



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS
FACULDADE DE MATEMÁTICA

ADRIANO DA SILVA CASTRO

SOBRE O TEOREMA DA APLICAÇÃO INVERSA E
O TEOREMA DA FUNÇÃO IMPLÍCITA

Belém, Pará
2023

ADRIANO DA SILVA CASTRO

SOBRE O TEOREMA DA APLICAÇÃO INVERSA E O
TEOREMA DA FUNÇÃO IMPLÍCITA

Trabalho de Conclusão de Curso, apresentado à Faculdade de Matemática do Instituto de Ciências Exatas e Naturais da Universidade Federal do Pará como requisito básico para a obtenção do grau de Licenciado em Matemática.

Orientador: Prof. Dr. Gelson Conceição Gonçalves dos Santos.

Belém, Pará

2023

ADRIANO DA SILVA CASTRO

SOBRE O TEOREMA DA APLICAÇÃO INVERSA E O
TEOREMA DA FUNÇÃO IMPLÍCITA

Trabalho de Conclusão de Curso, apresentado à
Faculdade de Matemática do Instituto de Ciências
Exatas e Naturais da Universidade Federal do Pará
como requisito básico para a obtenção do grau de
Licenciado em Matemática.

Data da Apresentação: 15 de Dezembro de 2023

Prof. Dr. Gelson Conceição Gonçalves dos Santos (Orientador)
Faculdade de Matemática, UFPA

Prof. Dr. Augusto César dos Reis Costa (Membro)
Faculdade de Matemática, UFPA

Prof. Dr. Geraldo Mendes de Araújo (Membro)
Faculdade de Matemática, UFPA

AGRADECIMENTOS

A Deus, pela vida, saúde, força e sabedoria ao longo dessa jornada, e também por ter me permitido chegar até aqui.

Aos meus pais: Maria Helena Santos da Silva e Rutnaldo Antônio Martins Castro e meus irmãos: André Felipe da Silva Castro e Antônio da Silva Castro, que sempre me deram todo o apoio necessário.

Ao meu orientador Prof. Dr. Gelson Conceição Gonçalves dos Santos, pela orientação, paciência e dedicação.

Aos professores Augusto César dos Reis Costa e Geraldo Mendes de Araújo, por aceitarem fazer parte da banca de minha defesa.

A todos os professores que me acompanharam ao longo do curso e contribuíram para minha formação, em especial: Jean Lelis, Amanda Suellen, Paulo Vilhena, João Rodrigues, Valter Borges, Irene Castro, Tânia Begazo e Marco Antônio, grandes profissionais de suas áreas que dedicam seus esforços para a melhoria do curso.

Ao meu amigo Emerson Dias Júnior, que me acompanhou ao longo desses anos de curso, e que também me indicou para a iniciação científica com o Prof. Gelson, sem ele, nada disso seria possível.

Aos meus amigos Marlon Barros, Railane Caroline, Fabíola Sodr , Felipe Macedo, Jennyfer Souza, Carolina Aiko, Alana Monteiro, Vin cius Lima, Jeanne Albuquerque, Augusto Cezar, Rafael Barbosa, Jo o Xavier, Wallace Ramos, Alice Costa, Felipe Farias, Israel Elias e Jonata do Socorro, que s o alguns dos amigos que me acompanharam nessa jornada ao longo do curso e me proporcionaram momentos incr veis e inesquec veis, me apoiaram e me ajudaram em momentos dif ceis. Certamente s o amizades que espero poder levar pelo resto de minha vida.

A Pr -Reitoria de Extens o pelo formento e oportunidade de desenvolvimento de pesquisa e a Universidade Federal do Par , pela oportunidade de estudo nesta renomada institui o.

Muito obrigado a todos!

RESUMO

Neste trabalho, estudaremos dois dos mais importantes teoremas da Análise no \mathbb{R}^n , mais precisamente, o Teorema da Aplicação Inversa e o Teorema da Função Implícita. Antes de apresentarmos formalmente os teoremas, faremos uma introdução com desenvolvimento histórico e também estudaremos definições, teoremas e lemas que nos ajudarão a entender e mostrar os dois teoremas desejados. Após isso, apresentaremos algumas aplicações destes teoremas afim de ressaltarmos suas importâncias.

Palavras-chaves: Teorema da Aplicação Inversa, Função Implícita, Análise no \mathbb{R}^n .

ABSTRACT

In this work, we will study two of the most important theorems of Analysis in \mathbb{R}^n , more precisely, the Inverse Application Theorem and the Implicit Function Theorem. Before we formally present the theorems, we will make an introduction with historical development and we will also study definitions, theorems and lemmas that will help us to understand and demonstrate the two desired theorems. After that, we will see some applications of these theorems in order to highlight their importance.

Key Words: Inverse Application Theorem, Implicit Function Theorem, Analysis in \mathbb{R}^n .

Lista de Figuras

2.1	Representação do gráfico de ξ em relação ao de f	44
3.1	Ilustração do Teorema da Aplicação Inversa.	58
3.2	Ilustração do teorema.	62
3.3	Ilustração da afirmação anterior.	63

Sumário

Introdução	10
1 Definições Preliminares	12
1.1 O Espaço \mathbb{R}^n	12
1.2 Bolas e Conjuntos Limitados	15
1.3 Sequências	16
1.4 Ponto de Acumulação	19
1.5 Continuidade	20
1.6 Continuidade Uniforme e Homeomorfismo	23
1.7 Limites	24
1.8 Conjuntos Abertos	27
1.9 Conjuntos Fechados	29
1.10 Compacidade	31
1.11 Conexidade	34
2 Diferenciabilidade	36
2.1 Diferenciabilidade para aplicações $f : \mathbb{R}^n \rightarrow \mathbb{R}$	36
2.2 Diferenciabilidade para aplicações $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$	46
2.3 Difeomorfismo	54
3 Os Teoremas da Aplicação Inversa e da Função Implícita	57
3.1 Teorema da Aplicação Inversa	57
3.2 Teorema da Função Implícita	61
A Noções de Álgebra Linear	67
A.1 Espaços Vetoriais	67
A.2 Transformações Lineares	69

A.3 Formas Bilineares	72
Referências Bibliográficas	72

Introdução

A Análise Real (ou Análise na reta) é um dos mais importantes ramos da matemática, ela se dedica a estudar propriedades do conjunto dos números reais, como ínfimo e supremo, sequências, funções de uma variável, limites, derivadas, integrais e séries de funções. Seus resultados são de extrema importância para a compreensão de conteúdos relacionados a Álgebra, Geometria e Matemática Aplicada.

A Análise no \mathbb{R}^n é uma extensão da Análise Real, ela estuda conceitos semelhantes aos da Análise real, mas agora estendido para o espaço de dimensão n , isto é, noções de conjuntos, sequências, funções de várias variáveis, derivadas parciais, superfícies, integrais múltiplas, etc...

A Análise surgiu durante o desenvolvimento e rigor matemático do Cálculo Diferencial e Integral, que foi inventado no século XVII por Isaac Newton (1643-1727) e por Gottfried Wilhelm Leibniz (1646-1716), este último desenvolveu alguns dos conceitos de derivada e integral que são utilizados até os dias de hoje. Esse desenvolvimento continuou graças a matemáticos como Leonhard Euler (1707-1783) e Jean-le-Rond d'Alembert (1717-1783). No entanto, devido as necessidades da época, não existia uma organização dos conceitos matemáticos (também chamado de “rigor matemático”), mas com o passar do tempo, surgiu a necessidade desta organização, a qual foi feita a partir do século XVIII por diversos matemáticos, alguns deles sendo Joseph Louis Lagrange (1736-1813), Bernhard Bolzano (1781-1848), Carl Friedrich Gauss (1777-1855), Augustin-Louis Cauchy (1789-1857) e Karl Weierstrass (1815-1897). Podemos dizer que a Análise Real e a Análise no \mathbb{R}^n surgiram graças a estes matemáticos, os quais possuem resultados que serão utilizados neste trabalho. Para mais informações sobre o desenvolvimento do Cálculo Diferencial e Integral, recomendamos as referências [3] e [1].

Dentre os diversos resultados importantes que compõem a Análise no \mathbb{R}^n , temos o Teorema da Aplicação Inversa e o Teorema da Função Implícita, os quais são os objetos de estudo deste trabalho.

O Teorema da Aplicação Inversa no espaço \mathbb{R}^n é uma generalização desde mesmo resultado em \mathbb{R} , este teorema afirma que se uma função $f : I \rightarrow J$, entre intervalos de \mathbb{R} , é diferenciável e bijetora, então ela é um difeomorfismo se, e somente se, $f'(x) \neq 0$ para todo $x \in I$. Já em \mathbb{R}^n , o teorema diz que dada uma função $f : U \rightarrow \mathbb{R}^n$ de classe

C^k e diferenciável, definida em um conjunto $U \subset \mathbb{R}^n$ aberto, se em um ponto $a \in U$, $f'(a) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é um isomorfismo, então localmente f é uma bijeção diferenciável com inversa f^{-1} diferenciável.

Já o Teorema da Função Implícita estabelece condições que permitem encontrar soluções de equações do tipo $f(x, y) = c$, e também permite definir y em função de x ou x em função de y , e obter a derivada da função implícita em função das derivadas da função dada $f(x, y)$. O Teorema da Função Implícita afirma que, dada uma função $f : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ de classe C^k , e o conjunto de pontos $(x, y) \in \mathbb{R}^{m+n}$ tais que $f(x, y) = c$, se uma restrição de f' é um isomorfismo, então existem abertos na vizinhança de um ponto (x_0, y_0) tal que $f(x_0, y_0) = c$, tais que existe uma aplicação $\xi \in \mathbb{R}^n$ em função de x ou em função de y tal que suas derivadas são obtidas por meio das derivadas de f e seu gráfico é obtido através do conjunto de nível c de f .

O Teorema da Aplicação Inversa e o Teorema da Função Implícita possuem algumas aplicações, eles podem ser usados no estudo de sistemas dinâmicos, estudo de curvas e superfícies diferenciáveis e implícitas, resoluções de Equações Diferenciais Parciais, etc... O que demonstra a importância e relevância destes resultados. Portanto, o objetivo deste trabalho será enunciar e demonstrar os Teoremas da Aplicação Inversa e da Função Implícita, para isso, serão desenvolvidas as definições preliminares necessárias para compreensão e demonstração destes teoremas, que serão feitas em dois capítulos.

Este trabalho está dividido da seguinte forma: No primeiro capítulo, serão abordados as primeiras definições sobre o espaço \mathbb{R}^n , sequências, noções de conjuntos, e conceitos envolvendo aplicações, como continuidade e limites. No segundo capítulo, serão abordados as definições de diferenciabilidade em \mathbb{R}^n . Finalmente, no terceiro capítulo, serão enunciados e demonstrados os dois teoremas desejados e serão abordados algumas aplicações destes teoremas. Após o terceiro capítulo, há um apêndice, que trata de algumas definições de Álgebra Linear utilizadas neste trabalho.

Para a elaboração deste trabalho, foram utilizados [4], [5], [6] e [9] como principais referências, e [7], [2], [10] e [8] como referências complementares.

Capítulo 1

Definições Preliminares

Neste primeiro capítulo, estudaremos os principais resultados referentes ao espaço \mathbb{R}^n . Faremos um breve resumo da topologia do \mathbb{R}^n , passando pelas definições de seqüências, continuidade de aplicações, homeomorfismo e limites.

1.1 O Espaço \mathbb{R}^n

Nesta primeira seção, construiremos o espaço \mathbb{R}^n e veremos suas primeiras propriedades.

Introdução

Seja n um número natural. O *espaço euclidiano n -dimensional* é o produto cartesiano de n fatores iguais a \mathbb{R} , onde \mathbb{R} é o conjunto dos números reais, logo

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}.$$

Cada ponto de \mathbb{R}^n é uma lista de n coordenadas de termos reais, representados na forma $x = (x_1, \cdots, x_n)$ onde x_1, \cdots, x_n são números reais.

Os elementos de \mathbb{R}^n serão chamados de *pontos* ou de *vetores*.

Definição 1.1. *Sejam $x = (x_1, \cdots, x_n)$ e $y = (y_1, \cdots, y_n)$ pontos de \mathbb{R}^n e α um número real, a soma $x + y$ e o produto $\alpha \cdot x$ são, respectivamente*

$$x + y = (x_1 + y_1, \cdots, x_n + y_n) \quad e \quad \alpha \cdot x = (\alpha \cdot x_1, \cdots, \alpha \cdot x_n).$$

A definição de soma e produto nos permite concluir que \mathbb{R}^n é um espaço vetorial, então, para os pontos $x, y, z \in \mathbb{R}^n$ e os números $\alpha, \beta \in \mathbb{R}$, valem as seguintes propriedades:

Comutatividade: $x + y = y + x$ e $\alpha x = x\alpha$;

Associatividade: $x + (y + z) = (x + y) + z$ e $\alpha(\beta x) = (\alpha\beta)x$;

Elemento Neutro: $x + 0 = x$ e $x \cdot 1 = x$;

Simétrico: $x + (-x) = 0$;

Distributiva: $(\alpha + \beta)x = \alpha x + \beta x$ e $\alpha(x + y) = \alpha x + \alpha y$.

Chamamos de *base canônica* do \mathbb{R}^n o conjunto dos vetores

$$e_1 = (1, 0, \dots, 0), e_2 = (0, 1, \dots, 0), \dots, e_n = (0, 0, \dots, 1)$$

que possuem uma única coordenada não nula, igual a 1. Isso significa que qualquer vetor $x = (x_1, x_2, \dots, x_n)$ pode ser escrito como

$$x = x_1 \cdot e_1 + x_2 \cdot e_2 + \dots + x_n \cdot e_n, \text{ isto é, } x = \sum_{i=1}^n x_i e_i.$$

Produto Interno

Definição 1.2. Chamamos de *produto interno do espaço euclidiano* \mathbb{R}^n a função bilinear $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, que faz corresponder a cada par de vetores $(x, y) \in \mathbb{R}^n$ o número real

$$\langle x, y \rangle = x_1 \cdot y_1 + \dots + x_n \cdot y_n.$$

Para $x, y, z \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, resulta da definição anterior, as seguintes propriedades:

- I) $\langle x, y \rangle = \langle y, x \rangle$;
- II) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$;
- III) $\langle \alpha \cdot x, y \rangle = \alpha \cdot \langle x, y \rangle = \langle x, \alpha \cdot y \rangle$;
- IV) $\langle x, x \rangle > 0$ e $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

Definição 1.3. Dados $x, y \in \mathbb{R}^n$, se $\langle x, y \rangle = 0$ então os vetores x e y são chamados *ortogonais*, representados pela notação: $x \perp y$.

Pela definição anterior, observamos que o vetor $0 = (0, \dots, 0)$ é ortogonal a todos os vetores de \mathbb{R}^n . Também o vetor e_i é ortogonal ao vetor e_j se $i \neq j$.

Exemplo 1.1. Dado o vetor $x \in \mathbb{R}^n$ não-nulo, para $y \in \mathbb{R}^n$, o vetor $z = y - \frac{\langle x, y \rangle}{\langle x, x \rangle} \cdot x$ é ortogonal a x . De fato, note que pela expressão de z que

$$\langle x, z \rangle = \langle x, y - \frac{\langle x, y \rangle}{\langle x, x \rangle} \cdot x \rangle = \langle x, y \rangle - \frac{\langle x, y \rangle}{\langle x, x \rangle} \cdot \langle x, x \rangle = \langle x, y \rangle - \langle x, y \rangle = 0.$$

Norma

Definição 1.4. Chamamos de *norma euclidiana em* \mathbb{R}^n , a função $|\cdot| : \mathbb{R}^n \rightarrow \mathbb{R}$, que associa cada $x \in \mathbb{R}^n$ o número

$$|x| = \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{\langle x, x \rangle}.$$

A norma euclidiana é o resultado da fórmula do comprimento de um vetor no plano em coordenadas cartesianas, ou seja, $|x|$ é o comprimento de x . Dessa forma,

$$|x - y| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

é o comprimento do segmento de reta de extremidades $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$, em outras palavras, é a *distância entre os pontos x e y* .

Quando $|x| = 1$, dizemos que x é um *vetor unitário*.

Teorema 1.1 (Desigualdade de Cauchy-Schwarz). *Para quaisquer $x, y \in \mathbb{R}^n$, tem-se $|\langle x, y \rangle| \leq |x| \cdot |y|$. Vale a igualdade se, e somente se, um dos vetores x, y é um múltiplo escalar do outro.*

Demonstração. Para $x = 0$ o resultado é trivial. Então seja $x \neq 0$. Pelo Exemplo 1.1, para $y \in \mathbb{R}^n$, o vetor $z = y - \frac{\langle x, y \rangle}{\langle x, x \rangle} \cdot x$ é ortogonal a x . Isolando y e fazendo $\alpha = \frac{\langle x, y \rangle}{\langle x, x \rangle}$, temos $y = z + \alpha \cdot x$. Segue daí que

$$|y|^2 = \langle z + \alpha \cdot x, z + \alpha \cdot x \rangle = |z|^2 + \alpha^2 \cdot |x|^2$$

$$|y|^2 \geq \alpha^2 \cdot |x|^2 = \frac{\langle x, y \rangle^2}{\langle x, x \rangle^2} \cdot |x|^2 = \frac{\langle x, y \rangle^2}{|x|^2}$$

$$|y|^2 \geq \frac{\langle x, y \rangle^2}{|x|^2} \Leftrightarrow |x|^2 \cdot |y|^2 \geq \langle x, y \rangle^2$$

Portanto, $|x| \cdot |y| \geq |\langle x, y \rangle|$, valendo a igualdade se $z = 0$, isto é, $y = \alpha \cdot x$. \square

Para $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, resulta da definição que a norma euclidiana tem as seguintes propriedades:

- I) $|x| = 0 \Leftrightarrow x = 0$;
- II) $|x| > 0$;
- III) $|\alpha \cdot x| = |\alpha| \cdot |x|$;
- IV) $|x + y| \leq |x| + |y|$ (Desigualdade Triangular).

As propriedades I), II) e III) resultam diretamente da definição, a propriedade IV), referindo-se a números não-negativos, é demonstrada observando que

$$|x + y|^2 = \langle x + y, x + y \rangle = |x|^2 + 2\langle x, y \rangle + |y|^2 \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2.$$

Então $|x + y|^2 \leq (|x| + |y|)^2$ e portanto, $|x + y| \leq |x| + |y|$. Isso ocorre em virtude do Teorema 1.1, que afirma que $\langle x, y \rangle \leq |x| \cdot |y|$.

Observação 1.1. *Qualquer função $|\cdot| : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ que tem as propriedades I), II), III) e IV) é chamada de norma.*

A norma euclidiana é uma forma de representação de uma norma no espaço \mathbb{R}^n e será a principal forma utilizada neste trabalho. No entanto, existem outras normas em \mathbb{R}^n , como a norma do máximo e a norma da soma, cujas respectivas expressões são:

$$|x|_M = \text{máx}\{|x_1|, \dots, |x_n|\} \quad \text{e} \quad |x|_S = |x_1| + \dots + |x_n|.$$

Ambas cumprem as quatro propriedades acima, e portanto são normas em \mathbb{R}^n . Quando necessário, essas normas serão utilizadas neste trabalho.

1.2 Bolas e Conjuntos Limitados

Nesta seção, apresentaremos as definições de bolas abertas e fechadas e a definição de conjuntos limitados.

Definição 1.5. *Sejam $a \in \mathbb{R}^n$ e um número real $r > 0$. Chamamos de bola aberta de centro a e raio r , o conjunto dos pontos de \mathbb{R}^n cuja distância destes ao centro a é menor que o raio r , ou seja,*

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

Definição 1.6. *Sejam $a \in \mathbb{R}^n$ e um número real $r > 0$. Chamamos de bola fechada de centro a e raio r , o conjunto dos pontos de \mathbb{R}^n cuja distância destes ao centro a é menor ou igual ao raio r , ou seja,*

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

Definição 1.7. *Sejam $a \in \mathbb{R}^n$ e um número real $r > 0$. Chamamos de esfera de centro a e raio r , o conjunto dos pontos de \mathbb{R}^n cuja distância destes ao centro a é igual ao raio r , ou seja,*

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

Pelas definições acima, observa-se que $B[a, r] = B(a, r) \cup S[a, r]$.

A bola fechada $B[0; 1]$ de centro na origem e raio 1 é também chamada *disco unitário* de \mathbb{R}^n . O conjunto dos vetores unitários de \mathbb{R}^n é chamado *esfera unitária de dimensão $n - 1$* , isto é,

$$S^{n-1} = \{x \in \mathbb{R}^n ; |x| = 1\}.$$

Definição 1.8. *Um conjunto $X \subset \mathbb{R}^n$ é dito limitado se está contido em alguma bola, isto é, quando dado $r > 0$, tem-se $|x| \leq r$, para todo $x \in X$.*

Definição 1.9. *Seja $X \subset \mathbb{R}^n$ um conjunto limitado. O diâmetro de X é o número*

$$\text{diam}.X = \sup\{|x - y|; x, y \in X\}.$$

Segue imediatamente dessa definição que se $\text{diam}.X = d$ e $x \in X$, então $X \subset B[x; d]$.

1.3 Sequências

Nesta seção, apresentaremos as definições de sequências de pontos em \mathbb{R}^n e também algumas propriedades, exemplos e estudaremos o conceito de limite de sequência.

Definição 1.10 (Sequência). *Uma sequência em \mathbb{R}^n é uma função $x : \mathbb{N} \rightarrow \mathbb{R}^n$ que associa a cada natural k um ponto $x_k \in \mathbb{R}^n$. Uma sequência em \mathbb{R}^n é representada por $(x_k)_{k \in \mathbb{N}}$ ou simplesmente por (x_k) .*

Para cada $i = 1, \dots, n$, x_{k_i} é a i -ésima coordenada de x_k . Como cada termo de uma sequência possui suas coordenadas, então uma sequência em \mathbb{R}^n é um lista de n sequências de números reais.

Definição 1.11. *Uma subsequência de (x_k) em \mathbb{R}^n é a restrição dessa sequência a um subconjunto infinito $\mathbb{N}' = \{k_1 < k_2 < \dots < k_i < \dots\} \subset \mathbb{N}$, denotada por $(x_{k_1}, \dots, x_{k_i}, \dots)$ ou $(x_{k_j})_{j \in \mathbb{N}}$.*

Definição 1.12. *Uma sequência é limitada quando existe uma bola que contenha todos os seus pontos. Isto é, existe $r > 0$, tal que $|x_k| \leq r$ para todo termo x_k de (x_k) .*

Resulta imediatamente da definição que se uma sequência (x_k) é limitada, então para todo $i = 1, \dots, n$, as n sequências de termos reais também são limitadas, a recíproca é verdadeira.

Definição 1.13 (Limite de Sequência). *Um ponto $a \in \mathbb{R}^n$ é chamado limite de uma sequência (x_k) quando para todo $\epsilon > 0$, existe um $k_0 \in \mathbb{N}$ tal que, $k > k_0 \Rightarrow |x_k - a| < \epsilon$ e escreve-se $\lim_{k \rightarrow \infty} x_k = a$. Em símbolos,*

$$\lim_{k \rightarrow \infty} x_k = a \Leftrightarrow \forall \epsilon > 0, \exists k_0 \in \mathbb{N}; k > k_0 \Rightarrow |x_k - a| < \epsilon$$

Dizer que $\lim_{k \rightarrow \infty} x_k = a$ é o limite de uma sequência significa dizer que qualquer bola aberta de centro a e raio ϵ contém todos os termos x_k da sequência, exceto possivelmente para um número finito de índices $k < k_0$, onde k_0 é escolhido em função de ϵ .

Para simplicidade de notação, será utilizado a notação $\lim x_k$ em vez de $\lim_{k \rightarrow \infty} x_k$.

Resulta dessa definição que

$$\lim x_k = a \Leftrightarrow |x_k - a| < \epsilon \Leftrightarrow |(x_k - a) - 0| < \epsilon \Leftrightarrow \lim(x_k - a) = 0.$$

Dessa forma, $\lim x_k = a$ se, e somente se, $\lim(x_k - a) = 0$.

Definição 1.14. *Uma sequência (x_k) é dita convergente quando existe $\lim x_k = a$. Caso não exista $\lim x_k$, então (x_k) é dita divergente.*

Resulta dessa definição que $k > k_0 \Rightarrow |x_k - a| < \epsilon$ pode ser escrito como $k > k_0 \Rightarrow x_k \in B(a; \epsilon)$. Assim, toda sequência convergente é limitada.

Teorema 1.2. *O limite de uma seqüência em \mathbb{R}^n quando existe, é unico.*

Demonstração. Considere a seqüência convergente (x_k) em \mathbb{R}^n . Sejam $a, b \in \mathbb{R}^n$ tais que $\lim x_k = a$ e $\lim x_k = b$ e suponhamos que exista $\epsilon > 0$ tal que $B(a; \epsilon)$ e $B(b; \epsilon)$ sejam disjuntos.

Como $\lim x_k = a$, então pela definição de limite, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow |x_k - a| < \epsilon$, ou seja, $x_k \in B(a; \epsilon)$, e então, $x_k \notin B(b; \epsilon)$, o que é uma contradição.

Então $a = b$. □

Teorema 1.3. *Se $\lim x_k = a$, então toda subsequência de (x_k) converge para a .*

Demonstração. Dada a seqüência (x_k) em \mathbb{R}^n , seja a subsequência $(x_{k_i})_{i \in \mathbb{N}}$. Para qualquer bola aberta B de centro em a , existe $k_0 \in \mathbb{N}$ tal que todos os termos x_k , com $k > k_0$, estão em B , em particular, todos os termos x_{k_i} , com $k_i > k_0$, estão em B , portanto, $\lim x_{k_i} = a$. □

Exemplo 1.2. *A seqüência (x_k) em \mathbb{R}^2 cujo termo geral é $x_k = \left(\frac{1}{k+1}, \operatorname{sen} \left(\frac{k\pi}{2} \right) \right)$ é divergente. Por mais que a primeira coordenada convirja para zero, ao observar a segunda coordenada, vemos que a seqüência de termos pares $x_{2k} \Rightarrow \operatorname{sen} \left(\frac{2k\pi}{2} \right) = \operatorname{sen}(k\pi)$ converge para 0, enquanto que a seqüência de termos ímpares $x_{2k-1} \Rightarrow \operatorname{sen} \left(\frac{(2k-1)\pi}{2} \right) = \operatorname{sen} \left(k\pi - \frac{\pi}{2} \right)$ converge para 1. No entanto, (x_k) é uma seqüência limitada, pois para todo $x_k \in \mathbb{R}^2$, tem-se $|x_k| \leq \sqrt{2}$.*

Teorema 1.4. *Seja uma seqüência (x_k) em \mathbb{R}^n , esta converge para um ponto $a = (a_1, \dots, a_n)$ se, e somente se, para $i = 1, \dots, n$, as suas coordenadas x_{k_i} convergem para as respectivas coordenadas a_i do ponto a .*

Demonstração. Seja (x_k) uma seqüência em \mathbb{R}^n , onde $k \in \mathbb{N}$. Adotando a norma do máximo, para cada $i = 1, \dots, n$, temos $|x_{k_i} - a_i| \leq |x_k - a|$. Pela definição de limite, temos $\lim x_k = a \Rightarrow \lim x_{k_i} = a_i$.

