕ como em (5.5.6), temos que

$$\phi(\mathbf{x} + \mathbf{h}) - \phi(\mathbf{x}) = \mathbf{h} + D^{-1}[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x} + \mathbf{h})] = \mathbf{h} - D^{-1}\mathbf{k}.$$

Por (5.5.7) temos $\|\mathbf{h} - D^{-1}\mathbf{k}\| \leq \|\mathbf{h}\|/2$ e então $\|D^{-1}\mathbf{k}\| \geq \|\mathbf{h}\|/2$. Logo

$$\|\mathbf{h}\| \leq 2\|D^{-1}\|\|\mathbf{k}\| \leq \frac{\|\mathbf{k}\|}{\lambda}.$$

FIG. 9. Conjunto $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$.

Mas por (5.5.5), e Lema 5.5.1, \mathbf{f}' é invertível em U . Denotando esta inversa por T , temos

$$\begin{aligned} \frac{\|\mathbf{g}(\mathbf{y} + \mathbf{k}) - \mathbf{g}(\mathbf{y}) - T\mathbf{k}\|}{\|\mathbf{k}\|} &= \frac{\|\mathbf{h} - T\mathbf{k}\|}{\|\mathbf{k}\|} \leq \frac{\|T(\mathbf{f}'(\mathbf{x})\mathbf{h} - \mathbf{k})\|}{\lambda\|\mathbf{h}\|} \\ &\leq \frac{\|T\|}{\lambda} \frac{\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}\|}{\|\mathbf{h}\|}. \end{aligned}$$

Tomando $\mathbf{k} \rightarrow 0$, temos $\mathbf{h} \rightarrow 0$, e o lado direito acima vai zero. Portanto o lado esquerdo também converge para zero, e $\mathbf{g}'(\mathbf{y}) = T$, que é a inversa de $\mathbf{f}'(\mathbf{x})$. Logo \mathbf{g} é diferenciável. Para concluir que $\mathbf{g} \in \mathcal{C}^1(V)$, usamos que $\mathbf{g}'(\mathbf{y}) = [\mathbf{f}'(\mathbf{g}(\mathbf{y}))]^{-1}$, e pelo Lema 5.5.1, \mathbf{f}' tem inversa contínua. \square

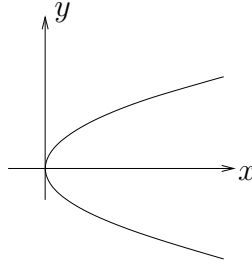
OBSERVAÇÃO. Note que o teorema acima tem caráter *local*. Em particular, é possível construir funções não injetivas em seu domínios que possuem matrizes jacobianas invertíveis em todos os pontos. Entretanto em *uma dimensão*, se a derivada não se anula em nenhum ponto de um intervalo aberto, a função é globalmente invertível.

5.5.2. 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 9. 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 ϕ 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 10. 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

FIG. 10. Conjunto $\{(x, y) \in \mathbb{R}^2 : x = y^2\}$.

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 5.5.3 (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 $\mathbf{F}(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}$. Se a transformação linear de \mathbb{R}^n em \mathbb{R}^n definida por $\mathbf{v} \mapsto \mathbf{F}'(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{0}, \mathbf{v})$ 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 5.5.2), 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,

$$(5.5.8) \quad \mathbf{x} = \phi_1(\mathbf{x}, \mathbf{y}), \quad \mathbf{y} = F(\mathbf{x}, \phi_2(\mathbf{x}, \mathbf{y})),$$

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 (5.5.8) que $F(\mathbf{x}, \phi(\mathbf{x})) = \mathbf{0}$. Como Φ é de classe \mathcal{C}^1 , então ϕ_2 , e portanto ϕ também é de classe \mathcal{C}^1 . \square

5.6. 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 demonstramos abaixo. Em ambas demonstrações que apresentamos se faz necessário o Lema da aplicação aberta (ou sobrejetiva), que apresentamos abaixo sem demonstrar [2].

LEMA 5.6.1 (aplicação aberta). *Seja Ω aberto em \mathbb{R}^m , e $f : \Omega \rightarrow \mathbb{R}^n$ de classe $\mathcal{C}^1(\Omega)$. Suponha que $\mathbf{f}'(\mathbf{x}) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ seja uma sobrejeção. Então existe U vizinhança aberta de \mathbf{x} tal que $\mathbf{f}(U)$ é aberto.*

Precisamos também de alguns conceitos de álgebra linear. Dizemos por exemplo que o conjunto de vetores $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ dum espaço vetorial é *linearmente dependente* se existem números $\lambda_1, \dots, \lambda_k$, não todos nulos, tais que $\lambda_1 \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_k = \mathbf{0}$. Estes mesmos conjunto é *linearmente independente* se não é linearmente dependente.

Um resultado importante de álgebra linear nos diz que se uma aplicação linear $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ dada por $A(\mathbf{x}) = (\mathbf{v}_1 \cdot \mathbf{x}, \dots, \mathbf{v}_n \cdot \mathbf{x})$ não é sobrejetiva, então $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ é linearmente dependente. De fato, como A não é sobrejetiva, então existe vetor não nulo $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ ortogonal a $A(\mathbf{x})$ para todo $\mathbf{x} \in \mathbb{R}^m$, i.e.,

$$0 = (\mathbf{v}_1 \cdot \mathbf{x}, \dots, \mathbf{v}_n \cdot \mathbf{x}) \cdot \boldsymbol{\xi} = (\xi_1 \mathbf{v}_1 + \dots + \xi_n \mathbf{v}_n) \cdot \mathbf{x}.$$

Mas então $\xi_1 \mathbf{v}_1 + \dots + \xi_n \mathbf{v}_n \in \mathbb{R}^m$ é ortogonal a todo vetor do \mathbb{R}^m , e isto só é possível se $\xi_1 \mathbf{v}_1 + \dots + \xi_n \mathbf{v}_n = 0$, como queríamos demonstrar.

