UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
LICENCIATURA PLENA EM MATEMÁTICA

Bruna Vitória Borges Bezerra

O teorema da função inversa e sua relação com as superfícies regulares

Recife-PE
Setembro de 2023

O teorema da função inversa e sua relação com as superfícies regulares

por

Bruna Vitória Borges Bezerra

Sob orientação de

Prof. Dr. Gilson Mamede de Carvalho

Trabalho de Conclusão de Curso submetido à Coordenação do Curso de Licenciatura plena em Matemática da Universidade Federal Rural de Pernambuco - Sede, como parte dos requisitos para a obtenção do grau de Licenciado em Matemática.

Recife-PE
Setembro de 2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

B574t Bezerra, Bruna Vitória Borges
O teorema da função inversa e sua relação com as superfícies regulares / Bruna Vitória Borges Bezerra. - 2023.
101 f. : il.

Orientador: Gilson Mamede de Carvalho.
Inclui referências.

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal Rural de Pernambuco, Licenciatura em Matemática, Recife, 2024.

1. Espaço Euclidiano. 2. Diferenciabilidade. 3. Teorema da função inversa. 4. Superfícies regulares. I. Carvalho, Gilson Mamede de, orient. II. Título

CDD 510

Bruna Vitória Borges Bezerra

O teorema da função inversa e sua relação com as superfícies regulares

Trabalho de Conclusão de Curso submetido à Coordenação do Curso de Licenciatura plena em Matemática da Universidade Federal Rural de Pernambuco - Sede, como parte dos requisitos para a obtenção do grau de Licenciado em Matemática.

Aprovado em: 18/09/2023.

BANCA EXAMINADORA

Prof. Dr. Gilson Mamede de Carvalho
Universidade Federal Rural de Pernambuco (UFRPE)

Prof. Dr. Eudes Mendes Barboza
Universidade Federal Rural de Pernambuco (UFRPE)

Prof. Dr. Diego Ferraz de Souza
Universidade Federal do Rio Grande do Norte (UFRN)

Recife-PE
Setembro de 2023

“Não que sejamos capazes, por nós, de pensar alguma coisa, como de nós mesmos; mas a nossa capacidade vem de Deus,”

BÍBLIA, 2° Coríntios, 3,5.

Agradecimentos

Diante de sua primazia, agradeço a Deus por seu cuidado inefável e infindável para comigo, especialmente em minha trajetória acadêmica, na qual é possível perceber e sentir sua providência e fidelidade em cada detalhe.

À minha mãe Veridiana Borges da Silva, que desde o início da minha existência tem sido o meu verdadeiro exemplo, minha ajudadora e maior incentivadora dos meus estudos. Bem como à toda minha família.

Ao meu orientador, professor Gilson Mamede de Carvalho, minha contínua gratidão por toda dedicação, disponibilidade, paciência, contribuição essencial para a elaboração deste trabalho, e sobretudo, pelas trocas de conhecimentos e por todo apoio e suporte que me ofertou em minha trajetória na graduação.

Aos membros da banca; professores Eudes Mendes Barboza e Diego Ferraz de Souza, pela honra de aceitarem o convite para compor a banca deste trabalho, disponibilidade e refinadas contribuições.

À professora Yane Lísley Ramos Araújo, por ter sido minha referência de uma maravilhosa profissional desde o início do curso, pela disponibilidade, acessibilidade e formidáveis contribuições quanto à correção deste trabalho.

A todos os professores do curso que tive a oportunidade de cursar as disciplinas, por todo apoio e contribuições valiosas na vida acadêmica e pessoal. E todos os professores que passaram por minha vida, desde o ensino infantil até o ensino médio, estes, fortes contribuidores da minha admiração pelo ofício de professor.

Aos colegas de curso, pelas trocas de conhecimentos, apoio e conversas. Em especial, à minha dupla acadêmica; Ana Catarine, que tornou essa trajetória mais leve.