Reciprocamente, se vale $\lim x_{k_i} = a_i$, então dado $\epsilon > 0$, para $i = 1, \dots, n$, existem $k_1, \dots, k_n \in \mathbb{N}$ tais que $k > k_i \Rightarrow |x_{k_i} - a_i| < \epsilon$. Tomando $k_0 = \max\{k_1, \dots, k_n\}$ e adotando a norma do máximo, temos $k > k_0 \Rightarrow |x_k - a| < \epsilon$. Portanto, $\lim x_k = a$. □

Teorema 1.5 (Propriedades Operatórias). *Sejam as seqüências (x_k) e (y_k) em \mathbb{R}^n tais que $\lim(x_k) = a$ e $\lim(y_k) = b$ e a seqüência (α_k) em \mathbb{R} tal que $\lim(\alpha_k) = \alpha$, valem as seguintes propriedades:*

- I) $\lim(x_k + y_k) = a + b$;
- II) $\lim(\alpha_k \cdot x_k) = \alpha \cdot a$;
- III) $\lim\langle x_k, y_k \rangle = \langle a, b \rangle$;
- IV) $\lim|x_k| = |a|$.

Demonstração. I) Esta propriedade é válida em \mathbb{R} (veja [4], página 28), logo, para cada $i = 1, \dots, n$, vale $\lim(x_{k_i} + y_{k_i}) = a_i + b_i$ e pelo Teorema 1.4, obtemos $\lim(x_k + y_k) = a + b$.

II) Como em I), esta propriedade é válida em \mathbb{R} (veja [4], página 28), logo, para cada $i = 1, \dots, n$, vale $\lim(\alpha_k \cdot x_{k_i}) = \alpha \cdot a_i$ e pelo Teorema 1.4, obtemos $\lim(\alpha_k \cdot x_k) = \alpha \cdot a$.

III) $\lim\langle x_k, y_k \rangle = \lim(x_{k_1} \cdot y_{k_1} + \dots + x_{k_n} \cdot y_{k_n}) = a_1 \cdot b_1 + \dots + a_n \cdot b_n = \langle a, b \rangle$.

IV) $\lim|x_k| = \lim\sqrt{\langle x_k, x_k \rangle} = \sqrt{\langle a, a \rangle} = |a|$. □

Teorema 1.6 (Teorema de Bolzano-Weierstrass). *Toda sequência limitada em \mathbb{R}^n possui subsequência convergente em \mathbb{R}^n .*

Demonstração. Seja (x_k) uma sequência limitada em \mathbb{R}^n . As primeiras coordenadas de seus termos formam uma sequência limitada $(x_{k_1})_{k \in \mathbb{N}}$ de números reais a qual pelo Teorema de Bolzano-Weierstrass em \mathbb{R} (veja [4], página 26), possui uma subsequência convergente, então existe um subconjunto infinito $\mathbb{N}_1 \subset \mathbb{N}$ e um número real a_1 tais que $\lim_{k \in \mathbb{N}_1} x_{k_1} = a_1$. Novamente, pelo Teorema de Bolzano-Weierstrass em \mathbb{R} , a sequência limitada $(x_k)_{k \in \mathbb{N}_1}$ possui uma subsequência convergente, então existem um subconjunto infinito $\mathbb{N}_2 \subset \mathbb{N}_1$ e um número real a_2 tais que $\lim_{k \in \mathbb{N}_2} x_{k_2} = a_2$.

Prosseguindo com este processo n -vezes, obtemos os conjuntos infinitos $\mathbb{N} \supset \mathbb{N}_1 \supset \mathbb{N}_2 \supset \dots \supset \mathbb{N}_n$ e os números reais a_1, a_2, \dots, a_n tais que $\lim_{k \in \mathbb{N}_i} x_{k_i} = a_i$, para $i = 1, \dots, n$. Fazendo $a = (a_1, a_2, \dots, a_n)$, pelo Teorema 1.4, temos $\lim_{k \in \mathbb{N}_i} x_k = a$, o que prova o teorema. □

Definição 1.15 (Sequência de Cauchy). *Seja (x_k) uma sequência em \mathbb{R}^n . Esta é chamada de sequência de Cauchy quando para todo $\epsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que $k, r > k_0 \Rightarrow |x_k - x_r| < \epsilon$. Em limite, $\lim|x_k - x_r| = 0$.*

Teorema 1.7. *Toda sequência de Cauchy em \mathbb{R}^n é limitada.*

Demonstração. Seja (x_k) uma sequência de Cauchy. Tomando a definição de limite e fazendo $\epsilon = 1$, existe $k_0 \in \mathbb{N}$ tal que $k, r > k_0 \Rightarrow |x_k - x_r| < 1$. Dai, percebe-se que para $k > k_0$, $x_k \in B(x_{k_0}; 1)$ e então

$$\{x_1, \dots, x_k, \dots\} \subseteq \{x_1, \dots, x_{k_0-1}\} \cup B(x_{k_0}; 1).$$

Isto é, todos os termos da sequência (x_k) estão dentro da união do conjunto dos termos x_j , com $j \leq k_0 - 1$, que possivelmente não estão na bola $B(x_{k_0}, 1)$ e do conjunto dos termos x_k que pertencem a bola. Portanto, (x_k) é limitada, pois tomando $r = \max\{|x_1|, \dots, |x_{k_0-1}|, |x_{k_0}| + 1\}$, temos $|x_k| \leq r, \forall k \in \mathbb{N}$. □

Toda sequência de Cauchy é limitada, mas nem toda sequência limitada é de Cauchy. A sequência vista no Exemplo 1.2 é limitada mas não de Cauchy, pois não existe $\lim|x_k - x_r|$.

Teorema 1.8 (Critério de Cauchy). *Uma sequência converge se, e somente se, é Cauchy.*

Demonstração. Seja (x_k) uma sequência de Cauchy em \mathbb{R}^n , pela proposição anterior, (x_k) é limitada e pelo Teorema de Bolzano-Weierstrass, possui uma subsequência convergente $(x_r)_{r \in \mathbb{N}'}$. Tomando $a \in \mathbb{R}^n$ tal que $\lim x_r = a$, temos $\lim_{r \in \mathbb{N}'} |x_r - a| = 0$ e, pela definição de sequência de Cauchy, $\lim_{k \in \mathbb{N}, r \in \mathbb{N}'} |x_k - x_r| = 0$. Pela desigualdade triangular, obtemos

$$|x_k - a| \leq |x_k - x_r| + |x_r - a|$$

dessa forma, $\lim |x_k - a| = 0$, então $\lim x_k = a$ e (x_k) é convergente.

Reciprocamente, seja (x_k) convergente em \mathbb{R}^n , com $\lim x_k = a$, observando que

$$|x_k - x_r| \leq |x_k - a| + |x_r - a|,$$

segue que $\lim |x_k - x_r| = 0$ e portanto, (x_k) é de Cauchy. □

1.4 Ponto de Acumulação

Nesta seção, estudaremos a noção de ponto de acumulação, conceito fundamental para o estudo de limites mais adiante.

Definição 1.16. Um ponto $a \in \mathbb{R}^n$ é chamado de ponto de acumulação de um conjunto $X \subset \mathbb{R}^n$ quando cada bola de centro a contém algum ponto de X diferente de a .

Um ponto de acumulação de X pode pertencer a X ou não. Se a não é um ponto de acumulação de X , este é chamado *ponto isolado* de X . Isto é, existe $r > 0$ tal que $B(a, r) \cap X = \{a\}$ ou $B(a, r) \cap X = \emptyset$. Se todos os pontos de um conjunto são isolados, este é chamado conjunto *discreto*.

Definição 1.17. O conjunto dos pontos de acumulação de X é chamado de conjunto derivado de X e é representado por X' .

Teorema 1.9. Sejam $a \in \mathbb{R}^n$ e $X \subset \mathbb{R}^n$, as seguintes afirmações são equivalentes:

- a) a é um ponto de acumulação de X .
- b) a é o limite de uma sequência de pontos $x_k \in X - \{a\}$.
- c) Toda bola de centro a contém uma infinidade de pontos de X .

Demonstração. Supondo a) verdadeiro, para todo $k \in \mathbb{N}$, tem um ponto $x_k \in X$ tal que $0 < |x_k - a| < \frac{1}{k}$, donde $x_k \neq a$ e $\lim x_k = a$. Então a) \Rightarrow b).

Supondo b) verdadeiro, para qualquer $k_0 \in \mathbb{N}$ o conjunto $x_k; k > k_0$ é infinito, então b) \Rightarrow c). Finalmente, é trivial que c) \Rightarrow a). □

1.5 Continuidade

Nesta seção, apresentaremos uma das mais importantes definições em \mathbb{R}^n , a continuidade de aplicações, serão vistos: definição, propriedades, exemplos e resultados.

Seja uma aplicação $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$, que associa a cada ponto $x = (x_1, \dots, x_n) \in X$ sua imagem $f(x) = (f_1(x), \dots, f_n(x))$. As funções $f_1(x), \dots, f_n(x)$ são chamadas *funções-coordenada* de $f(x)$.

Definição 1.18 (Continuidade). *Uma aplicação $f : X \subset \mathbb{R}^m \rightarrow Y \subset \mathbb{R}^n$ é contínua no ponto $a \in X$ quando, dado $\epsilon > 0$, obtém-se $\delta > 0$ tal que*

$$x \in X, |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

A aplicação f é dita contínua quando é contínua em todos os pontos de X .

A definição acima significa que se $f(x)$ é contínua no ponto a , então podemos tomar $f(x)$ tão próximo de $f(a)$ quanto desejamos, desde que x seja suficientemente próximo de a . De forma simples, a é um ponto fixo e x se aproxima dele, então $f(x)$ se aproxima de $f(a)$.

Na notação de bolas, a definição acima é equivalente a

$$x \in B(a; \delta) \cap X \Rightarrow f(x) \in B(f(a); \epsilon) \cap Y$$

o que implica $f(B(a; \delta) \cap X) \subset B(f(a); \epsilon) \cap Y$. De modo simples, $f(B(a; \delta)) \subset B(f(a); \epsilon)$.

Da definição acima, observa-se que se $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é contínua, então para $Y \subset X$ a restrição $f|_Y$ também é contínua.

Exemplo 1.3. *Vamos mostrar que as formas bilineares $s, p : \mathbb{R}^2 \rightarrow \mathbb{R}$ definidas por $s(x, y) = x + y$ e $p(x, y) = x \cdot y$ são contínuas, ou seja, mostraremos que a soma e o produto de números reais são funções contínuas.*

Adotando a norma do máximo em \mathbb{R}^2 , para $(a, b) \in \mathbb{R}^2$, temos $(x, y) \in B((a, b); \delta)$ se, e somente se, $|x - a| < \delta$ e $|y - b| < \delta$.

Continuidade da soma: dado $\epsilon > 0$, seja $\delta = \frac{\epsilon}{2}$, se $|x - a| < \frac{\epsilon}{2}$ e $|y - b| < \frac{\epsilon}{2}$ (isto é, $(x, y) \in B((a, b); \delta)$), então

$$|s(x, y) - s(a, b)| = |x + y - (a + b)| = |(x - a) + (y - b)| \leq |x - a| + |y - b| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Portanto, $s(x, y)$ é contínua.

Continuidade do produto: dado $\epsilon > 0$, temos $xy - ab = (x - a) \cdot (y - b) + (x - a) \cdot b + a \cdot (y - b)$, então tomando δ menor que os números $\sqrt{\frac{\epsilon}{3}}$, $\frac{\epsilon}{3 \cdot (|a| + 1)}$ e $\frac{\epsilon}{3 \cdot (|b| + 1)}$ vemos que se $|x - a| < \delta$ e $|y - b| < \delta$ (isto é, $(x, y) \in B((a, b); \delta)$), então

$$|p(x, y) - p(a, b)| = |xy - ab| \leq |x - a| \cdot |y - b| + |x - a| \cdot |b| + |a| \cdot |y - b| \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Portanto, $p(x, y)$ é contínua.

Exemplo 1.4. A norma, da Definição 1.4, é uma função contínua. De fato, seja a função $|\cdot| : \mathbb{R}^n \rightarrow \mathbb{R}$ com a correspondência $x \mapsto |x|$. Tomando $f(x) = |x|$, para $a \in \mathbb{R}^n$ tem-se $|f(x) - f(a)| = ||x| - |a|| \leq |x - a|$. Logo, dado $\epsilon > 0$, fazendo $\delta = \epsilon$ tem-se $|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon$, e portanto, f é contínua.

O próximo teorema afirma que a composta de duas aplicações contínuas é também contínua.

Teorema 1.10 (Composta de Aplicações). *Sejam $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, $f : X \rightarrow \mathbb{R}^n$ com $f(X) \subset Y$ e $g : Y \rightarrow \mathbb{R}^p$. Se f é contínua no ponto $a \in X$ e g é contínua no ponto $f(a)$, então $g \circ f : X \rightarrow \mathbb{R}^p$ é contínua no ponto a .*

Demonstração. Dado $\epsilon > 0$, aplicando a definição de continuidade para g , existe $\lambda > 0$ tal que para $y \in Y$, tem-se

$$|y - f(a)| < \lambda \Rightarrow |g(y) - g(f(a))| < \epsilon.$$

Agora, aplicando a definição em f , obtemos para $\lambda > 0$ a existência de um $\delta > 0$ tal que, para $x \in X$, temos

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \lambda.$$

Assim,

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \lambda \Rightarrow |g(f(x)) - g(f(a))| < \epsilon.$$

Portanto, $g \circ f : U \rightarrow \mathbb{R}^p$ é contínua no ponto a . □

Teorema 1.11. *A aplicação $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é contínua em a se, e somente se, para toda sequência de pontos $(x_k) \subset X$ com $\lim x_k = a$, se tem $\lim f(x_k) = f(a)$.*

Demonstração. Seja $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ contínua no ponto a . Dada a sequência de pontos $(x_k) \subset X$ com $\lim x_k = a$, para todo $\epsilon > 0$ existe $\delta > 0$ tal que $f(B(a; \delta)) \subset B(f(a); \epsilon)$. Correspondente a δ , existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow x_k \in B(a; \delta)$, assim $k > k_0 \Rightarrow f(x_k) \in B(f(a); \epsilon)$ e então, $\lim f(x_k) = f(a)$.

Reciprocamente, suponhamos por contradição que $\lim x_k = a \Rightarrow \lim f(x_k) = f(a)$ mas f não seja contínua em a . Então existe $\epsilon_0 > 0$ tal que, para cada $k \in \mathbb{N}$, obtemos $x_k \in X$ com

$$|x_k - a| < \frac{1}{k} \text{ e } |f(x_k) - f(a)| \geq \epsilon_0.$$

Então temos $\lim x_k = a$, mas $\lim f(x_k) \neq f(a)$, o que gera uma contradição. □

Teorema 1.12. A aplicação $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é contínua em um ponto $a \in X$ se, e somente se, suas funções-coordenadas $f_1, \dots, f_n : X \rightarrow \mathbb{R}$ são contínuas neste ponto.

Demonstração. Suponhamos que f é contínua em a e seja uma sequência (x_k) em \mathbb{R}^n tal que $\lim x_k = a$, pelo Teorema 1.11, temos $\lim f(x_k) = f(a)$, e pelo Teorema 1.4, temos $\lim f_i(x_k) = f_i(a)$, como o Teorema 1.11 é válido em \mathbb{R} (veja [4], página 77), para $i = 1, \dots, n$, as funções-coordenada f_1, \dots, f_n são contínuas em a .

Reciprocamente, suponhamos f_1, \dots, f_n contínuas em a . Sabemos que para cada $i = 1, \dots, n$, vale $\lim f_i(x_k) = f_i(a)$, então pelo Teorema 1.4, vale $\lim f(x_k) = f(a)$ e pelo Teorema 1.11, f é contínua em a . \square

Teorema 1.13 (Propriedades Operatórias). *Sejam o conjunto $X \subset \mathbb{R}^m$, as aplicações contínuas $f, g : X \rightarrow \mathbb{R}^n$ e a aplicação contínua $\alpha : X \rightarrow \mathbb{R}$. São contínuas as seguintes aplicações:*

- I) $f + g : X \rightarrow \mathbb{R}^n$ definida por $(f + g)(x) = f(x) + g(x)$;
- II) $\alpha \cdot f : X \rightarrow \mathbb{R}^n$ definida por $(\alpha \cdot f)(x) = \alpha(x) \cdot f(x)$.
- III) $\langle f, g \rangle : X \rightarrow \mathbb{R}$ definida por $\langle f, g \rangle(x) = \langle f(x), g(x) \rangle$;
- IV) $|f| : X \rightarrow \mathbb{R}$ definida por $|f|(x) = |f(x)|$.

Demonstração. Basta aplicar o Teorema 1.11 e utilizar as propriedades operatórias de limites de sequências do Teorema 1.5. \square

Definição 1.19 (Continuidade em relação a uma variável). *Uma aplicação $f : X \subset \mathbb{R}^m \rightarrow Y \subset \mathbb{R}^n$ diz-se contínua em relação a variável x_i ($1 \leq i \leq m$) quando, para cada $(a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_m)$ fixado, a aplicação parcial $t \mapsto f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_m)$ é contínua.*

Toda aplicação contínua $f : X \subset \mathbb{R}^m \rightarrow Y \subset \mathbb{R}^n$ é contínua em relação a cada uma de suas variáveis, pois suas aplicações parciais são compostas de f com uma aplicação contínua do tipo $t \mapsto f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_m)$. No entanto, uma aplicação contínua em relação a cada uma de suas variáveis não implica que ela seja contínua. O Exemplo abaixo mostrará isso.

Exemplo 1.5. *Seja a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2}, & \text{se } (x, y) \neq (0, 0) \\ 0, & \text{se } (x, y) = (0, 0). \end{cases}$*

A função f é contínua separadamente em relação a x e a y , pois $f(a, y) = \frac{ay}{a^2 + y^2}$ (para $a \neq 0$) e $f(0, y) = 0$, enquanto $f(x, b) = \frac{xb}{x^2 + b^2}$ (para $b \neq 0$) e $f(x, 0) = 0$. Mas f é descontínua na origem $(0, 0)$. De fato, seja $g : \mathbb{R} \rightarrow \mathbb{R}^2$ definida por $g(t) = (at, at)$ então

$$(f \circ g)(t) = \frac{(at)(at)}{(at)^2 + (at)^2} = \frac{(at)^2}{2(at)^2} = \frac{1}{2}, \text{ para todo } t \neq 0,$$

e $(f \circ g)(0) = 0$. Como g é evidentemente contínua, então f é descontínua na origem.

1.6 Continuidade Uniforme e Homeomorfismo

Nesta seção, estudaremos conceitos de continuidade uniforme e homeomorfismo, tais como definições, propriedades e exemplos.

Definição 1.20 (Continuidade Uniforme). *Uma aplicação $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é dita uniformemente contínua quando, dado $\epsilon > 0$, obtém-se $\delta > 0$ tal que*

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon, \forall x, y \in X.$$

Teorema 1.14. *A aplicação $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é uniformemente contínua se, e somente se, para as sequências $x_k, y_k \in X$ tais que $\lim |x_k - y_k| = 0$, se tem $\lim |f(x_k) - f(y_k)| = 0$.*

Demonstração. Sejam f uniformemente contínua e as sequências (x_k) e (y_k) com $\lim |x_k - y_k| = 0$. Assim, para $\epsilon > 0$ dado, existe $\delta > 0$ tal que para $x, y \in X$ se tem $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$.

Além disso, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow |x_k - y_k| < \delta$, assim $k > k_0 \Rightarrow |f(x_k) - f(y_k)| < \epsilon$ e então, $\lim |f(x_k) - f(y_k)| = 0$.

Reciprocamente, suponhamos por contradição que $\lim |f(x_k) - f(y_k)| = 0$ mas f não seja uniformemente contínua. Então existe $\epsilon_0 > 0$ tal que, para todo $k \in \mathbb{N}$, temos $x_k, y_k \in X$ com $|x_k - y_k| < \frac{1}{k}$ e $|f(x_k) - f(y_k)| \geq \epsilon_0$.

Então, temos $\lim |x_k - y_k| = 0$, mas $\lim |f(x_k) - f(y_k)| \neq 0$, o que é uma contradição. \square

Resulta da definição de continuidade uniforme que toda aplicação uniformemente contínua é contínua, mas nem toda aplicação contínua é uniformemente contínua. A função produto do Exemplo 1.3 é contínua mas não uniformemente contínua, pois naquele caso, δ não dependia apenas de ϵ , mas também do ponto (a, b) .

Definição 1.21 (Aplicação Lipschitziana). *Uma aplicação $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é chamada Lipschitziana quando existe $c > 0$ (chamado constante de Lipschitz) tal que*

$$|f(x) - f(y)| \leq c|x - y|, \forall x, y \in X.$$

Teorema 1.15. *Toda aplicação Lipschitziana é uniformemente contínua.*

Demonstração. Dado $\epsilon > 0$, basta tomar $\delta = \frac{\epsilon}{c}$ e pela definição anterior, temos

$$x, y \in X \text{ e } |x - y| < \delta \Rightarrow |x - y| < \frac{\epsilon}{c} \Rightarrow |f(x) - f(y)| \leq c|x - y| < c\delta < \epsilon.$$

\square

Quando $c = 1$, a aplicação Lipschitziana é chamada de contração fraca, se $0 < c < 1$, é chamada de contração e se $c > 1$, é chamada de dilatação.

Exemplo 1.6. A soma de vetores em \mathbb{R}^n é uma contração fraca.

De fato, seja a forma bilinear $s : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $s(x, y) = x + y$. Considerando a norma da soma em $\mathbb{R}^n \times \mathbb{R}^n$, temos pela desigualdade triangular,

$$|s(x, y) - s(a, b)| = |(x + y) - (a + b)| = |(x - a) + (y - b)| \leq |x - a| + |y - b| = |(x, y) - (a, b)|.$$

Exemplo 1.7. São contrações fracas as projeções $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ definidas por $\pi_i(x) = x_i$, onde $x = (x_1, \dots, x_n)$.

Com efeito, tem-se $|\pi_i(x) - \pi_i(y)| = |x_i - y_i| \leq |x - y|$.

Teorema 1.16. Uma aplicação é uniformemente contínua se, e somente se, suas coordenadas o são.

Demonstração. Análoga a prova do Teorema 1.12. □

Teorema 1.17. A composta de duas aplicações uniformemente contínuas é uniformemente contínua.

Demonstração. Análoga a prova do Teorema 1.10. □

Definição 1.22 (Homeomorfismo). Um homeomorfismo de um conjunto $X \subset \mathbb{R}^m$ sobre um conjunto $Y \subset \mathbb{R}^n$ é uma bijeção contínua $f : X \rightarrow Y$ cuja inversa $f^{-1} : Y \rightarrow X$ é também contínua.

Resulta da definição que a composta de dois homeomorfismos é um homeomorfismo, e a inversa de um homeomorfismo é um homeomorfismo.

Exemplo 1.8. Seja uma aplicação $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$, seu gráfico é o conjunto $G = \{(x, f(x)); x \in X\} \subset \mathbb{R}^m \times \mathbb{R}^n$. Se f é contínua, então seu gráfico G é homeomorfo ao seu domínio X .

Com efeito, a aplicação contínua $\varphi : X \rightarrow G$ dada por $\varphi(x) = (x, f(x))$ é um homeomorfismo cujo inverso $(x, f(x)) \mapsto x$ é a restrição de G da projeção de $\mathbb{R}^m \times \mathbb{R}^n$ sobre \mathbb{R}^m .

Exemplo 1.9. A bola aberta $B = B(0; 1) \subset \mathbb{R}^n$ é homeomorfa ao espaço \mathbb{R}^n .

De fato, as aplicações $f : \mathbb{R}^n \rightarrow B$ e $g : B \rightarrow \mathbb{R}^n$ definidas por $f(x) = \frac{x}{1 + |x|}$ e $g(y) = \frac{y}{1 - |y|}$ são contínuas e tem-se $f(g(y)) = y$ e $g(f(x)) = x$ para quaisquer $x \in \mathbb{R}^n$ e $y \in B$. Então $g = f^{-1}$.

1.7 Limites

Nesta seção, estudaremos outro conceito importante do espaço \mathbb{R}^n , o de limite de aplicações. Estudaremos a definição, algumas propriedades, resultados e apresentaremos alguns exemplos.

Definição 1.23 (Limite de Aplicação). *Sejam a aplicação $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ e $a \in \mathbb{R}^n$ um ponto de acumulação de X (ou seja, $a \in X'$). Dizemos que $b \in \mathbb{R}^n$ é o limite de $f(x)$ quando x tende para a quando para todo $\epsilon > 0$ dado, existe $\delta > 0$ tal que $x \in X$ e $0 < |x - a| < \delta$ implicam $|f(x) - b| < \epsilon$ e escreve-se $\lim_{x \rightarrow a} f(x) = b$. Em símbolos,*

$$\lim_{x \rightarrow a} f(x) = b \Leftrightarrow \forall \epsilon > 0, \exists \delta > 0; x \in X \text{ e } 0 < |x - a| < \delta \Rightarrow |f(x) - b| < \epsilon.$$

Dizer que $\lim_{x \rightarrow a} f(x) = b$ significa dizer que pode se tomar $f(x)$ suficientemente próximo de b desde que se tome x suficientemente próximo (mas diferente) de a . A restrição $0 < |x - a|$ significa que $x \neq a$, então no limite x não assume o valor de a .

O ponto a pode ou não pertencer ao conjunto X , se sim, então $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é contínua em a se, e somente se, $\lim_{x \rightarrow a} f(x) = f(a)$. Como em limite de seqüências, o limite de uma aplicação quando existe, é único.