Voltemos ao problema de minimização com restrições. Dadas funções reais f, g_1, \dots, g_k definidas num aberto Ω de \mathbb{R}^m , consideramos o problema de minimizar f restrita ao conjunto de raízes de g_1, \dots, g_k em Ω . O Teorema de Lagrange nos dá condições necessárias que um candidato a mínimo de tal problema tem que satisfazer.

TEOREMA 5.6.2 (Lagrange). *Seja $\Omega \subseteq \mathbb{R}^m$ aberto, e f, g_1, \dots, g_k funções reais definidas em Ω de classe $\mathcal{C}^1(\Omega)$. Suponha que exista um aberto $U \subseteq \Omega$ e $\mathbf{x}_* \in U$ e tal que*

$$f(\mathbf{x}_*) = \inf\{f(x) : \mathbf{x} \in U \text{ e } g_1(\mathbf{x}) = \dots = g_k(\mathbf{x}) = 0\}.$$

Então existem números $\mu, \lambda_1, \dots, \lambda_k$ não todos nulos e tais que

$$(5.6.1) \quad \mu f'(\mathbf{x}_*) = \lambda_1 g'_1(\mathbf{x}_*) + \dots + \lambda_k g'_k(\mathbf{x}_*).$$

Além disto, se $\{g'_1(\mathbf{x}_), \dots, g'_k(\mathbf{x}_*)\}$ é linearmente independente, então pode-se tomar $\mu = 1$.*

DEMONSTRAÇÃO. Seja $F : \Omega \rightarrow \mathbb{R}^{k+1}$ dada por $F(\mathbf{x}) = (f(\mathbf{x}), g_1(\mathbf{x}), \dots, g_k(\mathbf{x}))$. Então para $\mathbf{x} \in \Omega$, temos que $F'(\mathbf{x}) : \mathbb{R}^m \rightarrow \mathbb{R}^{k+1}$ é dada por

$$F'(\mathbf{x})(\mathbf{h}) = (f'(\mathbf{x})(\mathbf{h}), g'_1(\mathbf{x})(\mathbf{h}), \dots, g'_k(\mathbf{x})(\mathbf{h})).$$

Para $\mathbf{x} \in N = \{\mathbf{x} \in \Omega : g_1(\mathbf{x}) = \dots = g_k(\mathbf{x}) = 0\}$, temos que $F(\mathbf{x}) = (f(\mathbf{x}), 0, \dots, 0)$. Supondo agora que \mathbf{x}_* minimiza f restrita a N , então $F(\Omega)$ não contém pontos da forma $(f(\mathbf{x}_*) - \epsilon, 0, \dots, 0)$ para nenhum $\epsilon > 0$. Logo, pelo Lema da aplicação aberta (Lema 5.6.1), $F'(\mathbf{x}_*)$ não é uma sobrejeção. Temos portanto que $\{f'(\mathbf{x}_*), g'_1(\mathbf{x}_*), \dots, g'_k(\mathbf{x}_*)\}$ é linearmente dependente, e (5.6.1) vale.

Finalmente, se $\{g'_1(\mathbf{x}_*), \dots, g'_k(\mathbf{x}_*)\}$ é linearmente independente, então $\mu \neq 0$, pois caso contrário teríamos $\lambda_1 = \dots = \lambda_k = 0$. Logo podemos dividir os fatores da combinação linear por μ . \square

Os números $\lambda_1, \dots, \lambda_k$ acima são conhecidos por *multiplicadores de Lagrange*, e em muitas aplicações têm significado próprio.

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 5.6.3 (Kuhn–Tucker). *Seja $\Omega \subseteq \mathbb{R}^m$ aberto, e f, h_1, \dots, h_k funções reais definidas em Ω de classe $\mathcal{C}^1(\Omega)$. Suponha que exista um aberto $U \subseteq \Omega$ e $\mathbf{x}_* \in U$ e tal que*

$$f(\mathbf{x}_*) = \inf\{f(x) : \mathbf{x} \in U \text{ e } h_1(\mathbf{x}) \geq 0, \dots, h_k(\mathbf{x}) \geq 0\}.$$

Então as seguintes afirmativas são verdadeiras:

(1) existem números $\mu, \lambda_1, \dots, \lambda_k$ não todos nulos e tais que

$$\mu f'(\mathbf{x}_*) = \lambda_1 h'_1(\mathbf{x}_*) + \dots + \lambda_k h'_k(\mathbf{x}_*).$$

(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 $\lambda_1 \geq 0, \dots, \lambda_k \geq 0$.

DEMONSTRAÇÃO. (1) Neste caso, a demonstração é muito semelhante à do Teorema 5.6.2. Seja $F(\mathbf{x}) = (f(\mathbf{x}), h_1(\mathbf{x}), \dots, h_k(\mathbf{x}))$. Então para $\mathbf{x} \in \Omega$,

$$F'(\mathbf{x})(\mathbf{h}) = (f'(\mathbf{x})(\mathbf{h}), h'_1(\mathbf{x})(\mathbf{h}), \dots, h'_k(\mathbf{x})(\mathbf{h})).$$

Supondo agora que \mathbf{x}_* minimiza f restrita a $\{\mathbf{x} \in \Omega : h_1(\mathbf{x}) \geq 0, \dots, h_k(\mathbf{x}) \geq 0, \}$, então $F(\Omega)$ não contém pontos da forma $(f(\mathbf{x}_*) - \epsilon, h_1(\mathbf{x}_*), \dots, h_k(\mathbf{x}_*))$ para nenhum $\epsilon > 0$. Pelo Lema da aplicação aberta (Lema 5.6.1), $F'(\mathbf{x}_*)$ não é uma sobrejeção. Temos portanto que $\{f'(\mathbf{x}), h'_1(\mathbf{x}_*), \dots, h'_k(\mathbf{x}_*)\}$ é linearmente dependente.

(2) Se $h_{r+1}(\mathbf{x}_*) > 0, \dots, h_k(\mathbf{x}_*) > 0$ então considere um aberto U contendo \mathbf{x}_* tal que $h_{r+1}(\mathbf{x}_*) > 0, \dots, h_k(\mathbf{x}_*) > 0$ em U . Aplicando a primeira parte deste teorema com somente as primeiras r restrições $h_1(\mathbf{x}) \geq 0, \dots, h_r(\mathbf{x}) \geq 0$ obtemos o resultado.