A todos que, direta ou indiretamente, contribuíram de forma positiva para a minha graduação. Em especial, aos que sempre perguntavam empolgados sobre o andamento do curso e me apresentavam em suas orações.

À minha estimada UFRPE juntamente com o CNPq pelo apoio financeiro durante os estudos de iniciação científica que culminaram na elaboração deste trabalho.

Resumo

Este trabalho tem como principal objetivo estudar e estabelecer relações entre o Teorema da função inversa, que é apresentado no espaço Euclidiano $\mathbb{R}^n$, com as superfícies regulares, no contexto da Geometria Diferencial. Vamos mostrar como um resultado oriundo do contexto da Análise matemática pode servir como base para introduzir um dos principais objetos de estudo da Geometria diferencial. Para essa construção, inicialmente abordaremos conceitos básicos envolvendo a topologia do espaço Euclidiano $\mathbb{R}^n$, os quais se farão presente em todo o decorrer do texto. Em seguida, vamos apresentar as noções e resultados fundamentais sobre continuidade e diferenciabilidade no espaço Euclidiano n-dimensional e por fim, vamos introduzir as superfícies regulares juntamente com alguns resultados relevantes para estabelecer uma relação natural e esperada com a topologia dos espaços Euclidianos e o Teorema da função inversa.

Palavras-Chave: Espaço Euclidiano; Diferenciabilidade; Teorema da função inversa; Superfícies regulares.

Abstract

This work's main objective is to study and establish relationships between the inverse function theorem, which is presented in the Euclidean space $\mathbb{R}^n$, with regular surfaces, in the context of Differential Geometry. We will show how a result arising from the context of mathematical Analysis can serve as a basis for introducing one of the main objects of study in Differential Geometry. For this construction, we will initially address basic concepts involving the topology of the Euclidean space $\mathbb{R}^n$, which will be present throughout the text. Next, we will present the fundamental notions and results about continuity and differentiability in the n-dimensional Euclidean space and finally, we will introduce regular surfaces together with some relevant results to establish a natural and expected relationship with the topology of Euclidean spaces and o Inverse function theorem.

KeyWords: Euclidean Space; Differentiability; Inverse function theorem; Regular surfaces.

Sumário

Lista de Notações

- $\mathbb{R}^n$ denota o produto cartesiano de n fatores iguais a $\mathbb{R}$;
- $int.X$ denota o conjunto dos pontos interiores de X;
- $\overline{X}$ denota o conjunto dos pontos aderentes a X;
- $fr.X$ denota o conjunto dos pontos de fronteira de X;
- $G(f) = \{(x, f(x)) \in X \times Y : x \in X\}$ representa o gráfico da função $f : X \to Y$;
- $d(x, y)$ indica a distância de x para y;
- $|\cdot|$ representa o valor absoluto em $\mathbb{R}$;
- $\|\cdot\|$ representa uma norma em $\mathbb{R}^n$;
- $\|\cdot\|_M$ e $\|\cdot\|_S$ representam as normas do máximo e da soma em $\mathbb{R}^n$, respectivamente;
- $B(a; r)$ denota a bola aberta de centro a e raio r;
- $B[a; r]$ denota a bola fechada de centro a e raio r;
- $S[a; r]$ denota a esfera de centro a e raio r;
- $diam.X$ denota o diâmetro de X;
- ε e δ denotam constantes positivas;
- $x_n \to a$ indica que a sequência (x_n) converge para a;
- S^{n-1} indica a esfera unitária centrada na origem em $\mathbb{R}^n$.

Introdução

Um dos principais objetivos da Análise real na matemática pura é o de buscar impor o formalismo necessário aos conceitos e ideias advindas do cálculo diferencial e integral. A mesma surgiu no século XVII e com o passar do tempo, contou com colabarações de diversos matemáticos de destaque, os quais podemos citar: Leonhard Euler que, no século XVIII introduziu o conceito de função. No século XIX, Bernard Bolzano e Cauchy, o primeiro formulou a definição moderna de continuidade de funções e o segundo introduziu o cálculo infinitesimal. Jean-Baptiste Joseph Fourier, que foi um dos primeiros a investigar o conceito de derivada para funções de várias variáveis. Além do matemático alemão Weierstrass, que buscou constantemente um formalismo matemático em seus estudos e foi o primeiro a apresentar um exemplo de uma função contínua que não possui derivada em nenhum de seus pontos.