Exemplo 1.10. *Sejam a função $f : X \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x$, e um ponto $(a, b) \in X$. Então $\lim_{(x,y) \rightarrow (a,b)} f(x, y) = a$.*

De fato, adotando em \mathbb{R}^2 a norma do máximo, temos que $(x, y) \in B((a, b); \delta)$ se, e somente se, dado $\epsilon > 0$, existe $\delta > 0$ tal que $|x - a| < \delta$ e $|y - b| < \delta$, e pela definição de limite, vem $|f(x, y) - f(a, b)| = |x - a| < \epsilon$, e portanto, vale $\lim_{(x,y) \rightarrow (a,b)} f(x, y) = \lim_{(x,y) \rightarrow (a,b)} x = a$.

Exemplo 1.11. *A função constante tem limite igual a própria constante.*

Com efeito, seja $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $f(x) = c$ (onde $c \in \mathbb{R}$) e $a \in X$. Pela definição de limite, existe $\epsilon > 0$ tal que, para $\delta > 0$ e $x \in X$ tem-se $|x - a| < \delta \Rightarrow |f(x) - f(a)| = |c - c| < \epsilon$, e portanto, vale $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} c = c$.

Teorema 1.18. *Sejam $X \subset \mathbb{R}^m$ e $a \in X'$. Se as funções-coordenadas de $f : X \rightarrow \mathbb{R}^n$ são $f_1, \dots, f_n : X \rightarrow \mathbb{R}$, então tem-se $\lim_{x \rightarrow a} f(x) = b$ se, e somente se $\lim_{x \rightarrow a} f_i(x) = b_i$, para $i = 1, \dots, n$.*

Demonstração. Se $\lim_{x \rightarrow a} f(x) = b$ então para cada $i = 1, \dots, n$, tem-se $\lim_{x \rightarrow a} f_i(x) = b_i$, pois $|f_i(x) - b_i| \leq |f(x) - b|$.

Reciprocamente, se vale $\lim_{x \rightarrow a} f_i(x) = b_i$ para todo $i = 1, \dots, n$, então vale $\lim_{x \rightarrow a} f(x) = b$, pois $|f(x) - b| \leq \sum_{i=1}^n |f_i(x) - b_i|$. □

Teorema 1.19. *Sejam $X \subset \mathbb{R}^m$, $a \in X'$ e $f : X \rightarrow \mathbb{R}^n$. Tem-se $\lim_{x \rightarrow a} f(x) = b$ se, e somente se, para toda seqüência de pontos $(x_k) \subset X - \{a\}$ com $\lim x_k = a$, se tem $\lim_{x \rightarrow a} f(x_k) = b$.*

Demonstração. Suponhamos que $\lim_{x \rightarrow a} f(x) = b$ e que exista uma seqüência $(x_k) \subset X - \{a\}$ com $\lim x_k = a$.

Dado $\epsilon > 0$, existe $\delta > 0$ tal que $x \in X$, $0 < |x - a| < \delta \Rightarrow |f(x) - b| < \epsilon$. Existe também $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow 0 < |x_k - a| < \delta$. Como consequência, $k > k_0 \Rightarrow |f(x_k) - b| < \epsilon$, logo, vale $\lim_{x \rightarrow a} f(x_k) = b$.

Reciprocamente, suponhamos que exista uma sequência $(x_k) \subset X - \{a\}$ com $\lim x_k = a$, e isso implique $\lim_{x \rightarrow a} f(x_k) = b$, mas que não se tenha $\lim_{x \rightarrow a} f(x) = b$. Então existe $\epsilon_0 > 0$ tal que, para todo $k \in \mathbb{N}$, tenhamos $(x_k) \subset X$ tal que $0 < |x_k - a| < \frac{1}{k}$ com $|f(x_k) - b| \geq \epsilon_0$. Então teríamos $(x_k) \subset X - \{a\}$ com $\lim x_k = a$, mas $\lim_{x \rightarrow a} f(x_k) \neq b$, o que gera uma contradição. \square

Teorema 1.20 (Limite de composta). *Sejam $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, $a \in X'$, $b \in Y'$, $f : X \rightarrow Y$ uma função tal que $\lim_{x \rightarrow a} f(x) = b$ e $g : Y \rightarrow \mathbb{R}^p$ contínua em b , então $\lim_{x \rightarrow a} g(f(x)) = g(b)$.*

Demonstração. Dado $\epsilon > 0$, aplicando a definição de continuidade em g , existe $\lambda > 0$ tal que para $y = f(x) \in Y$, tem-se

$$|f(x) - b| < \lambda \Rightarrow |g(f(x)) - g(b)| < \epsilon.$$

Agora, como $\lim_{x \rightarrow a} f(x) = b$, aplicando a definição de limite em f , obtemos que, para $\lambda > 0$ existe $\delta > 0$ tal que, para $x \in X$, temos

$$|x - a| < \delta \Rightarrow |f(x) - b| < \lambda.$$

Assim,

$$|x - a| < \delta \Rightarrow |f(x) - b| < \lambda \Rightarrow |g(f(x)) - g(b)| < \epsilon.$$

Portanto, vale $\lim_{x \rightarrow a} g(f(x)) = g(b)$. \square

Teorema 1.21 (Propriedades Operatórias). *Sejam $X \subset \mathbb{R}^m$, $f, g : X \rightarrow \mathbb{R}^n$, $\alpha : X \rightarrow \mathbb{R}$ e $x_0 \in X'$. Se existem $\lim_{x \rightarrow x_0} f(x) = a$, $\lim_{x \rightarrow x_0} g(x) = b$ e $\lim_{x \rightarrow x_0} \alpha(x) = \alpha_0$. Então valem as seguintes propriedades:*

I) $\lim_{x \rightarrow x_0} (f(x) + g(x)) = a + b$;

II) $\lim_{x \rightarrow x_0} (\alpha(x) \cdot f(x)) = \alpha_0 \cdot a$;

III) $\lim_{x \rightarrow x_0} \langle f(x), g(x) \rangle = \langle a, b \rangle$;

IV) $\lim_{x \rightarrow x_0} |f(x)| = |a|$.

Demonstração. I) Considere a soma de vetores $s : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $s(x, y) = x + y$. Pelo Exemplo 1.6, s é uma aplicação contínua. Observando que $f(x) + g(x) = s(f(x), g(x))$, pelo Teorema 1.20, temos $\lim_{x \rightarrow a} [f(x) + g(x)] = \lim_{x \rightarrow a} f(x) + \lim_{x \rightarrow a} g(x) = a + b$.

As demonstrações de II), III) e IV) são análogas ao item I). \square

Teorema 1.22. *Sejam $X \subset \mathbb{R}^m$, $a \in X'$, $\alpha : X \rightarrow \mathbb{R}$ tal que $\lim_{x \rightarrow a} \alpha(x) = 0$ e $f : X \rightarrow \mathbb{R}^n$ uma função limitada. Então $\lim_{x \rightarrow a} [\alpha(x) \cdot f(x)] = 0$.*

Demonstração. Se $f(x)$ é limitada, então existe uma constante $c > 0$ tal que, $|f(x)| \leq c$, $\forall x \in X$.

Se $\lim_{x \rightarrow a} \alpha(x) = 0$, então existe $\epsilon > 0$ tal que, para $\delta > 0$ tem-se $|x - a| < \delta \Rightarrow |\alpha(x)| < \frac{\epsilon}{c}$.
Então, temos $|x - a| < \delta \Rightarrow |\alpha(x) \cdot f(x)| < \frac{\epsilon}{c} \cdot c = \epsilon$.

Portanto, $\lim_{x \rightarrow a} [\alpha(x) \cdot f(x)] = 0$. □

Exemplo 1.12. Seja $f : \mathbb{R}^2 - \{0\} \rightarrow \mathbb{R}$ definida por $f(x, y) = \frac{x^2 y}{x^2 + y^2}$. Então $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$.

De fato, temos $f(x, y) = \frac{xy}{x^2 + y^2} \cdot x$ e como $\lim_{(x,y) \rightarrow (0,0)} x = 0$, resta provar que $\frac{xy}{x^2 + y^2}$ é limitada. Basta observar que

$$\frac{xy}{x^2 + y^2} = \frac{x}{\sqrt{x^2 + y^2}} \cdot \frac{y}{\sqrt{x^2 + y^2}} = \cos(\theta) \cdot \sin(\theta),$$

onde θ é o ângulo que o vetor $z = (x, y)$ forma com o eixo das abscissas. Concluimos que $\left| \frac{xy}{x^2 + y^2} \right| \leq 1$.

Portanto, pelo Teorema 1.22, $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$.

Teorema 1.23. Sejam $X \subset \mathbb{R}^n$, $a \in X'$ e $f : X \rightarrow \mathbb{R}$ tal que existe $\lim_{x \rightarrow a} f(x)$. Se $\lim_{x \rightarrow a} f(x) > 0$, então existe $\delta > 0$ tal que, $\forall x \in X$ e $0 < |x - a| < \delta$ implicam $f(x) > 0$.

Demonstração. Seja b tal que $\lim_{x \rightarrow a} f(x) = b$. Se $b > 0$, então tomando $b = \epsilon$, existirá $\delta > 0$ tal que

$$x \in A \text{ e } 0 < |x - a| < \delta \Rightarrow b - \epsilon < f(x) < b + \epsilon \Rightarrow 0 < f(x) < 2b.$$

Portanto, se $\lim_{x \rightarrow a} f(x) = b > 0$, então $f(x) > 0$. □

Teorema 1.24. Sejam $X \subset \mathbb{R}^n$, $a \in X'$ e $f, g : X \rightarrow \mathbb{R}$. Se $f(x) \leq g(x)$ e existem os limites $\lim_{x \rightarrow a} f(x)$ e $\lim_{x \rightarrow a} g(x)$, então vale $\lim_{x \rightarrow a} f(x) \leq \lim_{x \rightarrow a} g(x)$.

Demonstração. Suponhamos por contradição que $\lim_{x \rightarrow a} f(x) > \lim_{x \rightarrow a} g(x)$, então teríamos $\lim_{x \rightarrow a} [f(x) - g(x)] > 0$ e pelo Teorema 1.23, existe $\delta > 0$ tal que, $\forall x \in X$ e $|x - a| < \delta$ implicam $f(x) - g(x) > 0 \Rightarrow f(x) > g(x)$. O que gera uma contradição. □

1.8 Conjuntos Abertos

Nesta seção, estudaremos o conceito de conjunto aberto, e veremos algumas propriedades e exemplos.

Definição 1.24. *Seja $X \subset \mathbb{R}^n$. Um ponto $x \in X$ é chamado ponto interior do conjunto X quando dado $r > 0$, a bola aberta $B(x, r)$ está contida em X , ou seja, todos os pontos suficientemente próximos de x estão em X .*

Definição 1.25. *O Conjunto $\text{int}.X$ dos pontos interiores a X chama-se interior do conjunto X . É claro que $\text{int}.X \subset X$ e quando $x \in \text{int}.X$, diz-se que X é vizinhança de x .*

O conjunto ∂X dos pontos cujas bolas de centro neles contêm pontos de X e pontos que não estão em X é chamado fronteira de X . Isto é, se $x \in \partial X$ então para $r > 0$, a bola $B(x; r)$ contém pontos de X e pontos que não estão em X , com isso, $x \in \partial X$ não significa que $x \in X$.

Definição 1.26 (Conjunto Aberto). *Um conjunto $X \subset \mathbb{R}^n$ é chamado conjunto aberto quando todos os seus pontos são pontos interiores, ou seja, quando $\text{int}X = X$.*

De outra forma, um conjunto X é um conjunto aberto se, e somente se, $X \cap \partial X = \emptyset$, ou seja, se nenhum de seus pontos pertencer a fronteira de X .

Desta definição, percebemos que $\text{int}.X$ é um conjunto aberto, enquanto ∂A não é conjunto aberto.

Exemplo 1.13. *O conjunto vazio \emptyset e o espaço \mathbb{R}^n são conjuntos abertos.*

Com efeito, um conjunto só não é aberto se contém algum ponto que não seja interior, como \emptyset não contém ponto algum, é um conjunto aberto.

É trivial que \mathbb{R}^n seja um conjunto aberto, pois para todo ponto $a \in \mathbb{R}^n$ e um real $r > 0$, tem-se $B(a; r) \subset \mathbb{R}^n$.

Exemplo 1.14. *Toda bola aberta em \mathbb{R}^n é um conjunto aberto.*

Com efeito, dada a bola $B = B(a; r)$ de centro $a \in \mathbb{R}^n$ e raio $r > 0$, seja $x \in B$. Então $|x - a| < r$, logo, $r - |x - a| > 0$, fazendo $s = r - |x - a|$, vamos provar que $B(x; s) \subset B(a; r)$.

De fato, $y \in B(x; s) \Rightarrow |y - x| < s \Rightarrow |y - x| < s = r - |x - a|$. Logo

$$|y - a| \leq |y - x| + |x - a| < r - |x - a| + |x - a| = r.$$

Concluimos que $y \in B(a; r)$. Portanto, $B(a; r)$ é um conjunto aberto.

Teorema 1.25. *São válidas as seguintes afirmações:*

- a) *Se X e Y são conjuntos abertos, então $X \cap Y$ é aberto.*
- b) *Sendo $(X_\lambda)_{\lambda \in L}$ uma família arbitrária de conjuntos abertos, então a união finita $X = \bigcup_{\lambda \in L} X_\lambda$ é aberta.*

Demonstração. a) *Seja $x \in X \cap Y$, então $x \in X$ e $x \in Y$, assim, existem $r, k > 0$ tais que $B(x; r) \subset X$ e $B(x; k) \subset Y$, então tomando $s = \min\{r, k\}$ tem-se $B(x; s) \subset X$ e*

$B(x; s) \subset Y$, logo $B(x; s) \subset X \cap Y$. Então, x é ponto interior de $X \cap Y$ e portanto, $X \cap Y$ é aberto.

b) Supondo $x \in X$, então $x \in X_\lambda$ para algum $\lambda \in L$. Como X_λ é aberto, existe $r > 0$ tal que $B(x; r) \subset X_\lambda \subset X$, assim x é ponto interior de X e como isso ocorrerá para qualquer ponto de X , conclui-se que X é aberto. \square

Teorema 1.26. *A função $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ é contínua se, e somente se, a imagem inversa $f^{-1}(Y)$ de todo conjunto aberto $Y \subset \mathbb{R}^n$ é aberto em X .*

Demonstração. Suponhamos f contínua, se Y é aberto, então para todo $x \in f^{-1}(Y)$ existe $\epsilon > 0$ tal que $B(f(x); \epsilon) \subset Y$. Pela continuidade de f , x é centro de uma bola aberta B_x tal que $f(B_x \cap X) \subset B(f(x); \epsilon) \subset Y$, logo $x \in B_x \cap X \subset f^{-1}(Y)$. Isto valendo para todo $x \in f^{-1}(Y)$, resulta que $f^{-1}(Y) \subset A_x \cap X \subset f^{-1}(Y)$, então $f^{-1}(Y) = A_x \cap X$, onde A_x é a reunião das bolas abertas B_x , $x \in f^{-1}(Y)$. Portanto, $f^{-1}(Y)$ é aberto.

Reciprocamente, suponhamos que, para Y aberto, $f^{-1}(Y)$ é aberto em X , ou seja $f^{-1}(X) = A \cap X$, com A aberto em \mathbb{R}^n . Então dados $x \in X$ e $\epsilon > 0$, tomamos $X = B(f(x); \epsilon)$ e obtemos $A \subset \mathbb{R}^n$ aberto tal que $A \cap X = f^{-1}(B(f(x); \epsilon))$. Certamente $x \in X$, então existe $\delta > 0$ tal que $B(x; \delta) \subset X$ e assim $f(B(x; \delta) \cap X) \subset B(f(x); \epsilon)$. Portanto, f é contínua em todos os pontos $x \in X$. \square

1.9 Conjuntos Fechados

Nesta seção, estudaremos o conceito de conjunto fechado, e veremos algumas propriedades e exemplos.

Definição 1.27. *Seja o ponto $a \in \mathbb{R}^n$. Este é dito ponto aderente a um conjunto $X \subset \mathbb{R}^n$ quando dado uma sequência de pontos (x_k) em X , tem-se $\lim(x_k) = a$.*

É importante observar que a ser aderente a X não significa que $a \in X$. Abaixo está um exemplo deste fato.

Exemplo 1.15. *Considere a bola $B(0; 1) \in \mathbb{R}^n$ de centro em $0 = (0, \dots, 0)$ e raio 1, o ponto $e_1 = (1, 0, \dots, 0)$ não pertence a $B(0; 1)$, mas se tomarmos uma sequência $(x_k) \in B(0; 1)$ (com $k \in \mathbb{N}$) de termo geral $x_k = \left(1 - \frac{1}{k}, 0, \dots, 0\right)$, vemos que $\lim x_k = e_1$. Então e_1 é aderente a $B(0; 1)$.*

Definição 1.28. *O conjunto dos pontos aderentes a X é chamado fecho de X e é representado por \overline{X} .*

Definição 1.29 (Conjunto Fechado). *Um conjunto $X \subset \mathbb{R}^n$ é chamado conjunto fechado quando todos os seus pontos são pontos aderentes, ou seja, quando $\overline{X} = X$.*

De outra forma, um conjunto X é dito fechado quando todo limite de uma sequência de pontos de X ainda é um ponto de X .

Exemplo 1.16. O conjunto vazio \emptyset e o espaço \mathbb{R}^n são conjuntos fechados.

De fato, um conjunto só não é fechado se contém algum ponto que não seja aderente, como o conjunto \emptyset não contém ponto algum, ele é um conjunto fechado.

É trivial que \mathbb{R}^n seja um conjunto fechado, pois para todo ponto $a \in \mathbb{R}^n$ podemos tomar uma sequência de pontos (x_k) em \mathbb{R}^n (com $k \in \mathbb{N}$) tal que $\lim x_k = a$.

Exemplo 1.17. A fronteira ∂X de um conjunto $X \subset \mathbb{R}^n$ é um conjunto fechado.

Com efeito, pela definição de fronteira, temos $a \in \partial X$ se, e somente se, a é aderente a X e $\mathbb{R}^n - X$, ou seja, $\partial X = \overline{X} \cap \overline{\mathbb{R}^n - X}$.

Então, todos os pontos de ∂X são pontos aderentes, e portanto, ∂X é um conjunto fechado.

Teorema 1.27. São válidas as seguintes afirmações:

- O ponto a é aderente ao conjunto $X \subset \mathbb{R}^n$ se, e somente se, toda bola de centro a contém algum ponto de X .
- Um conjunto $X \subset \mathbb{R}^n$ é fechado se, e somente se, seu complementar $\mathbb{R}^n - X$ é aberto.
- O fecho de qualquer conjunto é fechado, ou seja, $\overline{\overline{X}} = \overline{X}$.

Demonstração. a) Se a é aderente a X , então $\lim x_k = a$, com $x_k \in X$ para todo $k \in \mathbb{N}$. Portanto, qualquer bola $B(a; r)$ contém pontos de X , a saber, todos os x_k para k suficientemente grande.

Reciprocamente, se toda bola de centro a contém pontos de X , podemos escolher, para cada $k \in \mathbb{N}$, um ponto $x_k \in X$ que esteja na bola $B\left(a; \frac{1}{k}\right)$, isto é, $|x_k - a| < \frac{1}{k}$. Então, $\lim x_k = a$ e portanto, a é aderente a X .

b) Seja X um conjunto fechado, então se $x \in \mathbb{R}^n - X$ então x não é aderente a X e, pelo item a), existe $r > 0$ tal que $B(x; r) \subset \mathbb{R}^n - X$, então $\mathbb{R}^n - X$ é um conjunto aberto. A recíproca é provada de modo análogo.

Portanto, X é fechado se, e somente se, $\mathbb{R}^n - X$ é aberto.

c) Se $x \in \mathbb{R}^n - \overline{X}$, então pelo item a) existe uma bola aberta B de centro em x que não contém pontos de X , ou seja, $X \subset \mathbb{R}^n - B$. Logo, $\overline{X} \subset \overline{\mathbb{R}^n - B}$, mas pelo item b), $\mathbb{R}^n - B$ é fechado, então $\overline{X} \subset \mathbb{R}^n - B$, ou equivalentemente, $B \subset \mathbb{R}^n - \overline{X}$.

Assim, todo ponto $x \in \mathbb{R}^n - \overline{X}$ é um ponto interior, logo $\mathbb{R}^n - \overline{X}$ é aberto e pelo item b), \overline{X} é fechado. \square

Teorema 1.28. São válidas as seguintes afirmações:

- Se X e Y são conjuntos fechados, então $X \cup Y$ é fechado.
- Sendo $(X_\lambda)_{\lambda \in L}$ uma família arbitrária de conjuntos fechados, então a interseção finita $X = \bigcap_{\lambda \in L} X_\lambda$ é fechada.

Demonstração. a) Pelo item b) do Teorema 1.27, os conjuntos $X^c = \mathbb{R}^n - X$ e $Y^c = \mathbb{R}^n - Y$ são abertos, então pelo item a) do Teorema 1.25, temos $X^c \cap Y^c = \mathbb{R}^n - (X \cup Y)$ é aberto. Novamente pelo item b) do Teorema 1.27, $X \cup Y$ é fechado.

c) Para cada $\lambda \in L$, temos que $A_\lambda = \mathbb{R}^n - X_\lambda$ é aberto. Segue pela item b) do Teorema 1.27 que $A = \bigcup_{\lambda \in L} A_\lambda$ é aberto. Como $A = \mathbb{R}^n - X$, concluímos que X é fechado. \square

Teorema 1.29. *Uma função $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é contínua se, e somente se, a imagem inversa $f^{-1}(Y)$ de todo conjunto fechado $Y \subset \mathbb{R}^n$ é fechado em X .*

Demonstração. Se Y é fechado, então pelo item b) do Teorema 1.27, seu complementar, $A = \mathbb{R}^n - Y$ é aberto e pelo Teorema 1.26, f é contínua em A e $f^{-1}(A) = X - f^{-1}(Y)$ é aberto. Portanto, $f^{-1}(Y)$ é fechado. \square

1.10 Compacidade

Nesta seção, estudaremos o conceito de conjunto compacto, e veremos algumas propriedades e exemplos.

Definição 1.30 (Conjunto Compacto). *Um conjunto $X \subset \mathbb{R}^n$ é chamado compacto quando é limitado e fechado.*

Exemplo 1.18. *Toda bola fechada $B[a; r]$ e a esfera $S[a; r]$ são conjuntos compactos.*

Pelo Teorema 1.6, concluímos que um conjunto X é compacto se, e somente se, toda sequência de pontos $(x_k) \subset X$ possui uma subsequência que converge para algum ponto de X .

Decorre imediatamente da definição as seguintes propriedades:

- I) Se X e Y são conjuntos compactos, então $X \cup Y$ é compacto.
- II) Sendo $(X_\lambda)_{\lambda \in L}$ uma família arbitrária de conjuntos compactos, então a interseção finita $X = \bigcap_{\lambda \in L} X_\lambda$ é compacta.

Teorema 1.30. *A imagem $f(K)$ de um conjunto compacto $K \subset X$ por uma função contínua $f : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é um conjunto compacto.*

Demonstração. Seja (y_k) uma sequência de pontos em $f(K)$. Para cada $k \in \mathbb{N}$ existe $x_n \in K$ tal que $f(x_k) = y_k$. Como K é compacto, uma subsequência $(x_k)_{k \in \mathbb{N}'}$ converge para um ponto $a \in K$.

Sendo f contínua em a , e $a = \lim_{k \in \mathbb{N}'} (x_k)$, então pelo Teorema 1.11, temos $\lim_{k \in \mathbb{N}'} f(x_k) = f(a) \in f(K)$. Logo, toda sequência de pontos $(y_k) = f(x_k) \in f(K)$ possui uma subsequência $(y_k)_{k \in \mathbb{N}'}$ convergente para um ponto $a \in f(K)$. Portanto, $f(K)$ é compacto. \square

Corolário 1.30.1 (Teorema de Weierstrass). *Seja $K \subset \mathbb{R}^m$ um conjunto compacto. Se $f : K \rightarrow \mathbb{R}$ é uma função real contínua, então existem $x_0, x_1 \in K$ tais que $f(x_0) \leq f(x) \leq f(x_1), \forall x \in K$.*

Demonstração. Sendo $f(K) \subset \mathbb{R}$ compacto, então $f(K)$ é limitado, considerando os números $y_0 = \inf f(K)$ e $y_1 = \sup f(K)$, é claro que $y_0, y_1 \in f(K)$, então $y_0 = f(x_0)$ e $y_1 = f(x_1)$ para todo $x \in K$. \square

Teorema 1.31. *Seja $X \subset \mathbb{R}^m$ um conjunto compacto. Toda função contínua $f : X \rightarrow \mathbb{R}^n$ é uniformemente contínua.*

Demonstração. Suponhamos que f não seja uniformemente contínua. Então existe $\epsilon > 0$ tal que, para cada $k \in \mathbb{N}$ existem duas seqüências (x_k) e (y_k) tais que $\lim(x_k - y_k) = 0$ e $|f(x_k) - f(y_k)| \geq \epsilon$.

Como X é compacto, podemos supor que $\lim x_k = a \in X$ e como $y_k = (y_k - x_k) + x_k$, temos $\lim y_k = a$. Sendo f contínua em a , temos $\lim[f(x_k) - f(y_k)] = \lim f(x_k) - \lim f(y_k) = a - a = 0$, o que gera uma contradição. \square

Teorema 1.32. *Seja $X \subset \mathbb{R}^m$ um conjunto compacto. Toda função contínua injetiva $f : X \rightarrow \mathbb{R}^n$ é um homeomorfismo sobre sua imagem compacta $Y = f(X)$.*

Demonstração. Seja $g = f^{-1} : Y \rightarrow X$ a inversa de f , como Y é compacto, logo fechado, então pelo Teorema 1.29, g é contínua se, e somente se, para todo fechado $A \subset \mathbb{R}^m$, a imagem inversa $g^{-1}(A) = g^{-1}(A \cap Y)$ é um fechado em \mathbb{R}^n .

Mas $A \cap Y$ é compacto, então $g^{-1}(A \cap Y) = f(A \cap Y)$ é compacto pelo Teorema 1.30, logo é fechado. \square

Teorema 1.33 (Teorema de Cantor). *Sejam $K_1 \supset K_2 \supset \dots \supset K_n \supset \dots$ uma seqüência decrescente de conjuntos compactos não-vazios, existe pelo menos um ponto a que pertence a todos os esses K_k .*

Demonstração. Para cada $k \in \mathbb{N}$, seja $x_k \in K_k$, a seqüência (x_k) é limitada, então possui uma subsequência $(x_r)_{r \in \mathbb{N}'}$ convergente para o ponto $a = \lim_{r \in \mathbb{N}'} (x_r)$.