(3) Sem perda de generalidade, considere $V = \{h'_1(\mathbf{x}_*), \dots, h'_r(\mathbf{x}_*)\}$, onde $r \leq k$. Se V é linearmente independente, argumentamos como na demonstração do Teorema 5.6.2 para tomar $\mu = 1$. Além disto, dado $\epsilon > 0$, existem vetores $\mathbf{v}_1, \dots, \mathbf{v}_r \in \mathbb{R}^m$ tais que

$$h'_i(\mathbf{x}_*) \cdot \mathbf{v}_j = \begin{cases} 1 & \text{se } i = j, \\ \epsilon & \text{se } i \neq j. \end{cases}$$

Logo existe t suficientemente pequeno e \hat{t} e \bar{t} , ambos contidos em $(0, t)$ e tais que

$$h_i(\mathbf{x}_* + t\mathbf{v}_j) = h_i(\mathbf{x}_* + t\mathbf{v}_j) - h_i(\mathbf{x}_*) = th'_i(\mathbf{x}_* + \hat{t}\mathbf{v}_j)(\mathbf{v}_j) \geq 0.$$

Logo $f(\mathbf{x}_*) \leq f(\mathbf{x}_* + t\mathbf{v}_j)$. Concluimos então que

$$\begin{aligned} 0 &\leq \lim_{\substack{t \rightarrow 0 \\ t > 0}} \frac{f(\mathbf{x}_* + t\mathbf{v}_j) - f(\mathbf{x}_*)}{t} = f'(\mathbf{x}_*)(\mathbf{v}_j) = \lambda_1 h'_1(\mathbf{x}_*)(\mathbf{v}_j) + \dots + \lambda_r h'_r(\mathbf{x}_*)(\mathbf{v}_j) \\ &= \lambda_j h'_j(\mathbf{x}_*)(\mathbf{v}_j) + \sum_{i=1, i \neq j}^r \lambda_i h'_i(\mathbf{x}_*)(\mathbf{v}_j) = \lambda_j + \epsilon \sum_{i=1, i \neq j}^r \lambda_i. \end{aligned}$$

Como ϵ pode ser tomado arbitrariamente pequeno, necessariamente tem-se $\lambda_j \geq 0$. \square

5.7. Exercícios

EXERCÍCIO 5.1. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por

$$f(x) = \begin{cases} x^2 & \text{se } x \in \mathbb{Q}, \\ 0 & \text{se } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Calcule $f'(0)$.

EXERCÍCIO 5.2. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ e $c \in \mathbb{R}$ tal que $f(c) = 0$. Mostre então que $g(x) = |f(x)|$ é diferenciável em c se e somente se f é diferenciável em c e $f'(c) = 0$.

EXERCÍCIO 5.3. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = |x|$. Note que f atinge seu mínimo em $x = 0$. Pode-se concluir então que $f'(0) = 0$? Por que?

EXERCÍCIO 5.4. 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 local de f . Mostre que é único.

EXERCÍCIO 5.5. Dê exemplo de uma função uniformemente contínua em $[0, 1]$ que seja diferenciável em $(0, 1)$ mas cuja derivada não seja limitada em $(0, 1)$. Mostre porque que o seu exemplo funciona.

EXERCÍCIO 5.6. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ função diferenciável em \mathbb{R} e tal que $f'(x) \neq 0$ para todo $x \in \mathbb{R}$. Mostre que f é injetiva.

EXERCÍCIO 5.7. Seja I 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.

EXERCÍCIO 5.8. Mostre que se I é um intervalo e $f : I \rightarrow \mathbb{R}$ diferenciável com derivada limitada em I , então f é de Lipschitz.

EXERCÍCIO 5.9. Seja $I \subseteq \mathbb{R}$ intervalo aberto e $f : I \rightarrow \mathbb{R}$ quatro vezes diferenciável, com a quarta derivada contínua, numa vizinhança aberta de $x \in I$. Mostre então que existe uma constante c tal que

$$\left| f''(x) - \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \right| \leq ch^2,$$

para h suficientemente pequeno. A forma acima é utilizada para aproximar $f''(x)$, quando f é suave.

EXERCÍCIO 5.10. Mostre que dados quaisquer $x, y \in \mathbb{R}$ fixados, o resto da série de Taylor da função $\cos x$ centrada em x e calculada em y converge para zero quando $n \rightarrow +\infty$.

EXERCÍCIO 5.11. Seja $f : A \rightarrow \mathbb{R}$ duas vezes diferenciável em A , onde $A \subseteq \mathbb{R}$ aberto, e $x \in A$ ponto crítico de f tal que $f''(x) \neq 0$ (são chamados de *não degenerados*). Mostre que existe uma vizinhança aberta de x tal que x é o único ponto crítico.

EXERCÍCIO 5.12. Sejam f e A como no exercício 5.11. Mostre que cada compacto contido em A contém um número finito de pontos críticos não degenerados.

EXERCÍCIO 5.13. 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 5.14. 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 5.15 (Kevasan [5], 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 (5.3.1)).

EXERCÍCIO 5.16. 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 $\mathbf{f}'(\mathbf{x}) = \mathbf{g}'(\mathbf{x})$ 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 5.17. 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 5.18 (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 5.19. Mostre, usando o Teorema 5.4.3, 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 5.20. 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 5.21. 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 ∇f é chamado de *vetor gradiente de f* , e dá a direção de “maior crescimento” da função f no ponto \mathbf{x} .

CAPÍTULO 6

Sequência de Funções

¹ Seja $\Omega \subset \mathbb{R}^m$ e $f_i : \Omega \rightarrow \mathbb{R}^n$, onde $i \in \mathbb{N}$. Dizemos então que (f_i) define uma sequência de funções. Note que cada $\mathbf{x} \in \Omega$ define a sequência $(f_i(\mathbf{x}))$ em \mathbb{R}^n .