Aqui fazemos um estudo a respeito do espaço Euclidiano $\mathbb{R}^n$, bem como de funções e aplicações definidas nesse espaço, visando entender e aplicar o Teorema da função inversa. Este importante teorema, garante, que sob certas hipóteses a respeito da derivada em um ponto, uma função é localmente inversível. O primeiro matemático a apresentar uma demonstração para este resultado foi Joseph Louis Lagrange. Embora hoje já existam formulações mais gerais para este resultado, por exemplo, o mesmo pode ser estabelecido para funções definidas em espaços de Banach, como já dito anteriormente, aqui vamos nos ater as funções definidas no espaço $\mathbb{R}^n$.

Como principais aplicações deste estudo, podemos destacar o Teorema da função implícita, apresentado primeiramente pelo matemático Francês Augustin-Louis Cauchy, o qual garante que, sob certas condições de suavidade a respeito das derivadas parciais de uma função real f, de $m+n$ variáveis satisfazendo

$$f(x_1, ..., x_m, y_1, ..., y_n) = 0,$$

podemos escrever implicitamente as n últimas variáveis em função das m primeiras. Por fim, faremos uma relação do nosso estudo com a Geometria Diferencial, mostrando que a imagem inversa de um valor regular é uma superfície regular.

O presente trabalho é composto por 4 capítulos. O capítulo 1 trata da abordagem topológica do espaço Euclidiano $\mathbb{R}^n$, nele são apresentadas suas propriedades que vão exercer papel fundamental no desenvolvimento dos capítulos posteriores. Além disso, estabelecemos os conceitos de bolas, conjuntos limitados, conjuntos abertos, sequências no espaço Euclidiano que é essencial para definir o que vem a ser os conjuntos fechados, objetivando estabelecer uma característica importante de alguns conjuntos que é a compacidade.

Posteriormente, o capítulo 2 se atém às questões de continuidade. Iniciamos com as noções de limites de funções, apresentando seu comportamento e desenvolvendo ferramentas que irão nos auxiliar a trabalhar com funções contínuas e continuidade uniforme. Munidos com esses conhecimentos, passamos a lidar com homeomorfismos e com a conexidade de conjuntos.

Em seguida, o capítulo 3 traz consigo uma boa quantidade de informações sobre diferenciabilidade. Iniciando com as derivadas parciais que se revela como uma ferramenta indispensável para definirmos funções de classe C^1, bem como o Teorema de Schwarz e a fórmula de Taylor. Atrelado a isso, definimos pontos críticos e passamos a ver a derivada como uma transformação linear para apresentar dois importantes teoremas que estão intimamente relacionados: o Teorema da função inversa e o Teorema da função implícita.

Por fim, no capítulo 4, apresentamos o que vem a ser uma superfície regular e seus principais resultados a fim de estabelecer relações com o que foi tratado nos capítulos anteriores, agora no contexto da Geometria Diferencial, especialmente através do Teorema da função inversa.

Capítulo 1

Topologia do Espaço Euclidiano

Até o século XIX, a Geometria se debruçava fundamentalmente no estudo das propriedades das curvas e superfícies, que não variavam mesmo que sofressem deformações que conservavam a distância. Diante deste cenário, alguns geômetras como Bernhard Riemann (1826–1866) identificaram que era necessário a formulação de uma teoria geométrica para abranger as deformações que não preservavam a distância, essa teoria foi nomeada de Topologia. Os objetos de estudo da Topologia são os espaços topológicos, dentre eles está o espaço Euclidiano n-dimensional, o qual abordaremos neste capítulo. Embora não seja possível visualizar geometricamente os espaços maiores que o tridimensional, estudar espaços Euclidianos de dimensões maiores é de fundamental importância para as ciências básicas, em especial, na própria matemática e física-matemática.