Dado k , temos $K_r \subset K_k$ sempre que $r \in \mathbb{N}'$ e $r > k$. Então quando isso ocorre, temos $x_r \in K_k$. Segue que $a = \lim_{r \in \mathbb{N}'} (x_r)$ pertence a K_k , para cada $k \in \mathbb{N}$. \square

Definição 1.31. *Uma cobertura de um conjunto $X \subset \mathbb{R}^n$ é uma família $(C_\lambda)_{\lambda \in L}$ de subconjuntos $C_\lambda \subset \mathbb{R}^n$ tais que $X \subset \bigcup_{\lambda \in L} C_\lambda$.*

Definição 1.32. *Uma subcobertura é uma subfamília $(C_\lambda)_{\lambda \in L'}$, $L' \subset L$ de subconjuntos $C_\lambda \subset \mathbb{R}^n$ tal que ainda se tem $X \subset \bigcup_{\lambda \in L'} C_\lambda$.*

Uma cobertura é dita aberta quando todos os C_λ são abertos, ou finita quando L é um conjunto finito.

Lema 1.1. *Seja $K \subset \mathbb{R}^n$ compacto. Para todo $\epsilon > 0$ existe uma decomposição $K = K_1 \cup \dots \cup K_n$ onde cada K_i , ($1 \leq i \leq n$) é compacto e tem diâmetro $\leq \epsilon$.*

Demonstração. Consideremos $\alpha > 0$. Um cubo de aresta α é um produto cartesiano $C = \prod_{i=1}^n [a_i, a_i + \alpha]$ de n intervalos de comprimento α . Se $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ pertencem a C , então para cada $i = 1, \dots, n$, tem-se $|x_i - y_i| \leq \alpha$, logo $|x - y| = \sqrt{\sum (x_i - y_i)^2} \leq \alpha\sqrt{n}$. Tomando $x_i = a_i$ e $y_i = a_i + \alpha$, temos $|x - y| = \alpha\sqrt{n}$, então $\alpha\sqrt{n}$ é o diâmetro do cubo de aresta α em \mathbb{R}^n .

A Decomposição $\mathbb{R} = \bigcup_{m \in \mathbb{Z}} [m\alpha, (m+1)\alpha]$ da reta em intervalos adjacentes de comprimento α determina uma decomposição de \mathbb{R}^n como reunião de cubos adjacentes de aresta α . A saber, para cada $m = (m_1, \dots, m_n) \in \mathbb{Z}^n$, pomos $C_m = \prod_{i=1}^n [m_i\alpha, (m_i+1)\alpha]$ e temos $\mathbb{R}^n = \bigcup_{m \in \mathbb{Z}^n} C_m$. Para todo $X \subset \mathbb{R}^n$ tem-se $X = \bigcup_{m \in \mathbb{Z}^n} (X \cap C_m)$. Se X é limitado, apenas um número finito das interseções $X \cap C_m$ são não-vazias, logo, podemos escrever

$$X = X_1 \cup \dots \cup X_k$$

onde cada X_i é da forma $X \cap C_m$, logo tem diâmetro menor ou igual a $\alpha\sqrt{n}$. Se X for compacto, então cada X_i é compacto, o que demonstra o lema. \square

Teorema 1.34 (Teorema de Borel-Lebesgue). *Um conjunto $K \subset \mathbb{R}^n$ é compacto se, e somente se, toda cobertura aberta $K \subset \cup A_\lambda$ admite subcobertura finita $K \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_n}$.*

Demonstração. Seja $K \subset \mathbb{R}^n$ compacto. Suponhamos, por contradição, que $K \subset \bigcup A_\lambda$ seja uma cobertura aberta que não admite subcobertura finita. Através do lema 1.1, podemos exprimir K como uma reunião finita de compactos de diâmetro < 1 . Pelo menos um desses K_i , suponhamos $K_1 \subset \cup A_\lambda$, não admite subcobertura finita.

Exprimindo K_1 como uma reunião finita de compactos de diâmetro $< \frac{1}{2}$, então pelo menos um desses, supondo K_2 , não pode ser coberto por um número finito de A'_λ 's.

Prosseguindo com este argumento, obtemos uma sequência decrescente de compactos $K_1 \supset K_2 \supset \dots \supset K_n \supset \dots$ com $\text{diam} K_n < \frac{1}{n}$ e que nenhum deles está contido em uma reunião finita de A'_λ 's, em particular, todos os K_n são não-vazios.

Pelo Teorema 1.33, existe $a \in \bigcap K_n$. Para algum λ , tem-se $a \in A_\lambda$. Como A_λ é aberto, tem-se $B\left(a; \frac{1}{n}\right) \subset A_\lambda$ para algum n . Sendo $a \in K_n$ e $\text{diam} K_n < \frac{1}{n}$, concluímos que $K_n \subset B\left(a; \frac{1}{n}\right) \subset A_\lambda$, o que gera uma contradição. \square

1.11 Conexidade

Nesta seção, estudaremos o conceito de conjunto conexo, e veremos algumas propriedades e exemplos.

Definição 1.33. *Uma cisão do conjunto $X \subset \mathbb{R}^n$ é uma decomposição $X = A \cup B$ onde $\overline{A} \cup B = A \cup \overline{B} = \emptyset$, ou seja, nenhum ponto de A é aderente a B e nenhum ponto de B é aderente a A .*

Se $X = A \cup B$ é uma cisão então os pontos de X que são aderentes a A , não pertencendo a B , estão em A , então $A = \overline{A} \cap X$. Da mesma forma, $B = \overline{B} \cap X$. Então A e B são conjuntos fechados, e como $A = X - B$ e $B = X - A$, concluímos que A e B são conjuntos abertos.

Se X é aberto, A e B são abertos disjuntos. Se X é fechado, A e B são fechados disjuntos. Se X é compacto, A e B são compactos disjuntos.

A Cisão $X = A \cup \emptyset$ é chamada cisão trivial. Qualquer outra cisão que não seja esta, é chamada cisão não-trivial.

Definição 1.34 (Conjunto Conexos). *Um conjunto $X \subset \mathbb{R}^n$ é dito conexo quando admite apenas a cisão trivial. Ou seja, quando $X = A \cup B$ tem-se $A = \emptyset$ ou $B = \emptyset$.*

Se um conjunto $X \subset \mathbb{R}^n$ não possui a cisão trivial, ele é chamado desconexo.

Exemplo 1.19. *O Conjunto vazio, um ponto $\{x\}$ e \mathbb{R} são conjuntos conexos, pois todos podem ser representados por meio da cisão trivial.*

Exemplo 1.20. *Um conjunto discreto $X \subset \mathbb{R}^n$ com mais de um ponto é um conjunto desconexo. De fato, todo ponto $x \in X$ é isolado, então qualquer bola de centro em x está contida em X , assim, todo ponto x é um conjunto aberto e então a reunião A de alguns desses pontos é um conjunto aberto em X . Como $A \neq X \neq \emptyset$, tomando $B = X - A$, temos $X = A \cup B$, que é uma cisão não trivial. Portanto, X é desconexo.*

Teorema 1.35. *Valem as seguintes afirmações:*

- A imagem de um conjunto conexo $X \subset \mathbb{R}^m$ por uma aplicação contínua $f : X \rightarrow \mathbb{R}^n$ é um conjunto conexo.*
- A reunião $X = \bigcup_{\lambda \in L} X_\lambda$ de uma família qualquer de conjuntos conexos é um conjunto conexo.*
- O produto cartesiano $X \times Y \subset \mathbb{R}^{m+n}$ dos conjuntos $X \subset \mathbb{R}^m$ e $Y \subset \mathbb{R}^n$ é conexo se, e somente se, X e Y são conjuntos conexos.*
- Se $X \subset \mathbb{R}^n$ e $X \subset Y \subset \overline{X}$, então Y é conexo.*

Demonstração. a) Se $f(X) = A \cup B$ é uma cisão da imagem de X , então A e B são abertos e fechados em $f(X)$, além de disjuntos. Então $f^{-1}(A)$ e $f^{-1}(B)$ também são disjuntos,

abertos e fechados em X . Logo, $X = f^{-1}(A) \cup f^{-1}(B)$ é uma cisão, que é trivial pois X é conexo por hipótese.

Mas $A = f(f^{-1}(A))$ e $B = f(f^{-1}(B))$ pois A e B estão contidos em X . Então A ou B é vazio e a cisão $f(X) = A \cup B$ é trivial. Portanto, $f(X)$ é conexo.

b) Seja a tal que $a \in X_\lambda$ para todo $\lambda \in L$. Se $X = A \cup B$ é uma cisão então a está em A ou em B . Suponhamos $a \in B$. Para todo $\lambda \in L$, $X = (A \cap X_\lambda) \cup (B \cap X_\lambda)$ é uma cisão trivial, pois X_λ é conexo, como $a \in (B \cap X_\lambda)$, segue que $A \cap X_\lambda$ é vazio. Então $A = \cup_\lambda (A \cap X_\lambda)$ é vazio e a cisão $X = A \cup B$ é trivial. Portanto, X é conexo.

c) Se $X \times Y$ é conexo, então X e Y são conexos porque são as imagens das projeções contínuas $p_1(x, y) = x$ e $p_2(x, y) = y$.

Reciprocamente, se X e Y são conexos, então seja o ponto $c = (a, b) \in X \times Y$, para cada $z = (x, y)$ o conjunto $C_z = (X \times \{b\}) \cup (\{x\} \times Y)$ é conexo pois é reunião dos conjuntos conexos $X \times \{b\}$ e $\{x\} \times Y$ que possuem o ponto (x, b) em comum. Também $c = (a, b) \in C_z$ para todo $z \in X \times Y$ e $X \times Y = \cup_\lambda C_z$, então pelo item anterior, $X \times Y$ é conexo.

d) Seja $Y = A \cup B$ uma cisão, então $X = (A \cap X) \cup (B \cap X)$ também é uma cisão. Como X é conexo, então $(A \cap X) = \emptyset$ ou $(B \cap X) = \emptyset$. Suponhamos $(A \cap X) = \emptyset$. De $X \subset Y$ e $Y = X \cup B$ resulta $X \subset B$, logo $\overline{X} \subset \overline{B}$ e então $Y \subset \overline{B}$ pois $Y \subset \overline{X}$. Logo, $A = A \cap Y \subset X \cap \overline{Y} = \emptyset$, ou seja, $A = \emptyset$. Então $Y = A \cup B$ é uma cisão trivial e portanto, Y é conexo. \square

Capítulo 2

Diferenciabilidade

Neste capítulo, utilizando resultados estabelecidos no capítulo anterior, apresentaremos os resultados referentes a diferenciabilidade no espaço \mathbb{R}^n , começando pela diferenciabilidade de aplicações $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e depois estendendo esse conceito para as aplicações $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$. Serão apresentados conceitos como derivadas parciais e direcionais, definição de diferenciabilidade, classes de diferenciabilidade, regra da cadeia, matriz jacobiana e difeomorfismo.

2.1 Diferenciabilidade para aplicações $f : \mathbb{R}^n \rightarrow \mathbb{R}$

Nesta seção, apresentaremos o conceito de diferenciabilidade para aplicações do tipo $f : \mathbb{R}^n \rightarrow \mathbb{R}$, tal como seus resultados, propriedades e alguns exemplos.

Derivadas Parciais e Direcionais

Definição 2.1 (Derivada Parcial). *Sejam $f : U \rightarrow \mathbb{R}$ definida em um conjunto aberto $U \subset \mathbb{R}^n$ e $a \in U$. A i -ésima derivada parcial de f no ponto a é o número*

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t}, \text{ para } i = 1, \dots, n,$$

caso esse limite exista, onde e_i é o i -ésimo vetor da base canônica de \mathbb{R}^n , isto é, e_i é o vetor de \mathbb{R}^n que possui 1 na coordenada i -ésima e 0 nas demais.

A existência da derivada parcial $\frac{\partial f}{\partial x_i}$ mostra que a função f é contínua no ponto a em relação a variável x_i . No entanto, a existência de todas as derivadas parciais de f em um ponto a não implica que f seja necessariamente contínua neste ponto. Abaixo está um exemplo, no qual foi tomado a função do Exemplo 1.5.

Exemplo 2.1. *Seja a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2}, & \text{se } (x, y) \neq (0, 0) \\ 0, & \text{se } (x, y) = (0, 0). \end{cases}$*

Para todo x e y em \mathbb{R}^2 temos $f(0, y) = 0$ e $f(x, 0) = 0$. Então temos também

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} = 0 \quad e \quad \frac{\partial f}{\partial y}(0, 0) = \lim_{t \rightarrow 0} \frac{f(0, t) - f(0, 0)}{t} = 0,$$

assim existem as derivadas parciais de f na origem $(0, 0)$, mas pelo Exemplo 1.5, f não é contínua em $(0, 0)$.

Definição 2.2 (Derivada Direcional). *Sejam $f : U \rightarrow \mathbb{R}$ definida em um conjunto aberto $U \subset \mathbb{R}^n$, $a \in U$ e $v \in \mathbb{R}^n$. A derivada direcional de f no ponto a e na direção v é*

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}.$$

Com esta definição, as derivadas parciais se tornam casos particulares das derivadas direcionais, ou seja, $\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial e_i}$ é a derivada direcional de f no ponto a na direção do vetor e_i .

Seja $\delta > 0$ tal que o segmento de reta $(a - \delta v, a + \delta v)$ esteja contido em U . O caminho retilíneo $\delta : (-\delta, \delta) \rightarrow U$ dado por $\lambda(t) = a + tv$ é transformado por f no caminho $f \circ \lambda : t \mapsto f(a + tv)$ no espaço \mathbb{R}^n . A derivada direcional $\frac{\partial f}{\partial v}(a)$ é o vetor-velocidade de $f \circ \lambda$ em $t = 0$.

Novamente, consideremos a função do Exemplo 1.5, esta possui as derivadas direcionais $\frac{\partial f}{\partial v}$ para $v = (\alpha, 0)$ e $v = (0, \beta)$, as quais são nulas. Entretanto, ao tomarmos $v = (\alpha, \beta)$, com $\alpha, \beta \neq 0$, temos

$$\frac{\partial f}{\partial v}(0, 0) = \lim_{t \rightarrow 0} \frac{1}{t} \cdot \frac{t\alpha \cdot t\beta}{(t\alpha)^2 + (t\beta)^2} = \lim_{t \rightarrow 0} \frac{\alpha \cdot \beta}{t(\alpha^2 + \beta^2)}.$$

Esse limite não existe, então não existe a derivada direcional de f no ponto $(0, 0)$.

Como nas derivadas parciais, a existência da derivada direcional não implica em continuidade. O exemplo abaixo mostrará isso.

Exemplo 2.2. *Seja a função $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $h(x, y) = \begin{cases} \frac{x^3 y}{x^6 + y^2}, & \text{se } (x, y) \neq (0, 0) \\ 0, & \text{se } (x, y) = (0, 0). \end{cases}$*

Como $h(x, x^3) = \frac{1}{2}$ para todo $x \neq 0$, vemos que h é descontínua em $(0, 0)$. No entanto, tomando a derivada direcional de f em $(0, 0)$ na direção do $v = (\alpha, \beta)$ temos

$$\frac{\partial h}{\partial v}(0, 0) = \lim_{t \rightarrow 0} \frac{h(t\alpha, t\beta)}{t} = \lim_{t \rightarrow 0} \frac{t^4 \alpha^3 \beta}{t^7 \alpha^6 + t^3 \beta^2} = \lim_{t \rightarrow 0} \frac{t\alpha^3 \beta}{t^4 \alpha^6 + \beta^2} = 0.$$

Então, existe a derivada direcional de h em $(0, 0)$ na direção do vetor v . Mas h é descontínua em $(0, 0)$.

Diferenciabilidade e Classes de Diferenciabilidade

Seja $f : U \rightarrow \mathbb{R}$ uma função que possui as n derivadas parciais em todos os pontos do aberto $U \subset \mathbb{R}^n$. Então ficam definidas as n funções $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \rightarrow \mathbb{R}$.

Definição 2.3 (Diferenciabilidade). *Uma função $f : U \rightarrow \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$ é chamada diferenciável em um ponto $a \in U$ quando cumpre as seguintes condições:*

I) *Existem as derivadas parciais $\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)$.*

II) *Para todo $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ tal que $a + v \in U$, tem-se*

$$f(a + v) - f(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot \alpha_i + r(v), \text{ onde } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0.$$

Quando f é diferenciável em todos os pontos de U , dizemos que f é diferenciável.

Observação 2.1. *U deve ser um conjunto aberto, pois se $a \in U$ então existe $\delta > 0$ tal que $B(a, \delta) \subset U$. Logo, para todo $v \in \mathbb{R}^n$ tal que $|v| < \delta$ temos $a + v \in U$ e portanto, $f(a + v)$ está bem definido.*

Observação 2.2. *A igualdade acima é a definição da função resto $r(h)$. Considerando as derivadas parciais $\frac{\partial f}{\partial x_i}$, a diferenciabilidade de f no ponto a tem sua essência na afirmação de que $r(v)$ é infinitésimo em relação a v , ou então, $r(v)$ tende a zero mais rapidamente que $|v|$, o que resulta $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$. De modo explícito, para todo $\epsilon > 0$ existe $\delta > 0$ tal que $0 < |v| < \delta \Rightarrow |r(v)| < \epsilon|v|$.*

Observação 2.3. *Em alguns casos, para evitar as exceções causadas pelo denominador zero, é conveniente colocar o resto na forma $r(v) = \rho(v) \cdot |v|$, onde ρ é definida, para todo v tal que $a + v \in U$, por $\rho(v) = \frac{r(v)}{|v|}$ se $v \neq 0$ e $\rho(0) = 0$. Então, a diferenciabilidade de f no ponto a é representada por*

$$f(a + v) - f(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot \alpha_i + \rho(v) \cdot |v|, \text{ onde } \lim_{v \rightarrow 0} \rho(v) = 0.$$

Definição 2.4 (Função de Classe C^1). *Seja $f : U \rightarrow \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^n$, dizemos que f é uma função de classe C^1 quando suas derivadas parciais $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \rightarrow \mathbb{R}$ são contínuas.*

Teorema 2.1. *Dado um aberto $U \subset \mathbb{R}^n$. Toda função $f : U \rightarrow \mathbb{R}$ de classe C^1 é diferenciável.*

Demonstração. Sejam $c = (c_1, \dots, c_n) \in U$ e $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ tais que $c + v \in B \subset U$, onde B é uma bola de centro c . Tomando a função $r(v)$, podemos escrever

$$r(v) = r(v_1, \dots, v_n) = f(c_1 + v_1, \dots, c_n + v_n) - f(c_1, \dots, c_n) - \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot v_i$$

onde as derivadas são calculadas no ponto $c = (c_1, \dots, c_n)$. Podemos escrever

$$\begin{aligned} r(v) &= f(c_1 + v_1, \dots, c_n + v_n) - f(c_1, c_2 + v_2, \dots, c_n + v_n) \\ &\quad + f(c_1, c_2 + v_2, \dots, c_n + v_n) - f(c_1, \dots, c_n) - \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot v_i, \end{aligned}$$

e pelo Teorema do Valor Médio em \mathbb{R} (veja [4], página 98), existem $\theta_1, \dots, \theta_n \in (0, 1)$ tais que

$$\begin{aligned} r(v) &= \frac{\partial f}{\partial x_1}(c_1 + \theta_1 v_1, c_2 + v_2, \dots, c_n + v_n) \cdot v_1 \\ &\quad + \frac{\partial f}{\partial x_2}(c_1, c_2 + \theta_2 v_2, \dots, c_n + v_n) \cdot v_2 \\ &\quad \vdots \\ &\quad + \frac{\partial f}{\partial x_n}(c_1, c_2 + v_2, \dots, c_n + \theta_n v_n) \cdot v_n - \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot v_i \end{aligned}$$

logo,

$$\begin{aligned} r(v) &= \left[\frac{\partial f}{\partial x_1}(c_1 + \theta_1 v_1, c_2 + v_2, \dots, c_n + v_n) \cdot v_1 - \frac{\partial f}{\partial x_1}(c_1, \dots, c_n) \right] \\ &\quad + \left[\frac{\partial f}{\partial x_2}(c_1, c_2 + \theta_2 v_2, \dots, c_n + v_n) \cdot v_2 - \frac{\partial f}{\partial x_2}(c_1, \dots, c_n) \right] \\ &\quad \vdots \\ &\quad + \left[\frac{\partial f}{\partial x_n}(c_1, c_2 + v_2, \dots, c_n + \theta_n v_n) \cdot v_n - \frac{\partial f}{\partial x_n}(c_1, \dots, c_n) \right], \end{aligned}$$

e dividindo a expressão por $|v| = \sqrt{v_1^2 + \dots + v_n^2}$, obtemos

$$\begin{aligned} \frac{r(v)}{|v|} &= \left[\frac{\partial f}{\partial x_1}(c_1 + \theta_1 v_1, c_2 + v_2, \dots, c_n + v_n) \cdot v_1 - \frac{\partial f}{\partial x_1}(c_1, \dots, c_n) \right] \frac{v_1}{\sqrt{v_1^2 + \dots + v_n^2}} \\ &\quad + \left[\frac{\partial f}{\partial x_2}(c_1, c_2 + \theta_2 v_2, \dots, c_n + v_n) \cdot v_2 - \frac{\partial f}{\partial x_2}(c_1, \dots, c_n) \right] \frac{v_2}{\sqrt{v_1^2 + \dots + v_n^2}} \\ &\quad \vdots \\ &\quad + \left[\frac{\partial f}{\partial x_n}(c_1, c_2 + v_2, \dots, c_n + \theta_n v_n) \cdot v_n - \frac{\partial f}{\partial x_n}(c_1, \dots, c_n) \right] \frac{v_n}{\sqrt{v_1^2 + \dots + v_n^2}}. \end{aligned}$$

Ao tomar o limite com $v \rightarrow 0$, os termos dentro dos colchetes acima tendem a 0, devido a continuidade das derivadas $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$. Além disso, os termos fora dos colchetes têm valor absoluto ≤ 1 . Então, pelo Teorema 1.22, $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$ e portanto, f é diferenciável. \square

Seja $f : U \rightarrow \mathbb{R}$ uma função que possui as derivadas parciais $\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_1}(x)$ em todo ponto x do aberto $U \subset \mathbb{R}^n$. A j -ésima derivada parcial da função $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$ no ponto $x \in U$ será indicada por

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (x), \text{ onde } i, j = 1, \dots, n.$$

Se essas derivadas parciais de segunda ordem existirem em cada ponto $x \in U$, teremos n^2 funções $\frac{\partial^2 f}{\partial x_j \partial x_i} : U \rightarrow \mathbb{R}$. Quando tais funções forem contínuas, f é chamada função de classe C^2 .

Prosseguindo com este argumento k vezes ($k \geq 1$), a derivada de ordem k da função f é representada como

$$\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}}(x), \text{ onde } i = 1, \dots, n.$$

Se essas derivadas parciais de ordem k existirem em cada ponto $x \in U$, teremos n^k funções $\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}} : U \rightarrow \mathbb{R}$. Quando tais funções forem contínuas, f é chamada função de classe C^k .

Definição 2.5. *Seja $f : U \rightarrow \mathbb{R}$ definida no aberto $U \subset \mathbb{R}^n$ e diferenciável no ponto $a \in U$, o vetor*

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)$$

é chamado de gradiente de f em a .

Definição 2.6. *Seja $f : U \rightarrow \mathbb{R}$ uma função diferenciável definida no aberto $U \subset \mathbb{R}^n$, um ponto $a \in U$ é chamado de ponto crítico quando este é tal que $\nabla f(a) = 0$.*

Teorema 2.2 (Regra da Cadeia). *Sejam os conjuntos abertos $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$, uma aplicação $f : U \rightarrow \mathbb{R}^n$ cujas funções-coordenada f_1, \dots, f_n possuem derivadas parciais no ponto $a \in U$, e uma função $g : V \rightarrow \mathbb{R}$ diferenciável no ponto $b = f(a)$. Então $g \circ f : U \rightarrow \mathbb{R}$ possui derivadas parciais no ponto a e vale*

$$\frac{\partial (g \circ f)}{\partial x_i} = \sum_{k=1}^n \frac{\partial g}{\partial y_k} \cdot \frac{\partial f_k}{\partial x_i}, \quad i = 1, \dots, m,$$

onde as derivadas parciais relativas aos x_i são calculadas no ponto a e as relativas a y_k são calculadas no ponto $b = f(a)$. Além disso, se f e g são de classe C^k , então $g \circ f$ também é de classe C^k .