6.1. Convergência Pontual

DEFINIÇÃO 6.1.1. *Seja (f_i) uma sequência de funções, onde $f_i : \Omega \rightarrow \mathbb{R}^n$, e $\Omega \subset \mathbb{R}^m$. Dizemos que (f_i) converge pontualmente para uma função $f : \Omega_0 \rightarrow \mathbb{R}^n$ em $\Omega_0 \subset \Omega$ se para todo $\mathbf{x} \in \Omega_0$, a sequência $(f_i(\mathbf{x}))$ converge para $f(\mathbf{x})$.*

EXEMPLO 6.1. Sejam $f_i(x) = x/i$ e $f(x) = 0$. Então f_i converge pontualmente para f em \mathbb{R} , pois para todo $x \in \mathbb{R}$ tem-se $\lim_{i \rightarrow \infty} f_i(x) = \lim_{i \rightarrow \infty} x/i = 0$.

EXEMPLO 6.2. Sejam $g_i(x) = x^i$. Então

(1) Se $x \in (-1, 1)$, então $\lim_{i \rightarrow \infty} g_i(x) = \lim_{i \rightarrow \infty} x^i = 0$.

(2) Se $x = 1$, então $\lim_{i \rightarrow \infty} g_i(x) = \lim_{i \rightarrow \infty} 1 = 1$.

(3) Se $x = -1$, então $g_i(x) = (-1)^n = 1$ não converge.

(4) Se $|x| > 1$, então $g_i(x)$ não é limitada e portanto não converge.

Logo (g_i) converge pontualmente para g em $(-1, 1]$, onde

$$(6.1.1) \quad g(x) = \begin{cases} 0 & \text{se } -1 < x < 1, \\ 1 & \text{se } x = 1. \end{cases}$$

Note que

$$0 = \lim_{x \rightarrow 1^-} g(x) = \lim_{x \rightarrow 1^-} \lim_{i \rightarrow +\infty} g_i(x) \neq \lim_{i \rightarrow +\infty} \lim_{x \rightarrow 1^-} g_i(x) = 1.$$

Note que a definição de convergência pontual pode ser escrita da seguinte forma.

DEFINIÇÃO 6.1.2. *Uma sequência de funções (f_i) onde $f_i : \Omega \rightarrow \mathbb{R}^n$, e $\Omega \subset \mathbb{R}^m$ converge pontualmente para uma função $f : \Omega_0 \rightarrow \mathbb{R}^n$ em $\Omega_0 \subset \Omega$ se para dado $\epsilon > 0$ e $\mathbf{x} \in \Omega_0$, existe $N_0(\mathbf{x}, \epsilon)$ tal que*

$$n > N_0(\mathbf{x}, \epsilon) \implies |f_i(\mathbf{x}) - f(\mathbf{x})| < \epsilon.$$

O que fica claro na definição acima é que a “escolha de N_0 ” depende do ponto \mathbf{x} em consideração. Considere o exemplo 6.1, e seja $\epsilon = 1/10$. Então, para $x = 1$ e $N_0(x, \epsilon) = 10$, temos

$$n > N_0(x, \epsilon) = 10 \implies |f_i(x) - f(x)| = |1/n| < \epsilon.$$

Mas para $x = 2$, a escolha anterior de $N_0 = 10$ já não é suficiente e temos que escolher $N_0(x, \epsilon) \geq 20$.

¹Última Atualização: 30/05/2007

6.2. Convergência Uniforme

DEFINIÇÃO 6.2.1. *Dados $\Omega \subset \mathbb{R}^m$ e $i \in \mathbb{N}$, seja $f_i : \Omega \rightarrow \mathbb{R}^n$. Dizemos que a sequência de funções (f_i) , converge uniformemente para $f : \Omega \rightarrow \mathbb{R}^n$, se dado $\epsilon > 0$ existe $N_0(\epsilon)$ tal que*

$$n > N_0 \implies \|f_i(\mathbf{x}) - f(\mathbf{x})\| < \epsilon \text{ para todo } \mathbf{x} \in \Omega.$$

Observe que convergência uniforme implica em convergência pontual, mas que a afirmação recíproca não vale. Uma forma prática de se mostrar que uma sequência de funções não converge uniformemente é utilizando o resultado abaixo.

TEOREMA 6.2.2. *Seja $f_i : \Omega \rightarrow \mathbb{R}^n$ onde $\Omega \subset \mathbb{R}^m$ e $i \in \mathbb{N}$. Então a sequência de funções (f_i) não converge uniformemente para $f : \Omega \rightarrow \mathbb{R}^n$ se e somente se para algum $\epsilon > 0$ existir uma subsequência (f_{i_k}) e uma sequência de pontos (\mathbf{x}_k) em Ω tais que*

$$\|f_{i_k}(\mathbf{x}_k) - f(\mathbf{x}_k)\| \geq \epsilon \text{ para todo } k \in \mathbb{N}.$$

EXEMPLO 6.3. Sejam $f_i : \mathbb{R} \rightarrow \mathbb{R}$ e $f : \mathbb{R} \rightarrow \mathbb{R}$, onde $f_i(x) = x/i$ e $f(x) = 0$. Tome $\epsilon = 1/2$, $n_k = k$ e $x_k = k$. Então

$$|f_{n_k}(x_k) - f(x_k)| = 1 > \epsilon.$$

Logo não há convergência uniforme.

Uma forma de “medir” convergência uniforme é através da norma do supremo, que a cada função limitada associa o valor máximo do módulo desta. Formalmente temos a seguinte definição.

DEFINIÇÃO 6.2.3. *Seja $f : \Omega \rightarrow \mathbb{R}^n$, onde $\Omega \subset \mathbb{R}^m$, função limitada. Definimos a norma do supremo então por*

$$\|f\|_{\text{sup},\Omega} = \sup\{\|f(\mathbf{x})\| : \mathbf{x} \in \Omega\}.$$

Portanto, uma sequência de funções limitadas (f_i) , onde $\Omega \subset \mathbb{R}^m$, converge para $f : \Omega \rightarrow \mathbb{R}^n$, se e somente se $\lim_{i \rightarrow \infty} \|f_i - f\|_{\text{sup},\Omega} = 0$. Em particular, é possível mostrar que o espaço das funções contínuas e limitadas tem propriedades interessantes (i.e., é completo) quando se usa a norma do supremo. Ver exercício 6.5.