1.1 O Espaço Euclidiano n-dimensional

Definição 1.1. O espaço Euclidiano n-dimensional, com $n \in \mathbb{N}$, é o produto cartesiano de n fatores iguais a $\mathbb{R}$.
Notação: $\mathbb{R}^n = \mathbb{R} \times ... \times \mathbb{R}$.

Exemplo 1.2. Alguns espaços Euclidianos:

- $\mathbb{R}$ é o conjunto dos números reais;
- $\mathbb{R}^2$ é o modelo numérico do plano;
- $\mathbb{R}^3$ é o modelo do espaço Euclidiano tridimensional.

Definição 1.3. Os elementos de $\mathbb{R}^n$ são listas de n termos reais $x = (x_1, ..., x_n)$, onde o termo x_i é chamado de, i-ésima coordenada de x, para todo $i = 1, ..., n$.

Definição 1.4. Sejam $x = (x_1, ..., x_n)$ e $y = (y_1, ..., y_n)$, dizemos que $x = y$ se, e somente se, $x_1 = y_1, ..., x_n = y_n$. Ou seja, se as entradas corespondentes de x e y forem iguais.

Por convenção, em $\mathbb{R}^2$ escreveremos $z = (x, y)$ ao invés de $x = (x_1, x_2)$ e em $\mathbb{R}^3$, $w = (x, y, z)$ ao invés de $x = (x_1, x_2, x_3)$.

Por vezes, os elementos de $\mathbb{R}^n$ serão chamados de pontos, e em outras ocasiões, de vetores.

Em $\mathbb{R}^n$ estão bem definidas duas operações;

$$\begin{aligned} + : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ ((x_1, ..., x_n), (y_1, ..., y_n)) &\longmapsto (x_1 + y_1, ..., x_n + y_n) \end{aligned} \quad e \quad \begin{aligned} \cdot : \mathbb{R} \times \mathbb{R}^n &\to \mathbb{R}^n \\ (\alpha, x) &\mapsto (\alpha x_1, ..., \alpha x_n), \end{aligned}$$

chamadas de soma e produto por escalar, respectivamente, e estas operações satisfazem as seguintes propriedades:

(S1) Comutatividade: $x + y = y + x$;
(S2) Existência de elemento neutro: $x + 0 = x$;
(S3) Existência de Simétrico: $-x + x = 0$;
(S4) Associatividade: $x + (y + z) = (x + y) + z$;
(P1) Associatividade: $\alpha(\beta x){=}(\alpha\beta)x$;
(P2) Distributividade à direita: $\alpha(x + y){=}\alpha x{+}\alpha y$;
(P3) Distributividade à esquerda: $(\alpha{+}\beta)x{=}\alpha x{+}\beta x$;
(P4) Existência de elemento neutro: $1 \cdot x{=}x$.

Por isso temos que $\mathbb{R}^n$ munido com essas operações é um espaço vetorial normado. Além disso, podemos observar que $e_1, ..., e_n$, em que $e_i = (0, ..., 0, 1, 0, ..., 0)$, com 1 na i-ésima posição, é uma base para o espaço $\mathbb{R}^n$ chamada de base canônica de $\mathbb{R}^n$. Além das operações de soma e produto por escalar, está bem definida a operação;

$$\begin{aligned} \langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ ((x_1, ..., x_n), (y_1, ..., y_n)) &\longmapsto x_1 y_1 + ... + x_n y_n. \end{aligned}$$

Esta operação é um produto interno em $\mathbb{R}^n$, que satisfaz as seguintes propriedades:

(PI1) $\langle x, x \rangle \geq 0$, para todo $x \in \mathbb{R}^n$ e $\langle x, x \rangle = 0$ se, e só se $x = 0$;
(PI2) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$;
(PI3) $\langle x, y \rangle = \langle y, x \rangle$;
(PI4) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

Diz-se que dois vetores $x, y \in \mathbb{R}^n$ são ortogonais quando $\langle x, y \rangle = 0$ e denotamos isto por $x \perp y$.