Demonstração. Podemos escrever

$$g(f(a + te_i)) - g(f(a)) = \sum_{k=1}^n \frac{\partial g}{\partial y_k} \cdot [f_k(a + te_i) - f_k(a)] + \rho(t) \cdot |f(a + te_i) - f(a)|$$

onde, por simplicidade, escrevemos $\rho(t)$ em vez de $\rho(v)$ com $v = f(a + te_i) - f(a)$. A diferenciabilidade de g nos dá $\lim_{t \rightarrow 0} \rho(t) = 0$. Então

$$\frac{g(f(a + te_i)) - g(f(a))}{t} = \sum_{k=1}^n \frac{\partial g}{\partial y_k} \cdot \frac{f_k(a + te_i) - f_k(a)}{t} \pm \rho(t) \left| \frac{f(a + te_i) - f(a)}{t} \right|.$$

e como

$$\lim_{t \rightarrow 0} \rho(t) = 0 \text{ e } \lim_{t \rightarrow 0} \left| \frac{f(a + te_i) - f(a)}{t} \right| = \left| \frac{\partial f}{\partial x_i}(a) \right|,$$

temos

$$\frac{\partial(g \circ f)}{\partial x_i} = \lim_{t \rightarrow 0} \frac{g(f(a + te_i)) - g(f(a))}{t} = \sum_{k=1}^n \frac{\partial g}{\partial y_k} \cdot \frac{\partial f_k}{\partial x_i}.$$

Além disso, o fato de que $g \circ f \in C^k$ decorre da expressão de $\frac{\partial(g \circ f)}{\partial x_i}$ em termos das derivadas parciais de g e das f_k , que são contínuas. \square

Corolário 2.2.1. *Seja $f : U \rightarrow \mathbb{R}$ diferenciável no aberto $U \subset \mathbb{R}^n$, com $a \in U$. Dado o vetor $v = (\alpha_1, \dots, \alpha_n)$, se $\lambda : (-\delta, \delta) \rightarrow U$ é qualquer caminho diferenciável tal que $\lambda(0) = a$ e $\lambda'(0) = v$, tem-se*

$$(f \circ \lambda)'(0) = \langle \nabla f(a), v \rangle = \frac{\partial f}{\partial v}(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot \alpha_i.$$

Demonstração. Pela regra da cadeia, vem

$$(f \circ \lambda)' = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot \frac{d\lambda_i}{dt},$$

observando que, para $\lambda(t) = (\lambda_1(t), \dots, \lambda_n(t))$, tem-se $\alpha_i = \frac{d\lambda_i}{dt}(0)$. Note ainda que $\frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0)$ com $\lambda(t) = a + tv$, pois $\lambda'(0) = v$. \square

Teorema 2.3 (Teorema do Valor Médio). *Seja $f : U \rightarrow \mathbb{R}$ diferenciável no aberto $U \subset \mathbb{R}^n$. Se o segmento de reta $[a, a + v]$ está contido em U . Então existe $\theta \in (0, 1)$ tal que*

$$f(a + v) - f(a) = \frac{\partial f}{\partial v}(a + \theta v) = \langle \nabla f(a + \theta v), v \rangle = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a + \theta v) \cdot \alpha_i.$$

Demonstração. Considerando o caminho retilíneo $\lambda : [0, 1] \rightarrow U$, dado por $\lambda(t) = a + tv$, vemos que $f(a + v) - f(a) = (f \circ \lambda)(1) - (f \circ \lambda)(0)$. Pelo Teorema do Valor Médio em \mathbb{R} (veja [4], página 98), existe $\theta \in (0, 1)$ tal que $(f \circ \lambda)(1) - (f \circ \lambda)(0) = (f \circ \lambda)'(\theta)$, e pela regra da cadeia,

$$(f \circ \lambda)'(\theta) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a + \theta v) \cdot \alpha_i = \frac{\partial f}{\partial v}(a + \theta v) = \langle \nabla f(a + \theta v), v \rangle.$$

□

Corolário 2.3.1. *Seja $U \subset \mathbb{R}^n$ aberto e conexo. Se $f : U \rightarrow \mathbb{R}$ é tal que $\nabla f(x) = 0$ para todo $x \in U$, então f é constante.*

Demonstração. Sejam $a, b \in U$, a existência de $\nabla f(x)$ garante a continuidade de $f|_{[a, b]}$ para todo segmento de reta $[a, b]$ contido em U (veja [4], página 91). Resulta do Teorema do Valor Médio que $[a, b] \subset U$ implica $f(b) = f(a)$.

Devido a conexidade de U , qualquer ponto $x \in U$ pode ser ligado ao ponto a por uma poligonal contida em U , com vértices $a = a_0, a_1, \dots, a_k = x$. Então, temos $f(a) = f(a_1) = \dots = f(x)$. Logo, $f(x) = f(a)$ para todo $x \in U$ e portanto, f é constante. □

Definição 2.7. *Dada $f : U \rightarrow \mathbb{R}$ de classe C^1 , o conjunto $f^{-1}(c) = \{x \in U; f(x) = c\}$ é chamado de conjunto de nível c da função f .*

Agora, iremos enunciar e demonstrar um caso particular do Teorema da Função Implícita, que estabelece condições para que o conjunto de pontos $(x, y) \in \mathbb{R}^{n+1}$ que satisfazem $f(x, y) = c$ ($c \in \mathbb{R}$), sejam na vizinhança de um ponto (x_0, y_0) tal que $f(x_0, y_0) = c$, o gráfico de uma função ξ de variável x ou de variável y . O teorema diz que se f é de classe C^k ($k \geq 1$) e $\frac{\partial f}{\partial y} \neq 0$, ocorre o primeiro caso, e se $\frac{\partial f}{\partial x} \neq 0$, ocorre o segundo caso.

Nesse caso, os pontos do espaço \mathbb{R}^{n+1} serão representados da forma (x, y) , onde $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ e $y \in \mathbb{R}$.

Teorema 2.4 (Teorema da Função Implícita). *Sejam $f : U \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ uma função de classe C^k com $k \geq 1$, U um conjunto aberto e $(x_0, y_0) \in \mathbb{R}^{n+1}$ tal que $f(x_0, y_0) = c$ e $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Então, existem uma bola $B = B(x_0; \delta) \subset \mathbb{R}^n$ e um intervalo $J = (y_0 - \epsilon, y_0 + \epsilon) \subset \mathbb{R}$ tais que*

I) $B \times \bar{J} \subset U$ e $\frac{\partial f}{\partial y}(x, y) \neq 0$ para todo $(x, y) \in B \times \bar{J}$;

II) Para todo $x \in B$ existe um único $y = \xi(x) \in J$ tal que $f(x, y) = f(x, \xi(x)) = c$. Além disso, $\xi : B \rightarrow J$ é uma função de classe C^k , dada implicitamente, que tem derivadas parciais em cada $x \in B$ dadas por

$$\frac{\partial \xi}{\partial x_i}(x) = -\frac{\frac{\partial f}{\partial x_i}(x, \xi(x))}{\frac{\partial f}{\partial y}(x, \xi(x))}.$$

Demonstração. I) Vamos supor $\frac{\partial f}{\partial y}(x_0, y_0) > 0$. Pela continuidade de $\frac{\partial f}{\partial y}$, existem $\delta > 0$ e $\epsilon > 0$ tais que, tomando $B = B(x_0, \delta) \subset \mathbb{R}^n$ e $J = (y_0 - \epsilon, y_0 + \epsilon) \subset \mathbb{R}$, obtemos $B \times \bar{J} \subset U$

e $\frac{\partial f}{\partial y}(x, y) > 0$ para todo $(x, y) \in B \times \bar{J}$, o que prova o item I). O Caso $\frac{\partial f}{\partial y}(x_0, y_0) < 0$ é análogo.

II) **Existência e Unicidade de $\xi(x)$:** Pelo item I), para todo $x \in B$, a função $y \mapsto f(x, y)$ é crescente no intervalo \bar{J} . Como $f(x_0, y_0) = c$, segue que $f(x_0, y_0 - \epsilon) < c < f(x_0, y_0 + \epsilon)$.

Assim, sendo f contínua, podemos supor δ tão pequeno que $f(x, y_0 - \epsilon) < c < f(x, y_0 + \epsilon)$, para todo $x \in B$.

Então, pelo Teorema do Valor Intermediário (veja [4], página 78), para cada $x \in B$, deve existir um único $y = \xi(x) \in \bar{J}$ tal que $f(x, y) = c$. Temos $y \in J$, pois como f é contínua e crescente, ocorre $f(x, y_0 - \epsilon) < c < f(x, y_0 + \epsilon)$, o que implica $y_0 - \epsilon < y < y_0 + \epsilon$, fornecendo $y \in J$.

Continuidade de ξ : Sejam $F \subset \bar{J}$ um conjunto fechado e $(x_k) \subset B$ uma sequência tal que $\xi(x_k) \in F$ para todo $k \in \mathbb{N}$ e $\lim x_k = \bar{x} \in B$, então $\xi(\bar{x}) \in F$. Mas F é compacto, então (x_k) é limitada e, pelo Teorema de Bolzano-Weierstrass, possui uma subsequência $(x'_k) \subset B$ convergente, tal que $\lim \xi(x'_k) = a \in B$.

Logo, pela continuidade de f , temos $f(\bar{x}, a) = \lim f(x'_k, \xi(x'_k)) = 0$ mas como $f(\bar{x}, \xi(\bar{x})) = 0$, segue pela unicidade de ξ que $\xi(\bar{x}) = a \in F$.

Portanto, provamos que para todo conjunto fechado $F \subset \bar{J}$, a imagem inversa $\xi^{-1}(F)$ é fechada em B . Assim, ξ é contínua.

Derivadas Parciais de ξ : Tomemos $k = k(t) = \xi(x + te_i) - \xi(x)$, então $\xi(x + te_i) = \xi(x) + k$ logo,

$$f(x + te_i, \xi(x + te_i)) = f(x + te_i, \xi(x) + k) = f(x, \xi(x)) = 0.$$

Pelo Teorema 2.3, para todo t existe $\theta = \theta(t) \in (0, 1)$ tal que

$$f(x + te_i, \xi(x) + k) - f(x, \xi(x)) = 0,$$

assim

$$\frac{\partial f}{\partial x_i}(x + \theta te_i, \xi(x) + \theta k) \cdot t + \frac{\partial f}{\partial y}(x + \theta te_i, \xi(x) + \theta k) \cdot k = 0,$$

onde

$$k = -\frac{\frac{\partial f}{\partial x_i}(x + \theta te_i, \xi(x) + \theta k)}{\frac{\partial f}{\partial y}(x + \theta te_i, \xi(x) + \theta k)} \cdot t$$

Então, lembrando que $k = \xi(x + te_i) - \xi(x)$ e dividindo a expressão por t , obtemos

$$\frac{\xi(x + te_i) - \xi(x)}{t} = \frac{k}{t} = -\frac{\frac{\partial f}{\partial x_i}(x + \theta te_i, \xi(x) + \theta k)}{\frac{\partial f}{\partial y}(x + \theta te_i, \xi(x) + \theta k)}.$$

Como ξ é contínua, obtemos $\lim_{t \rightarrow 0} k(t) = 0$. A continuidade das derivadas parciais de f mostra que as derivadas parciais de ξ em todo $x \in B$ são dadas por,

$$\frac{\partial \xi}{\partial x_i}(x) = \lim_{t \rightarrow 0} \frac{\xi(x + te_i) - \xi(x)}{t} = -\frac{\frac{\partial f}{\partial x_i}(x, \xi(x))}{\frac{\partial f}{\partial y}(x, \xi(x))}, \text{ para } 1 \leq i \leq n.$$

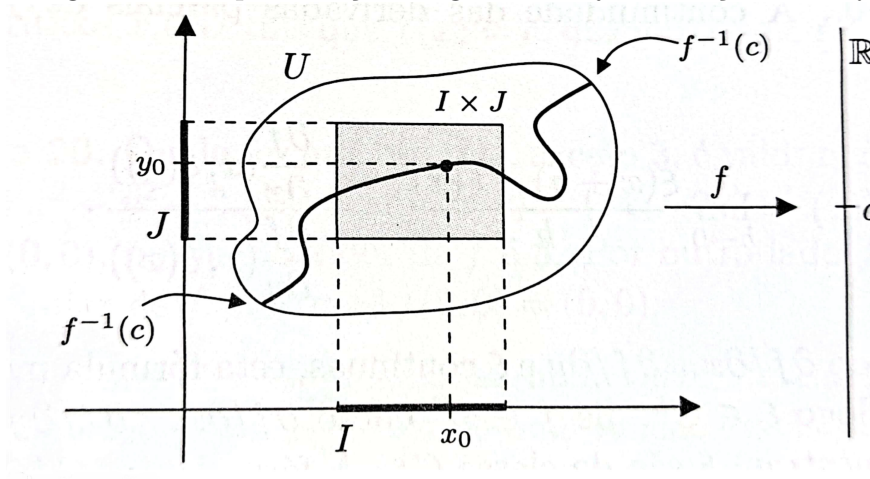
Essa expressão mostra que se f é de classe C^k , então as derivadas parciais de ξ são de classe C^{k-1} e portanto, ξ é de classe C^k . \square

Agora, considerando o conjunto aberto $V = B \times J \subset \mathbb{R}^{n+1}$, o teorema acima nos permite concluir que o conjunto

$$f^{-1}(c) \cap V = \{(x, \xi(x)) \in \mathbb{R}^n; x \in B\}$$

isto é, o conjunto dos pontos (x, y) que satisfazem $f(x, y) = c$, é o gráfico da função $\xi : B \rightarrow J$.

Figura 2.1: Representação do gráfico de ξ em relação ao de f .



Fonte: Lima, 2013

Exemplo 2.3. Considere a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 + y^2$. Para todo $(x, y) \in \mathbb{R}^2$, temos $\frac{\partial f}{\partial x}(x, y) = 2x$ e $\frac{\partial f}{\partial y}(x, y) = 2y$.

Se $c < 0$, a equação $x^2 + y^2 = c$ define o conjunto vazio, então não existe (x_0, y_0) tal que $f(x_0, y_0) = c$. Se $c = 0$, a equação $x^2 + y^2 = c$ é satisfeita somente no ponto $(0, 0)$, mas $\frac{\partial f}{\partial x}(0, 0) = 0$ e $\frac{\partial f}{\partial y}(0, 0) = 0$.

Finalmente, se $c > 0$, a equação $x^2 + y^2 = c$ define uma circunferência de centro na origem e raio \sqrt{c} , que não é gráfico de nenhuma função $y = \xi(x)$ ou $x = \zeta(y)$. No entanto, considerando os conjuntos

$$V_1 = \{(x, y) \in \mathbb{R}^2; y > 0\},$$

$$V_2 = \{(x, y) \in \mathbb{R}^2; y < 0\},$$

$$V_3 = \{(x, y) \in \mathbb{R}^2; x > 0\},$$

$$V_4 = \{(x, y) \in \mathbb{R}^2; x < 0\},$$

vemos que $f^{-1}(c) \cap V_1$ e $f^{-1}(c) \cap V_2$ são os gráficos das funções $\xi_1, \xi_2 : (-1, 1) \rightarrow \mathbb{R}$, dadas por $\xi_1(x) = \sqrt{c - x^2}$ e $\xi_2(x) = -\sqrt{c - x^2}$, enquanto $f^{-1}(c) \cap V_3$ e $f^{-1}(c) \cap V_4$ são os gráficos das funções $\zeta_1, \zeta_2 : (-1, 1) \rightarrow \mathbb{R}$, dadas por $\zeta_1(y) = \sqrt{c - y^2}$ e $\zeta_2(y) = -\sqrt{c - y^2}$.

Logo, em V_1 e V_2 , a equação $x^2 + y^2 = c$ define implicitamente y em função de x , enquanto que em V_3 e V_4 , a equação $x^2 + y^2 = c$ define implicitamente x em função de y .

Definição 2.8. Um conjunto $M \subset \mathbb{R}^{n+1}$ é chamado de hiperfície de classe C^k quando é localmente o gráfico de uma função real de n variáveis de classe C^k .

Mais precisamente, para cada $p \in M$, deve existir um aberto $V \subset \mathbb{R}^{n+1}$ e uma função $\xi : U \rightarrow \mathbb{R}$ de classe C^k em um aberto $U \subset \mathbb{R}^n$, tais que $p \in V$ e $V \cap M = \text{gráfico de } \xi$.

É claro que dada qualquer $f : U \rightarrow \mathbb{R}$ de classe C^k no aberto $U \subset \mathbb{R}^n$, seu gráfico é uma hiperfície $M = \{(x, f(x)) \in \mathbb{R}^{n+1}; x \in U\}$.

Exemplo 2.4. A esfera $S^n = \{x \in \mathbb{R}^{n+1}; \langle x, x \rangle = 1\}$ é uma hiperfície de classe C^∞ em \mathbb{R}^{n+1} . De fato, chamando de U a bola aberta de centro 0 e raio 1 em \mathbb{R}^n , para cada $i = 1, \dots, n+1$, $V_i = \{x \in \mathbb{R}^{n+1}; x_i > 0\}$, $W_i = \{x \in \mathbb{R}^{n+1}; x_i < 0\}$ e pondo $x^* = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})$, temos:

$$x \in S^n \cap V_i \Leftrightarrow |x^*| < 1 \text{ e } x_i = \sqrt{1 - \langle x^*, x^* \rangle};$$

$$x \in S^n \cap W_i \Leftrightarrow |x^*| < 1 \text{ e } x_i = -\sqrt{1 - \langle x^*, x^* \rangle}.$$

Então, tomando a função $\xi : U \rightarrow \mathbb{R}$, de classe C^∞ , definida por $\xi(u) = \sqrt{1 - \langle u, u \rangle}$, vemos que, para cada $i = 1, \dots, n+1$, $S^n \cap V_i$ é o gráfico da função $x_i = \xi(x^*)$ enquanto que $S^n \cap W_i$ é o gráfico de $x_i = -\xi(x^*)$. Como todo ponto $p \in S^n$ pertence a algum V_i ou algum W_i , concluímos que S^n é uma hiperfície de classe C^∞ em \mathbb{R}^{n+1} .

Definição 2.9. Seja $M \subset \mathbb{R}^{n+1}$ uma hiperfície de classe C^k . Para cada $p \in M$, associaremos o conjunto $T_p M$ formado por todos os vetores-velocidade $v = \lambda'(0)$ dos caminhos $\lambda : (-\delta, \delta) \rightarrow M$ que são diferenciáveis no ponto 0 e cumprem a condição $\lambda(0) = p$. O conjunto $T_p M$ é chamado de espaço vetorial tangente de M no ponto p .

Teorema 2.5. O conjunto $T_p M$ é um subespaço vetorial de dimensão n em \mathbb{R}^{n+1} .

Demonstração. Seja $\xi : U \rightarrow \mathbb{R}$ uma função de classe C^k no aberto $U \subset \mathbb{R}^n$, cujo gráfico, formado pelos pontos $(x, \xi(x)) \in \mathbb{R}^{n+1}$, $x \in U$, é a interseção $M \cap V$, onde $V \subset \mathbb{R}^{n+1}$ é um aberto que contém $p = (p_0, \xi(p_0))$, $p_0 \in U$. Para todo caminho $\lambda : (-\delta, \delta) \rightarrow M$, com $\lambda(0) = p$, tem-se $\lambda(t) = (x_1(t), \dots, x_n(t), \xi(x(t)))$, onde $x(t) = (x_1(t), \dots, x_n(t))$. Portanto,

$$\lambda'(0) = \left(\frac{dx_1}{dt}, \dots, \frac{dx_n}{dt}, \sum_{i=1}^n \frac{\partial \xi}{\partial x_i} \cdot \frac{dx_i}{dt} \right),$$

onde as derivadas $\frac{dx_i}{dt}$ são calculadas no ponto $t = 0$ e $\frac{\partial \xi}{\partial x_i}$ no ponto p_0 . Isso demonstra que todo $v = \lambda'(0)$ em $T_p M$ é uma combinação linear dos vetores

$$v_1 = \left(1, 0, \dots, 0, \frac{\partial \xi}{\partial x_1} \right), \dots, v_n = \left(1, 0, \dots, 0, \frac{\partial \xi}{\partial x_n} \right).$$

Reciprocamente, toda combinação linear $v = \sum_{i=1}^n \alpha_i v_i$ é o vetor-velocidade $\lambda'(0)$ do caminho $\lambda : (-\delta, \delta) \rightarrow M$ assim definido: tomamos $v_0 = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ e pomos $\lambda(t) = (p_0 + tv_0, \xi(p_0 + tv_0))$, sendo $\delta > 0$ escolhido de modo que o segmento de reta $(p_0 - \delta v_0, p_0 + \delta v_0)$ esteja contido em U . \square

Definição 2.10. Um número $c \in \mathbb{R}$ chama-se um valor regular de uma função $f : U \rightarrow \mathbb{R}$, de classe C^1 , quando não há pontos críticos de f no nível c . Isto é, quando $f(x) = c \Rightarrow \nabla f(x) \neq 0$. O número c também é chamado de nível regular de f .

Teorema 2.6. Se c é um valor regular da função $f : U \rightarrow \mathbb{R}$, de classe C^k no aberto $U \subset \mathbb{R}^{n+1}$, então $M = f^{-1}(c)$ é uma hipersuperfície de classe C^k , cujo espaço vetorial tangente $T_p M$ é, em cada ponto $p \in M$, o complemento ortogonal de $\nabla f(p)$.

Demonstração. Resulta do Teorema da Função Implícita que $f^{-1}(c)$ é uma hipersuperfície. Quanto ao espaço vetorial tangente $T_p M$, como M é uma superfície de nível da função f , vemos que todo vetor $v \in T_p M$ é ortogonal a $\nabla f(p)$, logo $T_p M \subset [\nabla f(p)]^\perp$. Sendo ambos subespaços de dimensão n em \mathbb{R}^{n+1} , conclui-se que $T_p M = [\nabla f(p)]^\perp$. \square

Exemplo 2.5. Pelo Teorema 2.6, a esfera unitária S^n é a hipersuperfície de nível 1 da função $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, dada por $f(x) = \langle x, x \rangle$. Como $\nabla f(x) = 2x$, vemos que zero é o único nível crítico de f . Em particular, 1 é valor regular, $S^n = f^{-1}(1)$ é uma hipersuperfície C^∞ e, para todo $o \in S^n$, tem-se $T_p S^n = [\nabla f(p)]^\perp = [p]^\perp$.

2.2 Diferenciabilidade para aplicações $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$

Nesta seção, apresentaremos o conceito de diferenciabilidade para aplicações do tipo $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$, alguns resultados, propriedades e exemplos.

Abaixo estão as definições de derivada parcial e direcional, que foram enunciadas na seção anterior, mas aqui, estas foram generalizadas para as aplicações de valores em \mathbb{R}^n .

Definição 2.11 (Derivada Parcial). Sejam $f : U \rightarrow \mathbb{R}^n$ definida em um conjunto aberto $U \subset \mathbb{R}^m$ e $a \in U$. A i -ésima derivada parcial de f no ponto a é o número

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t}, \text{ para } i = 1, \dots, m,$$

caso esse limite exista, onde e_i é o i -ésimo vetor da base canônica de \mathbb{R}^m .

Se $f = (f_1, \dots, f_n)$ então $\frac{\partial f}{\partial x_i}(a) = \left(\frac{\partial f_1}{\partial x_i}(a), \dots, \frac{\partial f_n}{\partial x_i}(a) \right)$. Logo, cada derivada parcial $\frac{\partial f}{\partial x_i}$ é um vetor de \mathbb{R}^n .

Definição 2.12 (Derivada Direcional). *Sejam $f : U \rightarrow \mathbb{R}^n$ definida em conjunto aberto $U \subset \mathbb{R}^m$, $a \in U$ e $v \in \mathbb{R}^m$. A derivada direcional de f no ponto a e na direção v é definida por*

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}.$$

Definição 2.13 (Diferenciabilidade). *Seja $U \subset \mathbb{R}^m$ um conjunto aberto, a aplicação $f : U \rightarrow \mathbb{R}^n$ é diferenciável em um ponto $a \in U$ se existe uma transformação linear $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$, tal que*

$$f(a + v) - f(a) = Tv + r(v), \text{ onde } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0.$$

Quando f é diferenciável em todos os pontos de U , dizemos que f é diferenciável em U , ou simplesmente, diferenciável.

Supondo que f é diferenciável no ponto a , para todo $v \in \mathbb{R}^m$ e qualquer $t \in \mathbb{R}$ suficiente pequeno, vale $f(a + tv) - f(a) = T(tv) + \rho(tv) \cdot |tv|$, onde $\lim_{v \rightarrow 0} \rho(tv) = 0$. Ora, temos $T(tv) = t \cdot Tv$ e $|tv| = |t| \cdot |v|$, então, para $t \neq 0$ segue que

$$\frac{f(a + tv) - f(a)}{t} = Tv \pm \rho(tv) \cdot |tv| \Rightarrow \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = Tv.$$

Logo, $Tv = \frac{\partial f}{\partial v}(a)$. Considerando $v = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$, para cada $i = 1, \dots, m$, vem

$$f(a + v) - f(a) = \sum_{i=1}^m \left[\frac{\partial f}{\partial x_i}(a) \cdot \alpha_i \right] + r(v), \text{ com } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0.$$

Em particular, vemos que T é a única transformação linear que fornece uma boa aproximação para o acréscimo $f(a + v) - f(a)$ na vizinhança do ponto a . Esta transformação T é chamada *derivada* de f em a e denotada por $f'(a)$. Assim, se $f : U \rightarrow \mathbb{R}^n$ definida em um aberto $U \subset \mathbb{R}^m$ é diferenciável no ponto a , sua derivada é a transformação linear $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ de forma que vale a igualdade

$$f(a + v) - f(a) = f'(a)v + r(v), \text{ onde } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$$

ou

$$f(a + v) - f(a) = f'(a)v + \rho(v) \cdot |v|, \text{ onde } \lim_{v \rightarrow 0} \rho(v) = 0.$$

Também se pode concluir deste desenvolvimento que se $f : U \rightarrow \mathbb{R}^n$ é diferenciável em um ponto $a \in U$, então existe a derivada direcional $\frac{\partial f}{\partial v}$ no ponto a , conseqüentemente, para $i = 1, \dots, m$, existem as derivadas parciais $\frac{\partial f}{\partial x_i}$ no ponto a . Também resulta que se as funções-coordenadas f_1, \dots, f_n de f são diferenciáveis em a , então existem as derivadas parciais $\frac{\partial f_1}{\partial x_i}, \dots, \frac{\partial f_n}{\partial x_i}$ no ponto a .

Teorema 2.7. *Seja $U \subset \mathbb{R}^m$ aberto. A aplicação $f : U \rightarrow \mathbb{R}^n$ é diferenciável no ponto $a \in U$ se, e somente se, suas funções-coordenada $f_1, \dots, f_n : U \rightarrow \mathbb{R}$ o são. Neste caso, vale $f'(a) = (f'_1(a), \dots, f'_n(a))$.*

Demonstração. Basta observar que a igualdade $f(a + v) - f(a) = f'(a)v + r(v)$ equivale a n igualdades numéricas $f_i(a + v) - f_i(a) = f'_i(a)v + r_i(v)$, onde $r(v) = (r_1(v), \dots, r_n(v))$, enquanto que o limite $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$ corresponde aos n limites numéricos $\lim_{v \rightarrow 0} \frac{r_i(v)}{|v|} = 0$. \square

Teorema 2.8. *Se a aplicação $f : U \rightarrow \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, é diferenciável num ponto $a \in U$, então é contínua neste ponto.*

Demonstração. Tomando a função resto $r(v) = f(a + v) - f(a) - f'(a)v$, que satisfaz $\lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$, percebe-se que

$$\lim_{v \rightarrow 0} r(v) = \lim_{v \rightarrow 0} \frac{r(v)}{|v|} \cdot |v| = 0,$$

então vale $\lim_{v \rightarrow 0} r(v) = 0$, sendo assim

$$\lim_{v \rightarrow 0} f(a + v) = \lim_{v \rightarrow 0} [f(a) + f'(a)v + r(v)] = f(a),$$

e portanto, f é contínua em a . \square

A transformação linear $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ tem relação com as bases canônicas de \mathbb{R}^m e \mathbb{R}^n . Para $j = 1, \dots, m$ e $i = 1, \dots, n$, temos

$$f'_i(a) \cdot e_j = \frac{\partial f_i}{\partial x_j} = \left(\frac{\partial f_i}{\partial x_1}, \dots, \frac{\partial f_i}{\partial x_m} \right),$$

que são m colunas de derivadas parciais onde cada uma dessas colunas é relativa a cada uma das variáveis x_j , e também temos

$$f'(a) \cdot e_j = f'_i(a) \cdot e_j = \frac{\partial f_i}{\partial x_j} = \left(\frac{\partial f_1}{\partial x_j}, \dots, \frac{\partial f_n}{\partial x_j} \right),$$

que são as n linhas de derivadas parciais relativas onde cada linha é relativa a cada uma das coordenadas de f . Assim, podemos definir uma matriz de dimensão $n \times m$ que contenha todas essas derivadas.