EXEMPLO 6.4. Se $g_i : [0, 1] \rightarrow \mathbb{R}$ é tal que $g_i(x) = x^i$, $g : [0, 1] \rightarrow \mathbb{R}$ é tal que

$$(6.2.1) \quad g(x) = \begin{cases} 0 & \text{se } x \in [0, 1), \\ 1 & \text{se } x = 1, \end{cases}$$

então

$$\|g_i - g\|_{\text{sup},[0,1]} = \sup(\{x^i : x \in [0, 1)\} \cup \{0\}) = 1$$

para todo $i \in \mathbb{N}$. Logo g_i não converge uniformemente para g . Observe entretanto no Exemplo 6.2 que há convergência pontual para a função definida em (6.2.1).

EXEMPLO 6.5. Se $f_i(x) = x/i$ e $f(x) = 0$ então

$$\|f_i - f\|_{\text{sup},[0,1]} = \sup\{x/i : x \in [0, 1]\} = 1/i.$$

Logo f_i converge uniformemente para a função identicamente nula.

EXEMPLO 6.6. Suponha que $f : \mathbb{R} \rightarrow \mathbb{R}$ seja uniformemente contínua em \mathbb{R} e defina $f_i(x) = f(x + 1/i)$. Então f_i converge uniformemente para f em \mathbb{R} . De fato, seja $\epsilon > 0$. Como f é uniformemente contínua, existe $\delta \in \mathbb{R}$ tal que

$$|x - y| < \delta \implies |f(x) - f(y)| < \epsilon.$$

Seja então $N^* \in \mathbb{N}$ tal que $N^* > 1/\delta$. Logo

$$i > N^* \implies |f_i(x) - f(x)| = |f(x + 1/i) - f(x)| < \epsilon,$$

para todo $x \in \mathbb{R}$. Portanto, f_i converge uniformemente para f .

TEOREMA 6.2.4 (Critério de Cauchy para convergência uniforme). *Seja $\Omega \subset \mathbb{R}^m$ e, para $i \in \mathbb{N}$, seja $f_i : \Omega \rightarrow \mathbb{R}^n$. Então a sequência (f_i) converge uniformemente para uma função $f : \Omega \rightarrow \mathbb{R}^n$ se e somente se dado $\epsilon > 0$, existe N_0 tal que*

$$(6.2.2) \quad \|f_i(\mathbf{x}) - f_j(\mathbf{x})\| < \epsilon.$$

para todo $i, j \geq N_0$, e $\mathbf{x} \in \Omega$.

DEMONSTRAÇÃO. (\implies) Basta usar que

$$\|f_j(\mathbf{x}) - f_i(\mathbf{x})\| \leq \|f_j(\mathbf{x}) - f(\mathbf{x})\| + \|f(\mathbf{x}) - f_i(\mathbf{x})\|$$

para todo $\mathbf{x} \in \Omega$.

(\impliedby) Por hipótese, dado $\epsilon > 0$, existe K_0 tal que

$$i, j \geq K_0 \implies \|f_i(\mathbf{x}) - f_j(\mathbf{x})\| < \frac{\epsilon}{2},$$

para todo $\mathbf{x} \in \Omega$. Mas então $(f_i(\mathbf{x}))$ é sequência de Cauchy em \mathbb{R}^n , e podemos definir $f(\mathbf{x}) = \lim_{i \rightarrow +\infty} f_i(\mathbf{x})$. Falta agora mostrar a convergência uniforme de f_i para f . Dado $\mathbf{x} \in \Omega$, seja $\bar{K} \in \mathbb{N}$ tal que

$$i \geq \bar{K} \implies \|f_i(\mathbf{x}) - f(\mathbf{x})\| < \frac{\epsilon}{2}.$$

Note que K depende somente de ϵ , mas \bar{K} depende também de \mathbf{x} . Então, seja $i \geq K$, e para cada $\mathbf{x} \in \Omega$, seja $j = \sup\{K, \bar{K}\}$. Logo

$$\|f(\mathbf{x}) - f_i(\mathbf{x})\| \leq \|f(\mathbf{x}) - f_j(\mathbf{x})\| + \|f_j(\mathbf{x}) - f_i(\mathbf{x})\| < \epsilon,$$

e (f_i) converge uniformemente para f . □

Finalmente concluímos esta seção mostrando que limite uniforme de funções contínuas é também uma função contínua. Lembre-se que esta propriedade não vale em geral se a convergência é só pontual.

TEOREMA 6.2.5 (Troca de Limites e Continuidade). *Seja (f_i) sequência de funções $f_i : \Omega \rightarrow \mathbb{R}^n$ contínuas em $\Omega \subset \mathbb{R}^m$, convergindo uniformemente para $f : \Omega \rightarrow \mathbb{R}^n$. Então f é contínua em Ω .*

DEMONSTRAÇÃO. Seja $\mathbf{x}_0 \in \Omega$. Dado $\epsilon > 0$ existe $N_0 \in \mathbb{N}$ tal que $\|f(\mathbf{x}) - f_{N_0}(\mathbf{x})\| < \epsilon/3$ para todo $\mathbf{x} \in \Omega$. Como f_{N_0} é contínua em Ω , existe $\delta > 0$ tal que

$$\mathbf{x} \in B_\delta(\mathbf{x}_0) \cap \Omega \implies \|f_{N_0}(\mathbf{x}) - f_{N_0}(\mathbf{x}_0)\| < \frac{\epsilon}{3}.$$

Logo se $\mathbf{x} \in \Omega$ e $\|\mathbf{x} - \mathbf{x}_0\| < \delta$, então

$$\|f(\mathbf{x}) - f(\mathbf{x}_0)\| \leq \|f(\mathbf{x}) - f_{N_0}(\mathbf{x})\| + \|f_{N_0}(\mathbf{x}) - f_{N_0}(\mathbf{x}_0)\| + \|f_{N_0}(\mathbf{x}_0) - f(\mathbf{x}_0)\| < \epsilon.$$

Logo f é contínua. □

6.3. Equicontinuidade

Nesta seção discutiremos os conceitos de equicontinuidade e enunciaremos o Teorema de Arzelá–Ascoli. Não apresentaremos demonstrações, que podem (devem) ser conferidas em [6], por exemplo.