Definição 2.14 (Matriz Jacobiana). *A matriz $n \times m$ de $f'(a)$ relativa as bases \mathbb{R}^m e \mathbb{R}^n , é chamada de matriz jacobiana de f em a e é denotada por*

$$Jf(a) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_m}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_m}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \cdots & \frac{\partial f_n}{\partial x_m}(a) \end{bmatrix}.$$

Também podemos utilizar a notação $Jf(a) = \left[\frac{\partial f_i}{\partial x_j}(a) \right]$, onde $i = 1, \dots, n$ e $j = 1, \dots, m$.

Exemplo 2.6. *Sejam $I \subset \mathbb{R}$ um intervalo aberto e $f : I \rightarrow \mathbb{R}^n$ um caminho diferenciável no ponto $a \in I$. Sua derivada no ponto a é a transformação linear $f'(a) : \mathbb{R} \rightarrow \mathbb{R}^n$ cuja matriz jacobiana tem por única coluna o vetor*

$$v = \left(\frac{df_1}{dt}(a), \dots, \frac{df_n}{dt}(a) \right),$$

que é o vetor-velocidade de f no ponto a .

Exemplo 2.7. *Seja $U \subset \mathbb{R}^m$ um conjunto aberto e $f : U \rightarrow \mathbb{R}$ uma função diferenciável no ponto $a \in U$. Sua derivada no ponto a é a transformação linear $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}$, isto é, um funcional linear, que associa a cada vetor $v = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$ o número*

$$f'(a)v = \frac{\partial f}{\partial x_1}(a) \cdot \alpha_1 + \cdots + \frac{\partial f}{\partial x_m}(a) \cdot \alpha_m = \frac{\partial f}{\partial v}(a) = \langle \nabla f(a), v \rangle.$$

Definição 2.15. *Seja $U \subset \mathbb{R}^m$ aberto, se $f : U \rightarrow \mathbb{R}^n$ é diferenciável. Então:*

- I) *Existe a aplicação derivada $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, que associa a cada ponto $x \in U$ a transformação linear $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$, derivada de f no ponto x ;*
- II) *Para cada $v \in \mathbb{R}^m$, é definida a aplicação $\frac{\partial f}{\partial v} : U \rightarrow \mathbb{R}^n$, cujo valor em um ponto $x \in U$ é a derivada direcional $\frac{\partial f}{\partial v}(x) = f'(x)v$.*

Suponhamos que f é diferenciável e que sua aplicação derivada $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ seja contínua, pelo Teorema 2.7, suas funções-coordenada f_1, \dots, f_n são diferenciáveis e possuem derivadas parciais $\frac{\partial f_i}{\partial x_j} : U \rightarrow \mathbb{R}$ contínuas. Consequentemente, a derivada direcional $\frac{\partial f}{\partial v} : U \rightarrow \mathbb{R}^n$ também é contínua, pois

$$\frac{\partial f}{\partial v} = \sum_{i=1}^m \left[\frac{\partial f}{\partial x_i} \cdot \alpha_i \right], \text{ para } v = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m.$$

Definição 2.16 (Aplicação de classe C^1). Seja $U \subset \mathbb{R}^m$ aberto, se $f : U \rightarrow \mathbb{R}^n$ é diferenciável e sua aplicação derivada $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é contínua, então dizemos que f é uma aplicação de classe C^1 e representamos por $f \in C^1$.

De modo equivalente, f é uma aplicação de classe C^1 quando a derivada direcional $\frac{\partial f_i}{\partial v} : U \rightarrow \mathbb{R}$ é contínua.

Definição 2.17. Seja $U \subset \mathbb{R}^m$ aberto, dizemos que $f : U \rightarrow \mathbb{R}^n$ é duas vezes diferenciável em um ponto $a \in U$ quando sua aplicação derivada $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é diferenciável em a .

De modo equivalente, f é duas vezes diferenciável em a quando a derivada direcional $\frac{\partial f_i}{\partial v} : U \rightarrow \mathbb{R}$ é diferenciável em a .

Se $f : U \rightarrow \mathbb{R}^n$ é duas vezes diferenciável em $a \in U$, sua derivada segunda (ou derivada de ordem 2) em a é a aplicação bilinear $f''(a) : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^n$, cujo valor no ponto $(v, w) \in \mathbb{R}^m \times \mathbb{R}^m$ é o vetor de \mathbb{R}^n :

$$f''(a)vw = \frac{\partial}{\partial w} \left(\frac{\partial}{\partial v} \right) (a) = \frac{\partial^2 f}{\partial w \partial v}(a).$$

Se $v = (\alpha_1, \dots, \alpha_m)$ e $w = (\beta_1, \dots, \beta_m)$ então

$$f''(a) v w = \frac{\partial}{\partial w} \left(\frac{\partial}{\partial v} \right) (a) = \frac{\partial}{\partial w} \left(\sum_{j=1}^m \alpha_j \cdot \frac{\partial}{\partial x_j} \right) (a) = \sum_{j,k=1}^m \alpha_j \beta_k \cdot \frac{\partial^2 f}{\partial x_k \partial x_j}(a)$$

é o vetor de \mathbb{R}^n cujas coordenadas são

$$f''_i(a) v w = \sum_{j,k=1}^m \alpha_j \beta_k \cdot \frac{\partial^2 f_i}{\partial x_k \partial x_j}(a), \text{ onde } i = 1, \dots, n.$$

De modo semelhante a Definição 2.16, se $f : U \rightarrow \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, é diferenciável num ponto $a \in U$, e sua derivada $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é de classe C^1 , então dizemos que f é de classe C^2 e representamos por $f \in C^2$.

Isso significa que f é de classe C^2 quando a derivada direcional $\frac{\partial f_i}{\partial v} : U \rightarrow \mathbb{R}$ é de classe C^1 .

Prosseguindo com este argumento k -vezes, podemos dizer que a aplicação $f : U \rightarrow \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$, é k vezes diferenciável em um ponto $a \in U$ quando f é diferenciável em U e a aplicação derivada $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é $k - 1$ vezes diferenciável no ponto a , o que equivale dizer que a derivada direcional $\frac{\partial f_i}{\partial v} : U \rightarrow \mathbb{R}$ é $k - 1$ vezes diferenciável.

Se isso ocorre, define-se a k -ésima derivada (ou derivada de ordem k) de f no ponto a como sendo a aplicação k -linear $f^{(k)}(a) : \mathbb{R}^m \times \dots \times \mathbb{R}^m \rightarrow \mathbb{R}^n$, cujo valor no ponto $(v_1, \dots, v_k) \in \mathbb{R}^m \times \dots \times \mathbb{R}^m$ é o vetor de \mathbb{R}^n :

$$f^{(k)}(a)v_1 \cdots v_k = \frac{\partial^k f}{\partial v_k \partial v_{k-1} \cdots \partial v_1}(a).$$

Então, podemos dizer que f é de classe C^k quando é diferenciável e quando sua derivada $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é de classe C^{k-1} . Ou seja, a derivada direcional $\frac{\partial f_i}{\partial v} : U \rightarrow \mathbb{R}$ é de classe C^{k-1} , então f é de classe C^k e representamos por $f \in C^k$.

Finalmente, f é de classe C^0 quando é contínua, e de classe C^∞ quando é de classe C^k para $k = 0, 1, 2, 3, \dots$. Também é trivial que $f \in C^k \Rightarrow f \in C^{k-1}$.

Exemplo 2.8. *Toda aplicação constante é de classe C^∞ e sua derivada é nula.*

Exemplo 2.9. *Toda transformação linear $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é diferenciável e $T'x = T$, para todo $x \in \mathbb{R}^m$. De fato, basta observar que $T(x+v) - Tx = Tv + 0$. A aplicação derivada $T : \mathbb{R}^m \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é constante, então $T \in C^\infty$.*

Exemplo 2.10. *Toda forma bilinear $b : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^p$ é diferenciável em cada ponto $(a, b) \in \mathbb{R}^m \times \mathbb{R}^n$ e sua derivada é a forma bilinear $b' : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^p$ definida por*

$$b'(a, b) \cdot (v, w) = b(v, b) + b(a, w).$$

De fato, a bilinearidade de b nos fornece

$$b(a+v, b+w) - b(a, b) = b(v, b) + a(a, w) + a(v, w).$$

Como toda forma bilinear é uma aplicação Lipschitziana (veja [6], página 23), existe $c > 0$ tal que $|b(v, w)| \leq c|v||w|$, para todo $v \in \mathbb{R}^m$ e $w \in \mathbb{R}^n$. Tomando a norma da soma, temos $|(v, w)| = |v| + |w|$. Logo

$$\frac{|b(v, w)|}{|(v, w)|} \leq \frac{c|v||w|}{|v| + |w|} \leq c|v| \text{ e daí, } \lim_{(v,w) \rightarrow (0,0)} \frac{b(v, w)}{|(v, w)|} = 0.$$

Portanto, b é diferenciável e vale $b'(a, b) \cdot (v, w) = b(v, b) + b(a, w)$. Como b' também é uma transformação linear, então $b \in C^\infty$.

Exemplo 2.11. *Dado $U \subset \mathbb{R}^m$, se $f, g : U \rightarrow \mathbb{R}^n$ são diferenciáveis no ponto $a \in U$, então a aplicação $(f, g) : U \rightarrow \mathbb{R}^n \times \mathbb{R}^n$, definida por $(f, g)(x) = (f(x), g(x))$ é diferenciável em a e sua derivada é $(f, g)'(a)v = (f'(a)v, g'(a)v)$, para $v \in \mathbb{R}^m$. Se f e g são de classe C^k , então (f, g) também é de classe C^k .*

Teorema 2.9 (Regra da Cadeia). *Sejam os conjuntos abertos $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$, uma aplicação $f : U \rightarrow \mathbb{R}^n$ diferenciável no ponto $a \in U$, com $f(U) \subset V$, e uma aplicação $g : V \rightarrow \mathbb{R}^p$ diferenciável no ponto $b = f(a) \in V$. Então $g \circ f : U \rightarrow \mathbb{R}^p$ é diferenciável no ponto a e vale $(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$.*

Demonstração. Pela Definição 2.13, para $v \in \mathbb{R}^m$ e $w \in \mathbb{R}^n$, temos

$$f(a+v) = f(a) + f'(a)v + r(v), \text{ com } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0$$

e

$$g(b+w) = g(b) + g'(b)w + s(w), \text{ com } \lim_{w \rightarrow 0} \frac{s(w)}{|w|} = 0.$$

Logo,

$$(g \circ f)(a+v) = g(f(a+v)) = g(f(a) + f'(a)v + r(v)).$$

Fazendo $w = f'(a)v + r(v)$ e lembrando que $b = f(a)$, temos

$$\begin{aligned} (g \circ f)(a+v) &= g(f(a) + w) = g(f(a)) + g'(f(a)) \cdot f'(a)v + g'(f(a)) \cdot r(v) + s(w) \\ &= (g \circ f)(a) + [g'(f(a)) \cdot f'(a)]v + C(v), \end{aligned}$$

onde $C(v) = g'(f(a)) \cdot r(v) + s(w)$. Ora, se $v \rightarrow 0$ então $w \rightarrow 0$ e ocorre que

$$\lim_{v \rightarrow 0} C(v) = \lim_{v \rightarrow 0} (g'(f(a)) \cdot r(v)) + \lim_{v \rightarrow 0} s(w) = \lim_{v \rightarrow 0} (g'(f(a)) \cdot r(v)) + \lim_{v \rightarrow 0} (f'(a)v + r(v)) = 0.$$

Portanto, $g \circ f : U \rightarrow \mathbb{R}^p$ é diferenciável no ponto a e vale $(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$. \square

Corolário 2.9.1. *Nas condições do Teorema 2.9, se f e g são de classe C^k , então $g \circ f$ é de classe C^k*

Demonstração. Considere $x \in U$, pelo Teorema 2.9, temos $(g \circ f)(x) = g'(f(x)) \cdot f'(x)$. Considerando as aplicações derivadas $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, $g' : V \rightarrow \mathcal{L}(\mathbb{R}^n; \mathbb{R}^p)$ e $(g \circ f)' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^p)$ a igualdade acima pode ser escrita como

$$(g \circ f)' = (g' \circ f) \cdot f' : U \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n).$$

Se f e g são de classe C^1 , a igualdade acima mostra que $(g \circ f)'$ é contínua, então $(g \circ f) \in C^1$. Da mesma forma, se f e g são de classe C^2 , então $(g \circ f)'$ é de classe C^1 e $(g \circ f) \in C^2$. Por indução, supondo que f e g são de classe C^k , temos $(g \circ f)' \in C^{k-1}$, portanto, $(g \circ f) \in C^k$. \square

Corolário 2.9.2. *Nas condições do Teorema 2.9, a matriz jacobiana de $g \circ f$ no ponto a é o produto da matriz jacobiana de g no ponto $f(a)$ pela matriz jacobiana de f no ponto a . Ou seja, $J(g \circ f)(a) = Jg(f(a)) \cdot Jf(a)$.*

Demonstração. Observe que Jf é uma matriz $n \times m$, Jg é $p \times n$ e $J(g \circ f)$ é $p \times m$. Como as matrizes jacobianas são as matrizes das derivadas, a igualdade $J(g \circ f)(a) = Jg(f(a)) \cdot Jf(a)$ é a regra da cadeia em termos de matrizes. Através da multiplicação de matrizes, para cada $i = 1, \dots, p$ e cada $j = 1, \dots, m$, obtemos

$$\frac{\partial (g_i \circ f)}{\partial x_j}(a) = \sum_{k=1}^n \frac{\partial g_i}{\partial y_k}(f(a)) \cdot \frac{\partial f_k}{\partial x_j}(a),$$

que é a regra da cadeia para as funções-coordenada $g_i \circ f$. \square

Teorema 2.10 (Regras de Derivação). *Sejam $U \subset \mathbb{R}^m$ aberto, as aplicações de classe C^k $f, g : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ diferenciáveis em $a \in U$, α um número real e a forma bilinear $b : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^p$. São diferenciáveis em $a \in U$, e de classe C^k , as seguintes aplicações:*

I) $(f + g) : U \rightarrow \mathbb{R}^n$, valendo a igualdade $(f + g)'(a) = f'(a) + g'(a)$;

II) $(\alpha \cdot f) : U \rightarrow \mathbb{R}^n$, valendo a igualdade $(\alpha \cdot f)'(a) = \alpha \cdot f'(a)$;

III) $b(f, g) : U \rightarrow \mathbb{R}^p$ definida por $B(f, g)(x) = B(f(x), g(x))$, cuja derivada em a é

$$[b(f, g)]'(a)v = b(f'(a)v, g(a)) + b(f(a), g'(a)v).$$

Demonstração. I) Considere a forma bilinear $S : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $S(x, y) = x + y$, pelo Exemplo 2.9, temos $S' = S$. Então, observando que $(f + g)(a) = S \circ (f(a), g(a))$ obtemos, pelo Teorema 2.9, que

$$(f + g)'(a) = S \circ (f'(a), g'(a)) = f'(a) + g'(a).$$

Como S é uma forma bilinear, é de classe C^∞ , logo, de classe C^k . Então, pelo Corolário 2.9.1, $f + g$ é de classe C^k .

II) Considere a transformação linear $\alpha^* : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $\alpha^*(x) = \alpha \cdot x$, pelo Exemplo 2.9, temos $(\alpha^*)' = \alpha^*$. Então, observando que $(\alpha \cdot f)(a) = \alpha^* \circ f(a)$ obtemos, pelo Teorema 2.9, que

$$(\alpha \cdot f)'(a) = \alpha^* \circ f'(a) = \alpha \cdot f'(a).$$

Como α^* é uma transformação linear, é de classe C^∞ , logo, de classe C^k . Então, pelo Corolário 2.9.1, $\alpha \cdot f$ é de classe C^k .

III) Considere $b(f, g) = b \circ (f, g)$, pelo Teorema 2.9, vale

$$\begin{aligned} [b(f, g)]'(a)v &= [b \circ (f, g)]'(a)v = b'(f(a), g(a)) \cdot (f'(a)v, g'(a)v) \\ &= b(f'(a)v, g(a)) + b(f(a), g'(a)v). \end{aligned}$$

Pelos Exemplos 2.10 e 2.11, podemos concluir que $b \in C^k$. □

Exemplo 2.12. *Sejam $f, g : U \rightarrow \mathbb{R}^n$ diferenciáveis no aberto $U \subset \mathbb{R}^m$, e de classe C^k . Tomando $\varphi : U \rightarrow \mathbb{R}$ definida por $\varphi(x) = \langle f(x), g(x) \rangle$, para todo $v \in \mathbb{R}^m$, vale*

$$\varphi'(x)v = \langle f'(x)v, g(x) \rangle + \langle f(x), g'(x)v \rangle,$$

ou seja, a igualdade acima representa a derivada do produto interno.

Teorema 2.11 (Desigualdade do Valor Médio). *Seja $f : U \rightarrow \mathbb{R}^n$ diferenciável em todos os pontos do segmento de reta $[a, a + v] \subset U$. Se, para todo $t \in [0, 1]$, tem-se $|f'(a + tv)| \leq M$ então $|f(a + v) - f(a)| \leq M \cdot |v|$.*

Demonstração. O caminho $\lambda : [0, 1] \rightarrow \mathbb{R}^n$, definido por $\lambda(t) = f(a + tv)$, é diferenciável, com $\lambda'(t) = f'(a + tv) \cdot v$, portanto, $|\lambda'(t)| \leq |f'(a + tv)| \cdot |v| \leq M \cdot |v|$, para todo $t \in [0, 1]$. Segue da Desigualdade do Valor Médio para caminhos (veja [5], página 44) que $|\lambda(1) - \lambda(0)| \leq M \cdot |v| \cdot (1 - 0)$, isto é, $|f(a + v) - f(a)| \leq M \cdot |v|$. \square

Definição 2.18 (Imersão). *Dado um conjunto aberto $U \subset \mathbb{R}^m$, uma aplicação diferenciável $f : U \rightarrow \mathbb{R}^n$ é chamada de imersão quando, para todo $x \in U$, sua derivada $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é uma transformação linear injetora.*

Para que isto ocorra, é necessário que $m \leq n$, pois por f' ser injetora, para todo ponto $x \in \mathbb{R}^m$, deve existir um $y \in \mathbb{R}^n$ tal que $f'(x) = y$. Se $m \geq n$, então existe algum x sem que tenhamos um y tal que $f'(x) = y$, o que é contradição.

Definição 2.19 (Submersão). *Dado um conjunto aberto $U \subset \mathbb{R}^m$, uma aplicação diferenciável $f : U \rightarrow \mathbb{R}^n$ é chamada de submersão quando, para todo $x \in U$, sua derivada $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é uma transformação linear sobrejetora.*

Para que isto ocorra, é necessário que $m \geq n$, pois por f' ser sobrejetora, para todo $y \in \mathbb{R}^n$, deve existir um $x \in \mathbb{R}^m$ tal que $f'(x) = y$. Se $m \leq n$, então existe algum y sem que tenhamos um x tal que $f'(x) = y$, o que é contradição.

2.3 Difeomorfismo

Nesta seção, apresentaremos a definição de difeomorfismo e alguns resultados, estes possuem bastante importancia para a demonstração do Teorema da Aplicação Inversa no capítulo seguinte.

Definição 2.20 (Difeomorfismo). *Sejam os conjuntos abertos $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$. Um difeomorfismo entre U e V é uma aplicação bijetora e diferenciável $f : U \rightarrow V$ cuja inversa $f^{-1} : V \rightarrow U$ é também bijetora e diferenciável.*

Quando necessário, um difeomorfismo será chamado de difeomorfismo global. Se $f : U \rightarrow V$ é um difeomorfismo, com $g = f^{-1} : V \rightarrow U$, então de $g \circ f = id_U$ e $f \circ g = id_V$ resulta, pelo Teorema 2.9, que $g'(f(x)) \cdot f'(x) = id_{\mathbb{R}^m}$ e $f'(g(x)) \cdot g'(x) = id_{\mathbb{R}^n}$, para todo $x \in U$. Portanto, $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é um isomorfismo cujo inverso é $g'(f(x)) : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Definição 2.21 (Difeomorfismo Local). *Dado o aberto $U \subset \mathbb{R}^m$, uma aplicação diferenciável $f : U \rightarrow \mathbb{R}^n$ chama-se difeomorfismo local quando, para $x \in U$ existe uma bola aberta $B = B(x; \delta) \subset U$ tal que f aplica B difeomorficamente sobre um aberto V contendo $f(x)$.*

Todo difeomorfismo é um difeomorfismo local. E como um difeomorfismo local é uma aplicação aberta (veja [6], página 277) concluímos que um difeomorfismo local é um difeomorfismo global se, e somente se, é uma aplicação injetiva.

Teorema 2.12. *Se o difeomorfismo $f : U \rightarrow V$ é de classe C^k ($k \geq 1$), então seu inverso $g = f^{-1} : V \rightarrow U$ é também de classe C^k .*

Demonstração. Para todo $y = f(x) \in V$, temos $g'(y) = [f'(x)^{-1}] = [f'(f^{-1}(y))]^{-1}$, portanto a aplicação $g' : V \rightarrow \mathcal{L}(\mathbb{R}^m)$ se exprime como a composta

$$g' = (Inv) \circ f' \circ f^{-1}$$

onde Inv leva todo operador invertível $X : \mathbb{R}^m \rightarrow \mathbb{R}^m$ no seu inverso X^{-1} , $f' : U \rightarrow \mathcal{L}(\mathbb{R}^m)$ leva todo ponto $x \in U$ na derivada invertível $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ e $f^{-1} : V \rightarrow U$ é a aplicação inversa de f .

Sabemos que $Inv \in C^\infty$. Portanto, se $f \in C^k$ então $f' \in C^{k-1}$ e, pela hipótese de indução, $f^{-1} \in C^{k-1}$, logo, $g' \in C^{k-1}$ como composta de três aplicações de classe C^{k-1} , por definição, $g \in C^k$. \square

Teorema 2.13. *Seja $f : U \rightarrow \mathbb{R}^n$ de classe C^k ($k \geq 1$) definida no aberto $U \subset \mathbb{R}^m$. Se para algum $a \in U$, a derivada $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é injetiva, então existem $\delta > 0$ e $c > 0$ tais que $B = B(a; \delta) \subset U$ e, para quaisquer $x, y \in B$ tem-se $|f(x) - f(y)| \geq c|x - y|$. Em particular, a restrição $f|_B$ é injetiva.*

Demonstração. A função $u \mapsto |f'(a) \cdot u|$ é positiva em todos os pontos u da esfera unitária S^{m-1} , que é compacta. Pelo Corolário 1.30.1, existe $c > 0$ tal que $|f'(a) \cdot u| \geq 2c$ para todo $u \in S^{m-1}$.

Por linearidade, segue-se que $|f'(a) \cdot v| \geq 2c \cdot |v|$ para todo $v \in \mathbb{R}^m$. Para todo $x \in U$, escrevamos

$$r(x) = f(x) - f(a) - f'(a)(x - a).$$

Então, para $x, y \in U$ quaisquer, temos

$$f(x) - f(y) \geq |f'(a)(x - y)| - |r(x) - r(y)|.$$

levando em conta que $|u + v| \geq |u| - |v|$, segue-se que

$$|f(x) - f(y)| \geq |f'(a) \cdot (x - y)| - |r(x) - r(y)| \geq 2c \cdot |x - y| - |r(x) - r(y)|.$$

Ora, a aplicação $r(x)$ definida, é de classe C^1 , com $r(a) = 0$ e $r'(a) = 0$. Pela continuidade de r' , existe $\delta > 0$ tal que $|x - a| < \delta \Rightarrow x \in U$ e $|r'(x)| < c$. Aplicando o Teorema 2.11 a r em $B = B(a; \delta)$, temos que se $x, y \in B$ então $|r(x) - r(y)| \leq c|x - y|$. Consequentemente, $x, y \in B \Rightarrow |f(x) - f(y)| \geq 2c|x - y| - c|x - y|$, ou seja, $|f(x) - f(y)| \geq c|x - y|$. \square

Teorema 2.14 (Diferenciabilidade do Homeomorfismo Inverso). *Seja $f : U \rightarrow \mathbb{R}^n$ um homeomorfismo de classe C^1 entre os abertos $U, V \subset \mathbb{R}^m$. Se para algum $x \in U$, a derivada $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é um operador invertível, então o homeomorfismo inverso $g = f^{-1} : V \rightarrow U$ é diferenciável no ponto $f(x)$, com $g'(f(x)) = [f'(x)]^{-1}$.*

Demonstração. Se $x, x + v \in U$, escrevemos $f(x) = y$ e $f(x + v) = y + w$, então

$$w = f(x + v) - f(x) = f'(x)v + r(v), \text{ onde } \lim_{v \rightarrow 0} \frac{r(v)}{|v|} = 0 \text{ e}$$

$$v = g(f(x + v)) - g(f(x)) = g(y + w) - g(y).$$

Para provar que $f(x)^{-1}$ é a derivada de g no ponto y , escrevamos

$$g(y + w) - g(y) = f'(x)^{-1}w + s(w)$$

e mostemos que $\lim_{w \rightarrow 0} \frac{s(w)}{|w|} = 0$.