Seja F um conjunto de funções de $\Omega \subset \mathbb{R}^m$ em \mathbb{R}^n . Chamamos o conjunto F de *equicontínuo em $\mathbf{x}_0 \in \Omega$* , se dado $\epsilon > 0$, existe $\delta > 0$ tal que

$$\mathbf{x} \in \Omega, \quad \|\mathbf{x} - \mathbf{x}_0\| < \delta \implies \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \epsilon \text{ para toda } f \in F.$$

Se F for equicontínuo em todos os pontos de Ω , dizemos simplesmente que F é *equicontínuo*.

O conceito de equicontinuidade num ponto pode ser generalizado de forma a que a escolha de δ não dependa mais do ponto em consideração i.e., seja uniforme. Dizemos então que F é *uniformemente equicontínuo*, se dado $\epsilon > 0$, existe $\delta > 0$ tal que

$$\mathbf{x}, \mathbf{x}_0 \in \Omega, \quad \|\mathbf{x} - \mathbf{x}_0\| < \delta \implies \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \epsilon \text{ para toda } f \in F.$$

De forma semelhante, chamamos F de *simplesmente limitado* se para cada $\mathbf{x} \in \Omega$ existe c tal que $\|f(\mathbf{x})\| < c$ para todo $f \in F$. Finalmente, dizemos que F é *uniformemente limitado* se existe c tal que $\|f(\mathbf{x})\| < c$ para cada $\mathbf{x} \in \Omega$ e para todo $f \in F$.

O resultado abaixo informa que se o domínio for compacto, então equicontinuidade e equicontinuidade uniforme são equivalentes. O mesmo acontece com limitação simples e uniforme, e as funções são contínuas.

LEMA 6.3.1. Seja F conjunto de funções de um compacto $K \subset \mathbb{R}^m$ em \mathbb{R}^n . Então, F é equicontínuo se e somente se é uniformemente equicontínuo. Além disto, se F for equicontínuo, então F é simplesmente limitado se e somente se for uniformemente limitado.

Temos então o Teorema de Arzelá–Ascoli, que de alguma forma generaliza o Teorema de Bolzano–Weierstrass para sequências de funções.

TEOREMA 6.3.2 (Teorema de Arzelá–Ascoli). *Seja F conjunto infinito de funções definidas em um compacto $K \subset \mathbb{R}^m$ e tomando valores em \mathbb{R}^n . Então F é equicontínuo e simplesmente limitado se e somente se toda sequência de funções tem subsequência que converge uniformemente.*

DEMONSTRAÇÃO. Sejam $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots\} = K \cap \mathbb{Q}^m$ o conjunto dos elementos de K com coordenadas racionais, e seja (f_i) sequência em F . Então a sequência $(f_i(\mathbf{x}_1))$ é limitada e possui subsequência denotada por $(f_i^1(\mathbf{x}_1))$ convergente. Analogamente, $(f_i^1(\mathbf{x}_2))$ possui subsequência $(f_i^2(\mathbf{x}_2))$ convergente. De forma geral, dada a sequência limitada $(f_i^{j-1}(\mathbf{x}_j))$, extrai-se uma subsequência convergente $(f_i^j(\mathbf{x}_j))$. Note que $(f_i^j(\mathbf{x}_l))$ é convergente para $l = 1, \dots, j$. Seja $g_i = f_i^j$. Então, por construção, $(g_i(\mathbf{x}_l))$ é convergente para todo $l \in \mathbb{N}$. De fato, $(g_i(\mathbf{x}_l))$ é subsequência de $(f_i^l(\mathbf{x}_l))$, que converge.

Como F é equicontínuo num compacto, então é uniformemente equicontínuo. Dado então $\epsilon > 0$, seja $\delta > 0$ tal que

$$\mathbf{x}, \mathbf{x}_0 \in \Omega, \quad \|\mathbf{x} - \mathbf{x}_0\| < \delta \implies \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \frac{\epsilon}{3} \text{ para toda } f \in F.$$

Como K é compacto, então existe $N_0 \in \mathbb{N}$ tal que $\{B_\delta(\mathbf{x}_l)\}_{l=1}^{N_0}$ é cobertura aberta de K .

Como $(g_i(\mathbf{x}_l))$ converge para $l = 1, \dots, N_0$, seja $N_1 \in \mathbb{N}$ tal que

$$i, j \geq N_1 \implies \|g_i(\mathbf{x}_l) - g_j(\mathbf{x}_l)\| < \frac{\epsilon}{3}$$

para $l = 1, \dots, N_0$.

Finalmente, tomando $N = \max\{N_0, N_1\}$ e $i, j \geq N$, temos

$$\|g_i(\mathbf{x}) - g_j(\mathbf{x})\| \leq \|g_i(\mathbf{x}) - g_i(\mathbf{x}_l)\| + \|g_i(\mathbf{x}_l) - g_j(\mathbf{x}_l)\| + \|g_j(\mathbf{x}_l) - g_j(\mathbf{x})\| < \epsilon$$

para todo $\mathbf{x} \in K$, onde $l \in \{1, \dots, N_0\}$ é tal que $\mathbf{x} \in B_\delta(\mathbf{x}_l)$. O resultado segue então do critério de Cauchy para convergência uniforme (Teorema 6.2.4). \square

Como aplicação mostramos alguns detalhes do belo exemplo apresentado em [6].

EXEMPLO 6.7. Seja F o conjunto das funções $f : [-1, 1] \rightarrow [0, 1]$, contínuas e tais que $f(-1) = f(1) = 1$. Considere $I(f) = \int_{-1}^1 f(x) dx$. É possível mostrar que não existe $\bar{f} \in F$ tal que $I(\bar{f}) = \min_{f \in F} I(f)$. Considere agora

$$F_c = \{f \in F : f \text{ é de Lipschitz com constante } c\}.$$

Então F_c é simplesmente limitado e equicontínuo. Seja então $\mu_c = \inf\{I(f) : f \in F_c\}$, e para cada $i \in \mathbb{N}$ seja $f_i \in F_c$ tal que

$$\mu_c \leq I(f_i) \leq \mu_c + \frac{1}{i}.$$

Pelo Teorema de Arzelá–Ascoli, (f_i) possui subsequência (f_{i_k}) uniformemente convergente para algum f_c . Pode-se mostrar que $f_c \in F_c$, e que $I(f_c) = \min_{f \in F_c} I(f)$. Portanto o problema de minimizar $I(\cdot)$ em F_c tem solução.

6.4. Exercícios

EXERCÍCIO 6.1. Mostre que $\|\cdot\|_{\text{sup}, \Omega}$, ver Definição 6.2.3, satisfaz as propriedades de norma.

EXERCÍCIO 6.2. Seja a sequência de funções (f_i) , onde $f_i(x) = \sin(ix)/(1+ix)$. Mostre que (f_i) converge pontualmente para todo $x \in [0, +\infty)$, uniformemente em $[a, +\infty)$ para $a > 0$, mas não converge uniformemente em $[0, +\infty)$.

EXERCÍCIO 6.3. Sejam $\Omega \subset \mathbb{R}^m$ e $f_i : \Omega \rightarrow \mathbb{R}^n$ sejam funções uniformemente contínuas. Mostre que se (f_i) converge uniformemente para f , então f é uniformemente contínua.

EXERCÍCIO 6.4. Ache exemplo de sequência (f_i) de funções que converge uniformemente em $(0, 1]$, mas não em $[0, 1]$.

EXERCÍCIO 6.5. Seja $\mathcal{C}_{\text{lim}}(\Omega)$ o espaço das funções de $\Omega \subset \mathbb{R}^m$ em \mathbb{R}^n , contínuas e limitadas. Mostre que $\mathcal{C}_{\text{lim}}(\Omega)$ é completo na norma do supremo $\|\cdot\|_{\text{sup}, \Omega}$, i.e., uma sequência (f_i) em $\mathcal{C}_{\text{lim}}(\Omega)$ é de Cauchy (satisfaz (6.2.2)) se e somente se existe $f \in \mathcal{C}_{\text{lim}}(\Omega)$ tal que $\|f_i - f\|_{\text{sup}, \Omega} \rightarrow 0$.

EXERCÍCIO 6.6. Demonstre o Lema 6.3.1.

EXERCÍCIO 6.7. Mostre que se $K \subset \mathbb{R}^m$ é compacto, e (f_i) é sequência de funções em $\mathcal{C}_{\text{lim}}(\Omega)$ (ver exercício 6.5) uniformemente convergente, então $\{f_i : i \in \mathbb{N}\}$ é equicontínuo.

EXERCÍCIO 6.8. Seja $f : [0, 1] \rightarrow \mathbb{R}$ contínua, e (f_n) sequência de funções contínuas de $[0, 1]$ em \mathbb{R} . Prove ou apresente contra-exemplo para a seguinte afirmação:

Se (f_n) converge uniformemente para f em $(0, 1]$, então (f_n) converge uniformemente para f em $[0, 1]$.

EXERCÍCIO 6.9. Seja K conjunto compacto, $f : K \rightarrow \mathbb{R}$ contínua, e (f_n) sequência de funções contínuas de K em \mathbb{R} . Prove ou apresente contra-exemplo para a seguinte afirmação:

Se (f_n) converge pontualmente para f em K , então (f_n) converge uniformemente para f em K .

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 [13].

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 exemplo²

$$(A.1.1) \quad 1 + 1 = 2,$$

$$(A.1.2) \quad 1 = 2.$$

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 $(X \text{ ou } Y)$ só é falsa

¹Última Atualização: 09/01/2008

²Suponho por enquanto, que as propriedades de conjuntos e dos números reais são conhecidos

se X e Y forem *ambas* falsas. Similarmente, $(X e Y)$ só é verdadeira se X e Y forem *ambas* verdadeiras.

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).

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*. 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*.

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 $\mathbb{N} = \{1, 2, 3, \dots\}$ ³. 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 teoremas⁴? 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.

³Alguns autores utilizam o símbolo $:=$ no lugar de $=$ em definições. Esta é provavelmente uma boa idéia pouco utilizada, e eu não a seguirei.

⁴Uma 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). ⁵ *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 [15].

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.

⁵Enunciado 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.)

APPENDIX B

Uma introdução não tão formal à teoria de conjuntos

1

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 [13], e do Paul Halmos [9].

A primeira dificuldade encontrada é definir o que é um conjunto. Uma saída (questionável) é simplesmente dizer que um conjunto é uma “coleção” 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).

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 B.0.5 (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 elementos x de A tais que $P(x)$ é verdade.

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 $\{x : P(x) \text{ é verdade}\}$.

Vale aqui descrever uma situação interessante dada pelo *Paradoxo de Russel*. É natural perguntar-se o quão grande podem ser 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 B.0.5) 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:

¹Última Atualização: 22/01/2008

- (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 B.0.6 (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 Morgan*, que diz que para conjuntos E_n , onde $n \in \mathbb{N}$, temos que

$$(B.0.9) \quad \mathcal{C}(\cup_{i \in \mathbb{N}} E_n) = \cap_{i \in \mathbb{N}} \mathcal{C}(E_n), \quad \mathcal{C}(\cap_{i \in \mathbb{N}} E_n) = \cup_{i \in \mathbb{N}} \mathcal{C}(E_n).$$

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.

B.1. Exercícios

EXERCÍCIO B.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) $\mathbb{R} \not\subset \{x \in \mathbb{R} : x^2 \geq 0\}$.

EXERCÍCIO B.2. Mostre a regra de *De Morgan* dada em (B.0.9).