Tomando a igualdade acima com as expressões de v e w obtidas, temos

$$v = f'(x)^{-1}[f'(x)v + r(v)] + s(w)$$

ou seja,

$$v = v + f'(x)^{-1} \cdot r(v) + s(w),$$

donde,

$$s(w) = -f'(x)^{-1} \cdot r(v), \text{ logo } \frac{s(w)}{|w|} = -f'(x)^{-1} \cdot \frac{r(v)}{|v|} \cdot \frac{|v|}{|w|},$$

isto é,

$$\frac{s(w)}{|w|} = -f'(x)^{-1} \cdot \frac{r(v)}{|v|} \cdot \frac{|v|}{|f(x + v) - f(x)|}.$$

Quando $w \rightarrow 0$, tem-se $v \rightarrow 0$ pela continuidade de g , então $\frac{r(v)}{|v|} \rightarrow 0$, e pelo Teorema 2.13, existem $\delta > 0$ e $c > 0$ tais que $|v| < \delta$ implica

$$|f(x + v) - f(x)| \geq c|v|, \text{ portanto, } \frac{|v|}{|f(x + v) - f(x)|} \leq \frac{1}{c}.$$

Assim, $\lim_{w \rightarrow 0} \frac{s(w)}{|w|} = 0$. Portanto, g é diferenciável e $g'(f(x)) = [f'(x)]^{-1}$. □

Capítulo 3

Os Teoremas da Aplicação Inversa e da Função Implícita

Neste capítulo, utilizando os resultados estabelecidos nos capítulos anteriores, serão enunciados e demonstrados o Teorema da Aplicação Inversa e o Teorema da Função Implícita.

3.1 Teorema da Aplicação Inversa

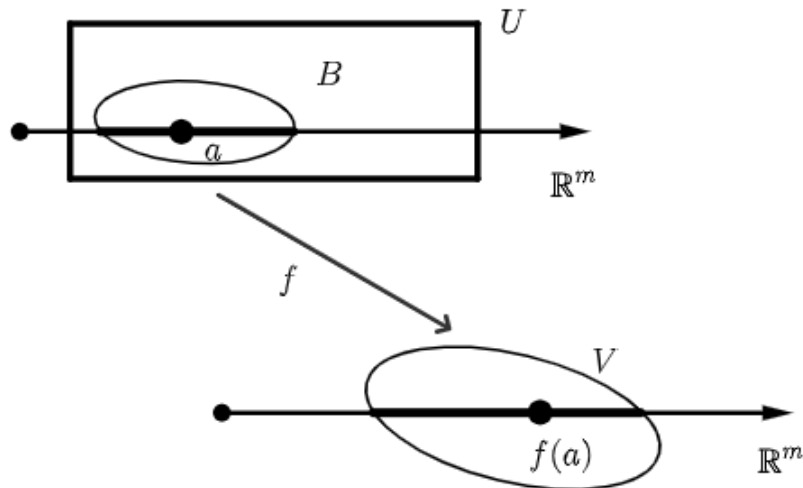
Nesta seção, iremos enunciar e demonstrar o Teorema da Aplicação Inversa, ele afirma que dado uma aplicação $f : U \rightarrow \mathbb{R}^n$ de classe C^k e definida no aberto $U \subset \mathbb{R}^m$, se existe um ponto $x \in U$ tal que $f'(x)$ é invertível, então existe um conjunto aberto em U tal que f restrita a este conjunto é um difeomorfismo sobre um aberto que contém $f(x)$. Antes de enunciarmos o teorema, iremos enunciar e demonstrar um lema com o objetivo de demonstrar o teorema desejado. Em seguida, apresentaremos uma aplicação e dois exemplos deste teorema.

Lema 3.1. *Sejam $U \subset \mathbb{R}^m$ aberto e $g : U \rightarrow \mathbb{R}^n$ diferenciável no ponto $a \in U$, com $g'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ sobrejetiva. Se a é um ponto de mínimo local de $|g(x)|$, $x \in U$, então $g(a) = 0$.*

Demonstração. Se a é um ponto de mínimo local para $|g(x)|$, será também um ponto de mínimo local para a função $\varphi : U \rightarrow \mathbb{R}$, definida por $\varphi(x) = |g(x)|^2 = \langle g(x), g(x) \rangle$, logo $\varphi'(a) = 0$ (veja [4], página 97), mas como $\varphi'(a) \cdot v = 2\langle g'(a) \cdot v, g(a) \rangle$, ou seja, $g(a)$ é ortogonal à imagem de $g'(a)$, que é \mathbb{R}^n . Logo $g(a) = 0$. \square

Teorema 3.1 (Teorema da Aplicação Inversa). *Sejam $f : U \rightarrow \mathbb{R}^m$ uma aplicação de classe C^k com $k \geq 1$ definida no aberto $U \subset \mathbb{R}^m$. Se $a \in U$ é tal que $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é invertível, então existe uma bola aberta $B = B(a; \delta) \subset U$ tal que a restrição $f|_B$ é um difeomorfismo sobre um aberto V que possui o ponto $f(a)$.*

Figura 3.1: Ilustração do Teorema da Aplicação Inversa.



Fonte: Autor

Demonstração. Pela hipótese, $B(a, \delta) \subset U$, diminuindo δ suficientemente, podemos admitir que $\bar{B} = B[a; \delta] \subset U$ e que f é injetiva no compacto \bar{B} (pelo Teorema 2.13), então pelo Teorema 1.30, f é um homeomorfismo de \bar{B} sobre $f(\bar{B})$, e conseqüentemente, f é um homeomorfismo de B sobre $f(B)$.

Além disso, como $f'(x)$ depende continuamente de x , podemos supor que, para todo $x \in B$, a derivada $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é um isomorfismo. Então, pelo Teorema 2.14, basta mostrar que $f(B) = V \subset \mathbb{R}^m$ é um conjunto aberto.

Seja $q = f(p)$, $p \in B$. Consideremos $S = S[a; \delta]$ a esfera que é a fronteira de \bar{B} , a injetividade de $f|_{\bar{B}}$ nos garante que $q \notin f(S)$, então existe $\epsilon > 0$ tal que $|f(x) - q| \geq 2\epsilon$ para todo $x \in S$, pois S é compacto. Agora basta mostrar que $B(q; \epsilon) \subset V$.

Seja $y \in B(q; \epsilon)$, colocando $g(x) = |f(x) - y|$, pelo Corolário 1.30.1, $g(x)$ possui máximos e mínimos. Devido a injetividade de $f|_{\bar{B}}$, temos que o mínimo de $g(x)$, para $x \in \bar{B}$, não é atingido por num ponto $x \in S$, pois $x \in S \Rightarrow |f(x) - y| \geq \epsilon$, enquanto $|f(p) - y| = |q - y| < \epsilon$, com $p \in B$, então o mínimo de $g(x) = |f(x) - y|$ é atingido por em um ponto $x_0 \in B$. Pelo Lema 3.1, esse mínimo é zero, então $f(x_0) = y$ onde $f(x_0) \in V$, então $y \in V$.

Logo, $B(q; \epsilon) \subset V$, então V é aberto e pelo Teorema 2.14, $f|_B$ é um difeomorfismo sobre o aberto V que possui $f(a)$. O que demonstra o teorema. \square

Exemplo 3.1. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $f(x, y) = (e^x \cos(y), e^x \sin(y))$. Sua matriz jacobiana é

$$Jf(x, y) = \begin{bmatrix} e^x \cos(y) & -e^x \sin(y) \\ e^x \sin(y) & e^x \cos(y) \end{bmatrix}$$

e seu determinante é e^{2x} , que é sempre diferente de 0. Logo, $f'(x, y) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é um isomorfismo, para todo $(x, y) \in \mathbb{R}^2$. Pelo Teorema da Aplicação Inversa, f é um difeomorfismo local (mas não um difeomorfismo global pois $f(x, y + 2\pi) = f(x, y)$), ou seja, f admite

inversa local em qualquer ponto de \mathbb{R}^2 .

Exemplo 3.2. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $f(x, y) = \left(\frac{x}{\sqrt{1+x^2+y^2}}, \frac{y}{\sqrt{1+x^2+y^2}} \right)$. Sua matriz jacobiana é

$$Jf(x, y) = \begin{bmatrix} \frac{1+y^2}{\sqrt{1+x^2+y^2}(1+x^2+y^2)} & -\frac{xy}{\sqrt{1+x^2+y^2}(1+x^2+y^2)} \\ -\frac{xy}{\sqrt{1+x^2+y^2}(1+x^2+y^2)} & \frac{1+x^2}{\sqrt{1+x^2+y^2}(1+x^2+y^2)} \end{bmatrix}$$

e seu determinante é $\frac{1}{(1+x^2+y^2)^2}$, que é sempre diferente de 0. Logo, $f'(x, y) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ é um isomorfismo, para todo $(x, y) \in \mathbb{R}^2$. Pelo Teorema da Aplicação Inversa, f é um difeomorfismo local, ou seja, f admite inversa local em qualquer ponto de \mathbb{R}^2 .

Aplicação: Lema de Morse

Como aplicação do Teorema da Aplicação Inversa, iremos enunciar e demonstrar o Lema de Morse, que afirma que na vizinhança de um ponto crítico não degenerado de uma função f é possível tomar um sistema de coordenadas em relação ao qual f se exprime como uma forma quadrática com coeficientes constantes.

Antes do lema, apresentaremos algumas definições essenciais para o entendimento deste resultado. Outras definições preliminares e exemplos sobre o Lema de Morse podem ser encontrados em [6], e para as definições de Álgebra Linear utilizadas aqui, recomendamos [7] e [2].

Definição 3.1. Uma forma quadrática $H : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função cujo valor no vetor $v = (\alpha_1, \dots, \alpha_n)$ é $\sum_{i,j=1}^n h_{ij}\alpha_i\alpha_j$, onde $[h_{ij}]$ é uma matriz simétrica $n \times n$.

Dada uma função $f : U \rightarrow \mathbb{R}$, de classe C^2 , definida no aberto $U \subset \mathbb{R}^n$, a forma quadrática hessiana $H(x) = (Hf)(x)$ de f em $x \in U$ é aquela cuja matriz é

$$[h_{ij}] = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right].$$

Definição 3.2. Seja $f : U \rightarrow \mathbb{R}$ diferenciável, de classe C^2 e definida no aberto $U \subset \mathbb{R}^n$, um ponto crítico $a \in U$ é chamado de ponto crítico não-degenerado quando a matriz hessiana de f é invertível.

Definição 3.3. Um sistema de coordenadas de classe C^k em um aberto $U \subset \mathbb{R}^n$ é um difeomorfismo $\xi : V \rightarrow U$ de classe C^k , definido em um aberto $V \subset \mathbb{R}^n$. As coordenadas de um ponto $p \in U$ no sistema ξ são os números y_1, \dots, y_n tais que $y = (y_1, \dots, y_n) \in V$ e $\xi(y) = p$.

Lema 3.2 (Lema de Morse). *Seja a um ponto crítico não-degenerado de uma função $f : U \rightarrow \mathbb{R}$ de classe C^k ($k \geq 3$) definida no aberto $U \subset \mathbb{R}^n$. Então existe um sistema de coordenadas $\xi : V \rightarrow W$ de classe C^{k-2} , com $a \in W \subset U$, $0 \in V$ e $\xi(0) = a$, tal que*

$$f(\xi(y)) - f(a) = \sum_{i,j=1}^n a_{ij} y_i y_j$$

para todo $y = (y_1, \dots, y_n) \in V$, onde

$$a_{ij} = \frac{1}{2} \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

Demonstração. Para simplificar a notação, suponhamos $a = 0$ e $f(a) = 0$. Seja W uma bola aberta de centro 0, contida em U . Pela fórmula de Taylor com resto integral (veja em [6], página 151), temos

$$x \in W \Rightarrow f(x) = \sum_{i,j=1}^n a_{ij} y_i y_j$$

com

$$a_{ij}(x) = \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(tx) dx.$$

Como vemos, cada a_{ij} é uma função de classe C^{k-2} , definida na bola W . Para cada $x \in W$ a matriz $A(x) = (a_{ij}(x))$ é simétrica, em virtude do Teorema de Schwarz (veja em [6], página 147). Podemos escrever $f(x) = \langle A(x) \cdot x, x \rangle$ para todo $x \in W$. Seja $A_0 = A(0) = \frac{1}{2} \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}(0)$. Dizer que o ponto crítico 0 é não degenerado significa afirmar que a matriz simétrica A_0 é invertível. Assim, para cada $x \in W$, podemos escrever $A(x)$ como produto de A_0 por uma matriz que depende de x em classe C^{k-2} e que, para $x = 0$ é a matriz identidade. Através do Exemplo 15 (Veja [6], página 283), podemos tomar o raio da bola W tão pequeno que $x \in W \Rightarrow A(x) = A_0 \cdot B(x)^2$, $B : W \rightarrow \mathbb{R}^{n^2}$ é de classe C^{k-2} . Como A_0 e $A(x)$ (para todo $x \in W$) são simétricas, tomando tranpostas obtemos:

$$A = A_0 \cdot B^2 = (B^*)^2 \cdot A_0, \text{ donde } B^2 = A_0^{-1} (B^*)^2 A_0 = (A_0^{-1} \cdot B^* \cdot A_0^2),$$

onde a variável x foi omitida por simplicidade, ainda pelo Exemplo 15 (Veja [6], página 283), se o raio de W for suficiente pequeno, $x \in W$ implicará que $B(x)$ e $A_0^{-1} \cdot (B^*)^2 \cdot A_0$ estão tão próximas da identidade que, por terem quadrados iguais, são elas iguais. Então $B = A_0^{-1} \cdot B^* \cdot A_0$, ou seja, $A_0 \cdot B = B^* \cdot A_0$ e daí $A = A_0 \cdot B^2 = B^* \cdot A_0 \cdot B$. Assim:

$$x \in W \Rightarrow f(x) = \langle A(x) \cdot x, x \rangle = \langle B(x)^* \cdot A_0 \cdot B(x) \cdot x, x \rangle = \langle A_0 \cdot B(x) \cdot x, B(x) \cdot x \rangle$$

Agora, basta demonstrarmos que, se o raio da bola W for tomado suficientemente pequeno, a aplicação $\varphi(x) : W \rightarrow \mathbb{R}^n$ definida por $\varphi(x) = B(x) \cdot x$ é um difeomorfismo de classe C^{k-2} sobre sua imagem.

Para todo $x \in W$ e todo $v \in \mathbb{R}^n$, temos $\varphi'(x) \cdot v = \frac{\partial \varphi}{\partial v}(x) = \frac{\partial B}{\partial v}(x) = x + B(x) \cdot v$.

Para $x = 0$, vem: $\varphi'(0) \cdot v = B(0) \cdot v = v$. Logo, $\varphi'(0) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é a transformação linear identidade. Resulta do Teorema da Aplicação Inversa que, se o raio de W for tomado suficiente pequeno, obtemos um difeomorfismo $\varphi : W \rightarrow V$ de classe C^{k-2} , com $\varphi(x) = B(x) \cdot x$, $\varphi(0) = 0$ e $f(x) = \langle A_0 \cdot \varphi(x), \varphi(x) \rangle$, para todo $t \in W$. Então $\xi = \varphi^{-1} : V \rightarrow W$ é um sistema de coordenadas de classe C^{k-2} tal que, para todo $y \in V$, $f(\xi(y)) = \langle A_0 \cdot y, y \rangle = \sum_{i,j=1}^n a_{ij} y_i y_j$. \square

3.2 Teorema da Função Implícita

Na primeira seção do capítulo anterior, apresentamos um caso particular do Teorema da Função Implícita, a qual se aplicava em funções do tipo $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, este permite encontrar uma função implícita $y = y(x)$ ou $x = x(y)$, associada a equação $f(x, y) = c$, com $c \in \mathbb{R}$.

Como resultado do Teorema da Aplicação Inversa, temos o caso geral do Teorema da Função Implícita, que pode ser aplicado para funções $f : U \subset \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$. Neste caso, os pontos do espaço $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$ serão representados da forma (x, y) , onde $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ e $y = (y_1, \dots, y_n) \in \mathbb{R}^n$.

Antes de apresentarmos o teorema, faremos algumas preliminares e apresentaremos alguns resultados. Para outras definições e exemplos, recomendamos [6], [9], [10] e [8].

Definição 3.4 (Difeomorfismo Vertical). *Um difeomorfismo $h : U \rightarrow V$ entre abertos $U, V \subset \mathbb{R}^{m+n}$ é chamado de difeomorfismo vertical quando é da forma $h(x, y) = (x, h_2(x, y))$, ou seja, quando deixar invariante a coordenada x .*

O inverso de um difeomorfismo vertical é também vertical.

Teorema 3.2 (Forma Local das Submersões). *Seja $f = (f_1, \dots, f_n)$ uma aplicação de classe C^k ($k \geq 1$) de um aberto $U \subset \mathbb{R}^{m+n}$ em \mathbb{R}^n . Se em um ponto $p = (a, b) \in U$, a matriz*

$$\left[\frac{\partial f_i}{\partial y_j}(p) \right] \quad (i, j = 1, \dots, n)$$

é invertível, então existem abertos $Z \ni p$ em \mathbb{R}^{m+n} , $V \ni a$ em \mathbb{R}^m , $W \ni c = f(p)$ em \mathbb{R}^n e um difeomorfismo vertical $h : V \times W \rightarrow Z$, de classe C^k , tal que $f(h(x, w)) = w$ para todo $x \in V$ e todo $w \in W$.

Demonstração. Seja $\varphi : U \rightarrow \mathbb{R}^m \times \mathbb{R}^n$ a aplicação de classe C^k definida por $\varphi(x, y) = (x, f(x, y))$. A matriz jacobiana de φ tem a forma

$$J\varphi = \begin{bmatrix} I & 0 \\ a & b \end{bmatrix},$$

onde I é a matriz identidade $m \times m$ e a matriz $n \times n$

$$b = b(z) = \left[\frac{\partial f_i}{\partial y_j}(z) \right],$$

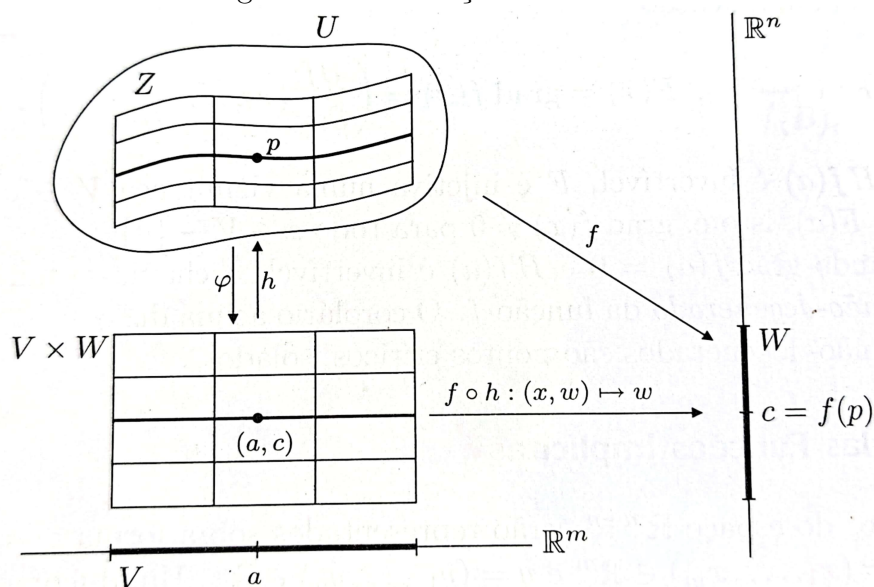
é, no ponto $p = (a, b)$, invertível.

Pelo Teorema da Aplicação Inversa, φ é um difeomorfismo de um aberto $Z \ni p$ sobre um aberto de $\mathbb{R}^m \times \mathbb{R}^n$, o qual podemos supor da forma $V \times W$, onde $V \subset \mathbb{R}^m$ e $W \subset \mathbb{R}^n$, com $a \in V$ e $c = f(a, b) \in W$. O difeomorfismo inverso $h : V \times W \rightarrow Z$ é da forma $h(x, w) = (x, h_2(x, w))$. Então, para qualquer $(x, w) \in V \times W$, tem-se

$$(x, w) = \varphi(h(x, w)) = \varphi(x, h_2(x, w)) = (x, f(x, h_2(x, w))) = (x, f(h(x, w))),$$

logo $f(h(x, w)) = w$ para qualquer $(x, w) \in V \times W$. □

Figura 3.2: Ilustração do teorema.



Fonte: Lima, 2013

Corolário 3.2.1. *Seja $f : U \rightarrow \mathbb{R}^n$ uma submersão de classe C^k , definida no aberto $U \subset \mathbb{R}^{m+n}$. Para cada ponto $z \in U$ existem abertos $Z \subset U$ contendo z , $W \subset \mathbb{R}^n$ contendo $c = f(z)$, $V \subset \mathbb{R}^m$ e um difeomorfismo $h : V \times W \rightarrow Z$ de classe C^k , tais que $f(h(x, w)) = w$ para todo $x \in V$ e todo $w \in W$.*

Demonstração. Como $f'(z) : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ é sobrejetiva, n das $m+n$ colunas são linearmente independentes, logo formam uma matriz invertível $n \times n$. Se essas forem as últimas colunas, então o corolário é o Teorema 3.2. Se não forem, basta modificar a demonstração daquele teorema, permutando as coordenadas em \mathbb{R}^{m+n} de modo que as n colunas linearmente independentes de $Jf(z)$ sejam as últimas. □

Teorema 3.3 (Teorema da Função Implícita). *Sejam o conjunto aberto $U \subset \mathbb{R}^{m+n}$, $f : U \rightarrow \mathbb{R}^n$ uma aplicação de classe C^k ($k \geq 1$) e $(x_0, y_0) \in U$ tais que $f(x_0, y_0) = c$, $c \in \mathbb{R}^n$. Nestas condições, se $f'(x_0, y_0)|_{\{0\} \times \mathbb{R}^m}$ é um isomorfismo sobre \mathbb{R}^n , então existem abertos $V \ni (x_0, y_0)$, $V \subset U$ e $A \ni x_0$, $A \subset \mathbb{R}^m$ tais que*

- I) Para todo $x \in A$, existe um único $y = \xi(x) \in \mathbb{R}^n$ tal que $(x, \xi(x)) \in V$ e $f(x, \xi(x)) = c$;
 II) A aplicação $\xi : A \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é de classe C^k e $\xi'(x_0) = -T_2^{-1}T_1$, onde $T_1 \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ e $T_2 \in \mathcal{L}(\mathbb{R}^n)$ são definidos por

$$T_1 h = f'(x_0, y_0)(h, 0) \text{ e } T_2 k = f'(x_0, y_0)(0, k),$$

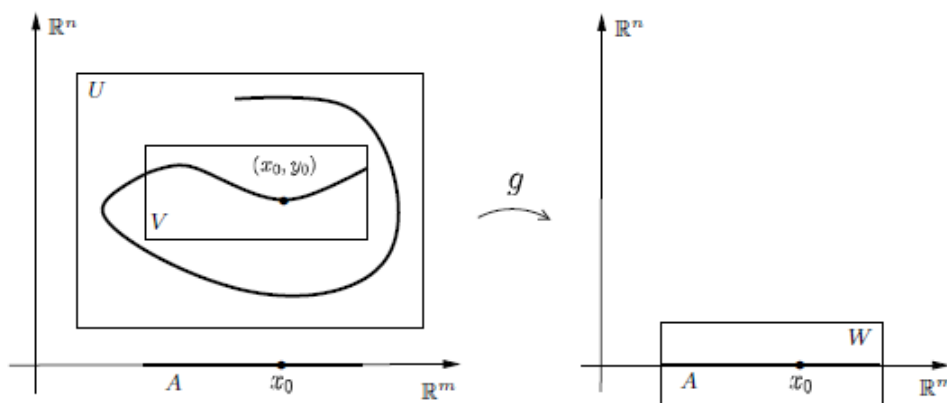
onde $h \in \mathbb{R}^m$ e $k \in \mathbb{R}^n$.

Demonstração. I) Sem perda de generalidade, suponhamos $c = 0$ e definamos $g : U \rightarrow \mathbb{R}^{m+n}$ de classe C^k onde $g(x, y) = (x, f(x, y))$ e

$$g'(x_0, y_0)(h, k) = (h, f'(x_0, y_0)(h, k)).$$

Em particular, se $g'(x_0, y_0)(h, k) = (0, 0)$, teremos $h = 0$ e $0 = f'(x_0, y_0)(h, k) = f'(x_0, y_0)(0, k)$, onde $k = 0$, pois da hipótese, $f'(x_0, y_0)|_{\{0\} \times \mathbb{R}^n}$ é um isomorfismo. Então $g'(x_0, y_0)$ é um isomorfismo de \mathbb{R}^{m+n} sobre si mesmo. Pelo Teorema da Aplicação Inversa, existem abertos $V \ni (x_0, y_0)$, $V \subset U$ e $W \ni (x_0, 0)$, $W \subset \mathbb{R}^{m+n}$, tais que $g|_V : V \rightarrow W$ é um difeomorfismo de classe C^k .

Figura 3.3: Ilustração da afirmação anterior.



Fonte: Freire, 2013

Novamente, sem perda de generalidade, suponhamos que $W = A \times B$, onde $A \subset \mathbb{R}^m$ e $B \subset \mathbb{R}^n$ são abertos e $x_0 \in A$. Então dado $x \in A$ existe um único $y = \xi(x) \in \mathbb{R}^n$ e $g(x, y) = g(x, f(x, y)) = (x, 0) \in W$, pois $g|_V : V \rightarrow W$ é bijetora. Então a aplicação ξ satisfaz $f(x, \xi(x)) = 0$, para todo $x \in A$.

II) Considere $(g|_V)^{-1} : W \rightarrow V$, pelo Teorema da Aplicação Inversa, $(g|_V)^{-1}$ é de classe C^k . Além disso, $(g|_V)^{-1}(x, 0) = (x, \xi(x))$. Definindo $P(x, y) = y$ e $\phi(x) = (x, 0)$,

com $x \in \mathbb{R}^m$, temos $\xi = P \circ (g|V)^{-1} \circ \phi$, onde ξ é de classe C^k . Note que, para todo $(h, k) \in \mathbb{R}^{m+n}$, temos

$$f'(x_0, y_0)(h, k) = f'(x_0, y_0)((h, 0) + (0, k)) = f'(x_0, y_0)(h, 0) + f'(x_0, y_0)(0, k) = T_1 h + T_2 k.$$

Derivando ambos os membros de $f(x, \xi(x)) = 0$ e tomando $x = x_0$, obtemos

$$0 = f'(x_0, \xi(x_0))(h, \xi'(x_0)h) = f'(x_0, y_0)(h, \xi'(x_0)h) = T_1 h + T_2 \xi'(x_0)h, \forall h \in \mathbb{R}^m.$$

De onde obtemos que $\xi'(x_0) = -T_2^{-1}T_1$, o que conclui a demonstração. \square

Observação 3.1. Se $c \neq 0$, poderíamos definir uma função $F(x, y) = f(x, y) - c$, e basear toda a demonstração sobre F , por isso, não se perde a generalidade ao supor $c = 0$.

Se $f : U \subset \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ é uma submersão de classe C^k , conjunto

$$B((x_0, y_0); \delta) \cap \{f^{-1}(c)\} = \{(x, y) \in \mathbb{R}^{m+n}; f(x, y) = c\}$$

é o gráfico da aplicação ξ de classe C^k .

Exemplo 3.3. Considere a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por $f(x, y, z) = \sin(xy) + \sin(xz) + \sin(yz)$. Vamos encontrar um ponto (x_0, y_0, z_0) tal que $f(x_0, y_0, z_0) = 0$ possua uma solução $z = \xi(x, y)$.

Derivando f em relação a z , obtemos

$$\frac{\partial f}{\partial z}(x, y, z) = x \cdot \cos(xz) + y \cdot \cos(yz).$$

Essa derivada só é igual a 0 se $x = y = 0$ e como temos $f(x, y, z) = 0$ se duas das três coordenadas forem iguais a 0, concluímos que os pontos das formas $(x_0, 0, 0)$ e $(0, y_0, 0)$, com $x_0 \neq 0$ e $y_0 \neq 0$, são pontos dos quais vale o Teorema da Função Implícita.

De fato, para o ponto $(x_0, 0, 0)$, temos $\frac{\partial f}{\partial z}(x_0, 0, 0) = x_0$ e $f(x_0, 0, 0) = 0$. Então, pelo Teorema da Função Implícita, na vizinhança de $(x_0, 0, 0)$, existe uma solução $z = \xi(x, y)$ para $f(x, y, z) = 0$.

Agora, para o ponto $(0, y_0, 0)$, temos $\frac{\partial f}{\partial z}(0, y_0, 0) = y_0$ e $f(0, y_0, 0) = 0$. Então, pelo Teorema da Função Implícita, na vizinhança de $(0, y_0, 0)$, existe uma solução $z = \xi(x, y)$ para $f(x, y, z) = 0$.

Aplicação: Multiplicadores de Lagrange

Como aplicação do Teorema da Função Implícita, iremos enunciar e demonstrar o método dos Multiplicadores de Lagrange, este se aplica quando se deseja encontrar os extremos locais da restrição de uma função diferenciável f a um subconjunto $M \subset U \subset \mathbb{R}^m \times \mathbb{R}^n$, que também é uma submersão de classe C^1 , $g : U \rightarrow \mathbb{R}^n$. Nestas condições, para que o

ponto $p \in M$ seja um extremo local de $f|M$, é necessário que o gradiente de f em p esteja no espaço gerado pelos gradientes das coordenadas de $g = (g_1, \dots, g_n)$ em p , isto é, devem existir $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ tais que

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p) + \dots + \lambda_n \nabla g_n(p),$$

onde $\lambda_1, \lambda_2, \dots, \lambda_n$ são chamados de multiplicadores de lagrange.

Teorema 3.4 (Multiplicadores de Lagrange). *Sejam $U \subset \mathbb{R}^{m+n}$ um conjunto aberto, $f : U \rightarrow \mathbb{R}$ uma função diferenciável e $M \subset U$ um conjunto de nível de uma submersão de classe C^1 , $g = (g_1, \dots, g_n) : U \rightarrow \mathbb{R}^n$. Então, se $p \in M$ é um extremo local de $f|M$, existem $\lambda_1, \lambda_2, \dots, \lambda_n$ tais que*

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p) + \dots + \lambda_n \nabla g_n(p)$$

Demonstração. Seja $p = (x_0, y_0) \in M$ um extremo local de $f|M$. Uma vez que g é uma submersão, temos que $g'(p) \in \mathcal{L}(\mathbb{R}^{m+n}, \mathbb{R}^n)$ é sobrejetora, de onde podemos supor, sem perda de generalidade, que $g'(p)|\{0\} \times \mathbb{R}^n$ é um isomorfismo sobre \mathbb{R}^n . Logo, pelo Teorema da Função Implícita, existem um aberto de \mathbb{R}^m , $A \ni x_0$ e uma aplicação de classe C^1 , $\xi : A \rightarrow \mathbb{R}^n$, tais que o gráfico de ξ está contido em M , ou seja, $(x, \xi(x)) \in M$ para todo $x \in A$.

Considere o gráfico de $\xi(x_0)$, $G = \{(h, \xi'(x_0)h); h \in \mathbb{R}^m\} \subset \mathbb{R}^{m+n}$ e observe que este é um subespaço vetorial de $\mathbb{R}^m \times \mathbb{R}^n$. Sendo M um conjunto de nível de g , temos que $(x, \xi(x))$ é constante. Em particular, dado um vetor $h \in \mathbb{R}^m$, tem-se $g'(x_0, \xi(x))(h, \xi'(x_0)h) = 0$, isto é, $g'(p)(G) = 0$. Logo, para todo $v = (h, \xi'(x_0)h) \in G$, teremos

$$0 = g'(p)v = (g'_1(p)v, g'_2(p)v, \dots, g'_n(p)v) = (\langle \nabla g'_1(p), v \rangle, \langle \nabla g'_2(p), v \rangle, \dots, \langle \nabla g'_n(p), v \rangle),$$

de onde concluímos que para cada $j \in \{1, 2, \dots, n\}$, $\nabla g'_j(p)$ é ortogonal a G . Porém, os vetores-linha da matriz jacobiana de g em p são, justamente, $\nabla g'_1(p), \nabla g'_2(p), \dots, \nabla g'_n(p)$, e portanto são linearmente independentes, pois a imagem de $g'(p)$ tem dimensão n . Segue daí que o complemento ortogonal de G em \mathbb{R}^{m+n} é o subespaço gerado pelos vetores $\nabla g'_1(p), \nabla g'_2(p), \dots, \nabla g'_n(p)$.

Uma vez que p é um extremo local de $f|M$, temos que x_0 é um extremo local de $f(x, \xi(x))$, $x \in A$. Então dado $v = (h, \xi'(x_0)h) \in G$, teremos

$$0 = f'(x_0, \xi(x_0))(h, \xi'(x_0)h) = \langle \nabla f(p), v \rangle,$$

onde $f(p) \in G^\perp$. Logo, existem reais $\lambda_1, \lambda_2, \dots, \lambda_n$ tais que

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p) + \dots + \lambda_n \nabla g_n(p).$$

Como queríamos demonstrar. □

Exemplo 3.4. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = ax + by$, com $a^2 + b^2 \neq 0$. Temos $\nabla f(x, y) = (a, b) = v$, onde v é um vetor ortogonal às linhas de nível $ax + by = c$, que são retas, duas a duas paralelas. f não possui pontos críticos, mas se $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ for dada por $\varphi(x, y) = x^2 + y^2$, então $\nabla \varphi(x, y) = (2x, 2y)$, 1 é o valor regular de φ e $M = \varphi^{-1}(1)$ é a circunferência unitária $x^2 + y^2 = 1$. Como M é compacta, a restrição $f|M$ possui pelo menos dois pontos críticos, nos quais assume seus valores mínimo e máximo. Os pontos críticos de $f|M$ são as soluções (x, y) do sistema

$$\nabla f(x, y) = \lambda \cdot \nabla \varphi(x, y), \varphi(x, y) = 1.$$

Ou seja,

$$2\lambda x = a, 2\lambda y = b \text{ e } x^2 + y^2 = 1.$$

Portanto (x, y) é um ponto crítico de $f|M$ se, e somente se, o vetor unitário $z = (x, y)$ é um múltiplo do vetor $v = (a, b)$, o que nos dá

$$(x, y) = \left(\frac{a}{\sqrt{a^2 + b^2}}, \frac{b}{\sqrt{a^2 + b^2}} \right) \text{ ou } (x, y) = \left(\frac{-a}{\sqrt{a^2 + b^2}}, \frac{-b}{\sqrt{a^2 + b^2}} \right).$$

Estes são os pontos nos quais $f(x, y)$ assume seus valores máximo e mínimo em $M = S^1$.

Apêndice A

Noções de Álgebra Linear

Neste apêndice, com o objetivo de fornecer uma base para as definições de álgebra linear utilizadas neste trabalho, estudaremos algumas definições e resultados referentes a espaços vetoriais, transformações lineares e formas bilineares.

Para um estudo mais profundo, podem ser consultadas as referências [7] e [2].

A.1 Espaços Vetoriais

Definição A.1 (Espaço Vetorial). *Um conjunto E , cujos elementos são chamados de vetores, é chamado de espaço vetorial quando nele estão definidas duas operações:*

- a) A soma, que a cada par $u, v \in E$ faz corresponder o elemento $u + v \in E$.*
- b) O produto com um número real, que a cada $\alpha \in \mathbb{R}$ e $u \in E$, faz corresponder o elemento $\alpha \cdot u \in E$.*

Essas operações estão definidas em E de modo que, para $u, v, w \in E$ e $\alpha, \beta \in \mathbb{R}$, valem as seguintes propriedades:

Comutatividade: $u + v = v + u$ e $\alpha u = u\alpha$;

Associatividade: $u + (v + w) = (u + v) + w$ e $\alpha(\beta u) = (\alpha\beta)u$;

Elemento neutro: existe $0 \in E$ tal que $u + 0 = u$;

Multiplicação por 1: $u \cdot 1 = u$;

Inverso aditivo: para todo $u \in E$, existe $-u \in E$ tal que $u + (-u) = 0$;

Distributividade: $(\alpha + \beta)u = \alpha u + \beta u$ e $\alpha(u + v) = \alpha u + \alpha v$.

Observação A.1. *Os elementos 0 e $-u$ são chamados de vetor nulo e oposto de u , respectivamente.*

Resulta da Definição A.1 os seguintes resultados, cujas demonstrações podem ser encontradas em [2], página 50:

1. Para todo $u \in E$, tem-se $0 \cdot u = 0$.
2. Dados $\alpha \in \mathbb{R}$ e $u \in E$, tem-se $\alpha u = 0$ se, e somente se, $\alpha = 0$ ou $u = 0$.

3. Se $w + u = w + v$, então $u = v$.
4. Para $\alpha \in \mathbb{R}$ e $u \in E$, vale $(-\alpha)u = \alpha(-u) = -(\alpha u)$.
5. Para $\alpha, \beta \in \mathbb{R}$ e $u \in E$, vale $(\alpha - \beta)u = \alpha u - \beta u$.

Definição A.2 (Subespaço Vetorial). Dado um espaço vetorial E , um subconjunto $F \subset E$ é chamado subespaço vetorial quando nele, valem as seguintes propriedades:

- a) $0 \in F$;
- b) Se $u, v \in F$, então $u + v \in F$;
- c) Se $u \in F$, então para todo $\alpha \in \mathbb{R}$, tem-se $\alpha u \in F$.

Para um subespaço vetorial ser considerado um espaço vetorial, ele deve satisfazer as 6 propriedades da Definição A.1.

Definição A.3. Seja E um espaço vetorial, tomando o subconjunto $S = \{u_1, \dots, u_n\} \subset E$, o subconjunto

$$[S] = \{\alpha_1 u_1 + \dots + \alpha_n u_n \mid \alpha_1, \dots, \alpha_n \in \mathbb{R}\}$$

é chamado subespaço gerado por S .

Teorema A.1. O Conjunto $[S]$ é um subespaço vetorial.

Demonstração. a) $0 = 0u_1 + \dots + 0u_n$, então $0 \in S$.

b) Sendo $v = \alpha_1 u_1 + \dots + \alpha_n u_n$ e $w = \beta_1 u_1 + \dots + \beta_n u_n$ pertencentes a S , temos

$$v + w = (\alpha_1 + \beta_1)u_1 + \dots + (\alpha_n + \beta_n)u_n,$$

que também pertence a S .

c) Sendo $\beta \in \mathbb{R}$ e $v = \alpha_1 u_1 + \dots + \alpha_n u_n \in S$, temos

$$\beta v = (\beta \alpha_1)u_1 + \dots + (\beta \alpha_n)u_n,$$

que também pertence a S . □

Definição A.4. Um espaço vetorial E é dito finitamente gerado se existe $S \subset E$, com S finito, tal que $E = [S]$.

Definição A.5. Seja E um espaço vetorial, o subconjunto $L = \{u_1, \dots, u_n\} \subset E$ é chamado linearmente independente (L.I) se, e somente se, a igualdade

$$\alpha_1 u_1 + \dots + \alpha_n u_n = 0$$

só ocorre quando $\alpha_1 = \dots = \alpha_n = 0$.

Definição A.6. Seja E um espaço vetorial, o subconjunto $L = \{u_1, \dots, u_n\} \subset E$ é chamado linearmente dependente (L.D) se, e somente se, não é L.I, isto é, a igualdade

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0$$

ocorre sem que $\alpha_1 = \cdots = \alpha_n = 0$.

Definição A.7 (Base). *Seja E um espaço vetorial finitamente gerado, uma base de E é um subconjunto $B \subset E$ tal que, valem as seguintes propriedades:*

- a) $[B] = E$;
- b) B é linearmente independente.

Isso significa que qualquer vetor de E pode ser representado como uma combinação linear dos vetores da base B .

Definição A.8 (Dimensão). *Seja E um espaço vetorial finitamente gerado, Chamamos de dimensão o número de vetores de quaisquer de suas bases, e denotamos por $\dim E$.*

Exemplo A.1. *O conjunto de vetores de n coordenadas*

$$\{(1, 0, \dots, 0); (0, 1, \dots, 0); \dots; (0, 0, \dots, 1)\}$$

é a base canônica de \mathbb{R}^n , isso é claro pois considerando o vetor $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$, temos

$$(a_1, a_2, \dots, a_n) = a_1(1, 0, \dots, 0) + a_2(0, 1, \dots, 0) + \cdots + a_n(0, 0, \dots, 1).$$

E podemos concluir que $\dim \mathbb{R}^n = n$.

A.2 Transformações Lineares

Definição A.9 (Transformação Linear). *Sejam E e F espaços vetoriais, uma aplicação $T : E \rightarrow F$ que associa a cada vetor $u \in E$, um vetor $T(u) = Tu \in F$, é chamada de transformação linear se, e somente se, para $u, v \in E$ e $\alpha \in \mathbb{R}$, valem as seguintes igualdades:*

- a) $T(u + v) = Tu + Tv$;
- b) $T(\alpha u) = \alpha \cdot Tu$.

Caso $E = F$, então $T : E \rightarrow E$ é chamado de *operador linear*. Uma transformação linear $\varphi : E \rightarrow \mathbb{R}$ é chamada de *funcional linear*.

Os espaços vetoriais E e F são chamados *domínio* e *contra-domínio* de T , respectivamente.

Dado $U \subset E$, é chamado de *imagem* de U por T o conjunto $F(U) = \{Tu \mid u \in U\} \subset F$. Se $U = E$, então $F(U)$ é a imagem de T e é denotada por $Im(T)$.

Exemplo A.2. A aplicação $0 : E \rightarrow F$ definida por $0u = 0$ é uma transformação linear.

De fato,

a) $0(u + v) = 0 = 0 + 0 = 0u + 0v,$

b) $0(\alpha u) = 0 = \alpha 0 = \alpha \cdot 0u.$

Essa aplicação é chamada de transformação linear nula.

Exemplo A.3. A aplicação $I : E \rightarrow E$ definida por $Iu = u,$ é uma transformação linear.

Com efeito,

a) $I(u + v) = u + v = Iu + Iv,$

b) $I(\alpha u) = \alpha u = \alpha \cdot Iu.$

Essa transformação linear é chamada de operador identidade.

Resulta da Definição A.9 os seguintes resultados, cujas demonstrações podem ser encontradas em [2], página 105:

1. $T0 = 0$
2. $T(-u) = -Tu, \forall u \in E.$
3. $T(u - v) = Tu - Tv, \forall u, v \in E.$

Definição A.10. Uma transformação linear $T : E \rightarrow F$ é chamada injetora se, e somente se, para $u, v \in E,$ vale $u \neq v \Leftrightarrow Tu \neq Tv$ ou $u = v \Leftrightarrow Tu = Tv.$

Definição A.11. Uma transformação linear $T : E \rightarrow F$ é chamada sobrejetora se, e somente se, $Im(T) = F,$ isto é, para todo $v \in F,$ existe $u \in E$ tal que $Tu = v.$

Definição A.12. Uma transformação linear $T : E \rightarrow F$ é chamada bijetora se, e somente se, é injetora e sobrejetora.

Se a transformação linear $T : E \rightarrow F$ é bijetora, então cada elemento de F é do tipo $Tu,$ com $u \in E,$ logo, fazendo a associação $Tu \mapsto u,$ teremos uma transformação linear de F em $E,$ essa aplicação é chamada de inversa de T e é denotada por $T^{-1}.$ Nesse caso, dizemos que T é invertível.

Definição A.13 (Isomorfismo). Se uma transformação linear $T : E \rightarrow F$ é bijetora, então ela é chamada de isomorfismo de E em $F.$

Caso $E = F,$ então $T : E \rightarrow E$ é chamado de *automorfismo* de $E.$

Dados os espaços vetoriais E e $F,$ o conjunto de todas as transformações lineares $T : E \rightarrow F$ será denotado por $\mathcal{L}(E; F).$ No caso dos operadores lineares $T : E \rightarrow E,$ a notação é $\mathcal{L}(E).$

Definição A.14 (Propriedades Operatórias). Sejam E e F espaços vetoriais, dadas as transformações lineares $S, T \in \mathcal{L}(E; F),$ o vetor $u \in E$ e o número real $\alpha \in \mathbb{R},$ definimos, respectivamente, a soma $S + T$ e o produto αT como sendo

$S + T : E \rightarrow F$ definida por $(S + T)u = Su + Tu$,

e

$\alpha \cdot T : E \rightarrow F$ definida por $(\alpha \cdot T)u = \alpha \cdot Tu$.

Observamos facilmente que $S + T : E \rightarrow F$ é uma transformação linear, pois

- a) $(S + T)(u + v) = S(u + v) + T(u + v) = Su + Tu + Sv + Tv = (S + T)u + (S + T)v$;
- b) $(S + T)(\alpha u) = S(\alpha u) + T(\alpha u) = \alpha Su + \alpha Tu = \alpha(Su + Tu) = \alpha(S + T)u$.

Também $\alpha \cdot T : E \rightarrow F$ é uma transformação linear, pois

- a) $(\alpha \cdot T)(u + v) = \alpha \cdot T(u + v) = \alpha \cdot (Tu + Tv) = \alpha \cdot Tu + \alpha \cdot Tv = (\alpha \cdot T)u + (\alpha \cdot T)v$;
- b) $(\alpha \cdot T)(\beta u) = \alpha \cdot T(\beta u) = \alpha \cdot \beta \cdot Tu = \beta \cdot \alpha \cdot Tu = \beta \cdot (\alpha \cdot T)u$.

Sendo assim, observamos que a soma e o produto de $\mathcal{L}(E; F)$ satisfazem as 6 propriedades da Definição A.1, portanto, $\mathcal{L}(E; F)$ é um espaço vetorial.

Definição A.15 (Composta). *Sejam E, F e G espaços vetoriais, se $T : E \rightarrow F$ e $S : F \rightarrow G$ são transformações lineares, é chamado de aplicação composta de T e S a aplicação*

$S \circ T : E \rightarrow G$ definida por $(S \circ T)u = S(Tu)$, $\forall u \in E$.

Ora, é fácil ver que $(S \circ T) \in \mathcal{L}(E; G)$, pois

- a) $(S \circ T)(u + v) = S(T(u + v)) = S(Tu + Tv) = S(Tu) + S(Tv) = (S \circ T)u + (S \circ T)v$.
- b) $(S \circ T)(\alpha u) = S(T(\alpha u)) = S(\alpha \cdot Tu) = \alpha \cdot S(Tu) = \alpha \cdot (S \circ T)u$.

Se $T : E \rightarrow F$ é invertível, então existe a transformação inversa $T^{-1} : F \rightarrow E$, logo, para $u \in E$ e $v \in F$, temos

$$(T^{-1} \circ T)u = T^{-1}(Tu) = u = Iu \quad \text{e} \quad (T \circ T^{-1})v = T(T^{-1}v) = v = Iv.$$

Ou seja, a composta de uma transformação linear com a sua inversa é o operador identidade.

Teorema A.2. *Se $T : E \rightarrow F$ é um isomorfismo, então sua inversa $T^{-1} : F \rightarrow E$ também é um isomorfismo.*

Demonstração. Primeiro, mostremos que T^{-1} é uma transformação linear.

a) Sejam $v_1, v_2 \in F$ e suponhamos $T^{-1}(v_1 + v_2) = u$. Como T é sobrejetora, existem $u_1, u_2 \in E$ tais que $Tu_1 = v_1$ (se, e somente se, $T^{-1}v_1 = u_1$) e $Tu_2 = v_2$ (se, e somente se, $T^{-1}v_2 = u_2$). Substituindo esses resultados na igualdade inicial, temos

$$u = T^{-1}(v_1 + v_2) = T^{-1}(Tu_1 + Tu_2) = T^{-1}(T(u_1 + u_2)) = u_1 + u_2 = T^{-1}v_1 + T^{-1}v_2,$$

e pela igualdade inicial, resulta $T^{-1}(v_1 + v_2) = T^{-1}v_1 + T^{-1}v_2$.

b) Sejam $v \in F$ e $\alpha \in \mathbb{R}$ e suponhamos $T^{-1}(\alpha v) = u$. Como T é sobrejetora, existe $u_1 \in E$ tal que $Tu_1 = v$. Logo (se, e somente se, $T^{-1}v = u_1$)

$$u = T^{-1}(\alpha v) = T^{-1}(\alpha \cdot Tu_1) = T^{-1}(T(\alpha u_1)) = \alpha u_1 = \alpha \cdot T^{-1}v.$$

Portanto, $T^{-1}(\alpha v) = \alpha \cdot T^{-1}v$ e T^{-1} é uma transformação linear. Agora, basta demonstrarmos que T^{-1} é injetora e sobrejetora.

Injetora: Suponhamos $v_1, v_2 \in F$ tais que $T^{-1}v_1 = T^{-1}v_2 = u$. Então $Tu = v_1$ e $Tu = v_2$. Logo, $v_1 = v_2$ e T^{-1} é injetora.

Sobrejetora: Dado $u \in E$, supondo $v = Tu$, temos $T^{-1}v = T^{-1}(Tu) = u$. □

A.3 Formas Bilineares

Definição A.16. *Sejam E e F espaços vetoriais, uma forma bilinear $b : E \times F \rightarrow \mathbb{R}$ é uma função que é linear em cada uma de suas variáveis $u \in E$ e $v \in F$, de modo que, para $u_1, u_2 \in E$, $v_1, v_2 \in F$ e $\alpha \in \mathbb{R}$, valem as seguintes propriedades:*

- a)** $b(u_1 + u_2, v) = b(u_1, v) + b(u_2, v)$;
- b)** $b(u, v_1 + v_2) = b(u, v_1) + b(u, v_2)$;
- c)** $b(\alpha u, v) = \alpha b(u, v)$;
- d)** $b(u, \alpha v) = \alpha b(u, v)$.

O conjunto das formas bilineares $b : E \times F \rightarrow \mathbb{R}$ é denotado por $\mathcal{B}(E \times F)$.

Dados $f, g \in \mathcal{B}(E \times F)$ e $\lambda \in \mathbb{R}$, definimos a soma $f + g : E \times F \rightarrow \mathbb{R}$ como $(f + g)(u, v) = f(u, v) + g(u, v)$ e o produto $\lambda f : E \times F \rightarrow \mathbb{R}$ como $(\lambda f)(u, v) = \lambda f(u, v)$.

$(f + g)$ e λf pertencem a $\mathcal{B}(E \times F)$, a demonstração pode ser encontrada em [2], página 221. Então observamos que a soma e o produto em $\mathcal{B}(E \times F)$ satisfazem as 6 propriedades da Definição A.1, portanto, $\mathcal{B}(E \times F)$ é um espaço vetorial.

Definição A.17. *Uma forma bilinear $f : E \times E \rightarrow \mathbb{R}$ é chamada simétrica se $f(u, v) = f(v, u)$, para todo $(u, v) \in E \times E$.*

Se f e g são simétricas, então $f + g$ e λf também são simétricas.

Definição A.18. *Uma forma bilinear $f : E \times E \rightarrow \mathbb{R}$ é chamada anti-simétrica se $f(u, v) = -f(v, u)$, para todo $(u, v) \in E \times E$.*

Se f e g são anti-simétricas, então $f + g$ e λf também são anti-simétricas.

Referências Bibliográficas

- [1] BOYER, Carl. **História da Matemática**. 1.ed. São Paulo: Editora da USP, 1974.
- [2] CALLIOLI, Carlos A.; DOMINGUES, Hygino H.; COSTA, Roberto C.F. **Álgebra Linear e Aplicações**. 6.ed. São Paulo: Atual, 2009.
- [3] EVES, Howard. **Introdução à História da Matemática**. 5.ed. Campinas: Editora da Unicamp, 2011.
- [4] LIMA, Elon L. **Análise Real Vol.1, Funções de uma Variável**. 8.ed. Rio de Janeiro: IMPA, 2006.
- [5] LIMA, Elon L. **Análise Real Vol.2, Funções de N Variáveis**. 6.ed. Rio de Janeiro: IMPA, 2013.
- [6] LIMA, Elon L. **Curso de Análise Vol.2**. 11.ed. Rio de Janeiro: IMPA, 2015.
- [7] LIMA, Elon L. **Álgebra Linear**. 10.ed. Rio de Janeiro: IMPA, 2014.
- [8] LIMA, Leandro L. **Teorema da Função Implícita e da Função Inversa e abordagens de Funções Reais de Várias Variáveis no Ensino Médio**. 2021. Dissertação (Mestrado Profissional em Matemática do Programa de Pós-Graduação em Matemática em Rede Nacional) - Instituto de Geociências e Ciências Exatas, UEP, Rio Claro, 2021.
- [9] LIMA, Ronaldo F. **Topologia e Análise no Espaço \mathbb{R}^n** . 1.ed. Natal: UFRN, 2013.
- [10] ZACARIAS, André P.S. **O Teorema da Aplicação Inversa e Aplicações**. 2022. Dissertação (Mestrado Profissional em Matemática do Programa de Pós-Graduação em Matemática em Rede Nacional) - Centro de Ciências, UFC, Fortaleza, 2022.