APPENDIX C

Listas de Exercícios

1

Lista I: Resolver problemas B.1, 1.2, 1.4, 1.6, 2.3, 2.4, 2.6, 2.11, 2.14, 2.15. (entrega dia 10/04/08)

Lista II: Resolver problemas 2.21, 2.22, 2.29, 2.30, 2.31. (entrega dia 28/04/08)

Lista III: Resolver problemas 3.3, 3.6, 3.7, 3.9. (entrega dia 8/05/08)

Lista IV: Resolver problemas 4.5, 4.9, 4.10, 4.18. (entrega dia 28/05/08)

Lista V: Resolver problemas 5.4, 5.6, 5.9, 5.17, 5.20. (entrega dia 4/06/08)

¹Última Atualização: 18/04/2008

Index

- Aberto 13
- Afirmativa 91
- Aproximações sucessivas 35, 61
- Arzelá–Ascoli, Teorema 88
- Axioma 93
 - da especificação 95
 - da união 96
- Base canônica 9
- Bola aberta 13
- Bola fechada 13
- Bijeção 2
- Bolzano–Weiertrass, Teorema 16, 30
- Cauchy, sequência 30
- Celas encaixantes 16
- Cobertura 18
- Conjunto 95
 - aberto 13
 - compacto 18
 - convexo 75
 - enumerável 3
 - fechado 14
 - finito 3
 - infinito 3
 - limitado 5, 16
 - linearmente dependente, independente 80
- Contagem diagonal 3
- Contração 35, 61
- Contração, Teorema 35
- Contradomínio 2
- Convergência
 - pontual 85
 - uniforme 86
- Cota inferior 5
 - superior 5
- Definição 93
- Demonstração 94
 - por indução 1
 - contradição 1
- Dependência linear 80
- Derivada direcional 65
- Derivada parcial 65
- Desigualdade
 - de Bernoulli 4
 - de Cauchy-Schwartz 11
- Densidade dos racionais nos reais 6
- Diferenciabilidade 55, 64
- Domínio 2
- Espaços Vetoriais 8
- Equicontinuidade 88
- Extensão contínua 54
- Fecho 21
- Função 2
 - bijetiva 2
 - biunívoca, injetiva, um a um 2
 - composta 46
 - contínua 43
 - convexa 75

- crescente, estritamente crescente 60
- decrecente, estritamente decrescente 60
- diferenciável 55, 64
- inversa 2
- limitada 48
- de Lipschitz 53
- sobre, sobrejetiva 2
- uniformemente contínua 51

- Heine–Borel, Teorema 18
- Hessiana 71, 71, 73

- Imagem 2
 - inversa 2
- Ínfimo 6
- Implicação 92
- Intervalos 7
 - encaixantes 7
- Independência linear 80

- Lei do paralelograma 20
- Lema 93
- Limite
 - inferior 39
 - superior 39
 - de sequência 23
 - de funções 44

- Matriz (semi-)definida negativa/positiva 72, 73
- Matriz hessiana 71, 71, 73
- Matriz jacobiana 66
- Máximo estrito local, local 72
- Método de Newton 74
- Mínimo estrito local, local 72

- Norma 10
 - do supremo 86
- Números naturais, inteiros, racionais 1

- Par ordenado 96
- Ponto
 - de acumulação 15
 - crítico 72, 73
 - exterior 15
 - extremo (estrito) local 58, 63, 72
 - fixo 35, 61
 - fronteira, de 15
 - interior 15
- Produto
 - interno 10
 - cartesiano 96
- Propriedade
 - do supremo dos reais 6
 - Arquimediana 6
- Prova: *ver* demonstração

- Regra de De Morgam 96

- Subespaço vetorial 32, 40
- Subsequência 28
- Supremo 6
- Sequência 23
 - de Cauchy 30
 - contrátil 34
 - limitada 26
 - monótona 38

- Teorema 93
 - aplicação aberta 80
 - Arzelá–Ascoli 88
 - Bolzano–Weiertrass 16, 30
 - celas encaixantes 17
 - contração 35
 - Heine–Borel 18
 - intervalos encaixantes 8
 - Kuhn–Tucker 81
 - Lagrange 81
 - Ponto extremo interior 58
 - Preservação de compacidade 48
 - Rolle 58
 - Taylor 62, 71
 - Valor Intermediário 50

Valor Médio	59
Teste da razão	37
Valor absoluto	7
Vizinhança aberta	15

Bibliography

- [1] M. S. Bazaraa, H. D. Sherali, C. M. Shetty, *Nonlinear programming. Theory and algorithms*, Third edition, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2006.
- [2] R. G. Bartle, *The elements of real analysis*, Second edition, John Wiley & Sons, New York, 1976.
- [3] R. G. Bartle and D. R. Sherbert, *Introduction to real analysis*, Second edition, Wiley, New York, 1992.
- [4] R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985.
- [5] S. Kevasan, *Nonlinear Functional Analysis*, Texts and readings in Mathematics **28**, Hindustan Book Agency, 2004.
- [6] E. L. Lima, *Curso de análise. Vol. 1*, Inst. Mat. Pura Apl., Rio de Janeiro, 1976.
- [7] E. L. Lima, *Curso de análise. Vol. 2*, Inst. Mat. Pura Apl., Rio de Janeiro, 1981.
- [8] E. L. Lima, *Espaços Métricos*, Inst. Mat. Pura Apl., Rio de Janeiro, 1977.
- [9] P. R. Halmos, *Naive set theory*, Reprint of the 1960 edition, Springer, New York, 1974.
- [10] D. G. Luenberger, *Introduction to linear and nonlinear programming*, Addison-Wesley, Reading, MA, 1973.
- [11] D. G. Luenberger, *Optimization by vector space methods*, John Wiley & Sons, Inc., New York-London-Sydney 1969.
- [12] W. Rudin, *Principles of mathematical analysis*, Third edition, McGraw-Hill, New York, 1976.
- [13] T. Tao, *Analysis I*, Texts and Readings in Mathematics, Hindustan Book Agency, New Delhi, 2006.
- [14] T. Tao, *Analysis II*, Texts and Readings in Mathematics, Hindustan Book Agency, New Delhi, 2006.
- [15] A. Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. of Math. (2)* **141** (1995), no. 3, 443–551.