Exemplo 1.5. Seja $x \in \mathbb{R}^n$ não nulo, para todo $y \in \mathbb{R}^n$, o vetor $z = y - \frac{\langle x,y \rangle}{\langle x,x \rangle} \cdot x$ é ortogonal a x.

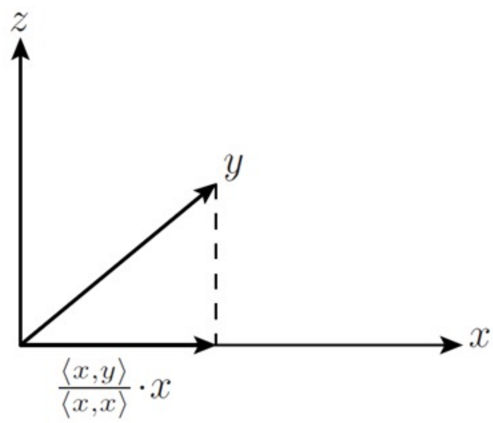
Perceba que $\dfrac{\langle x, y\rangle}{\langle x, x\rangle} \in \mathbb{R}$, e quando fazemos $\langle x, z\rangle$, temos

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

Logo, $x \perp z$.

Perceba que se $x, y \in \mathbb{R}^n$, o vetor $\alpha x = \dfrac{\langle x, y\rangle}{\langle x, x\rangle} \cdot x$ é a projeção ortogonal de y sobre x. Geometricamente, temos;

Figura 1.1: Projeção Ortogonal.



Fonte: LIMA, 2016, p. 3

Definição 1.6. A norma ou comprimento do vetor $x \in \mathbb{R}^n$ é o número $\|x\| = \sqrt{\langle x, x\rangle}$, ou seja, $\|x\| = \sqrt{x_1^2 + ... + x_n^2}$.

Um vetor $x \in \mathbb{R}^n$ é dito unitário quando $\|x\| = 1$. Para todo $x \neq 0 \in \mathbb{R}^n$, o vetor $u = \frac{x}{\|x\|}$ é unitário.

Teorema 1.7. *(Pitágoras) Se $x \perp y$, então $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.*

Demonstração: Note que

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \\ &= \langle x, x\rangle + 0 + \langle y, y\rangle \\ &= \|x\|^2 + \|y\|^2.\end{aligned}$$

■

Teorema 1.8. *(Desigualdade de Schwarz) Para quaisquer $x, y \in \mathbb{R}^n$, tem-se $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$, valendo a igualdade se, e somente se, um dos vetores é múltiplo do outro.*

Demonstração: - Se $x = 0$, temos $|\langle 0, y\rangle| = 0 = 0 \cdot \|y\| = 0$.
- Se $x \neq 0$, podemos escrever $y{=}\alpha x + z$, onde $z \perp x$ e $\alpha = \frac{\langle x,y\rangle}{\langle x,x\rangle} = \frac{\langle x,y\rangle}{\|x\|^2}$. Pelo Teorema de Pitágoras, $\|y\|^2{=}\alpha^2\|x\|^2 + \|z\|^2$, então

$$\|y\|^2 \geq \alpha^2\|x\|^2. \tag{1.1}$$

Afirmamos que a igualdade é válida se, e somente se, $y{=}\alpha x$. De fato, da desigualdade (1.1) temos que

$$\|y\|^2 \geq \frac{\langle x, y\rangle^2}{\|x\|^4}\|x\|^2.$$

Logo, $\|y\|^2 \geq \frac{\langle x,y\rangle^2}{\|x\|^2}$. Assim, $\langle x, y\rangle^2 \leq \|y\|^2\|x\|^2$, o que nos fornece $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$ e a igualdade é válida se, e somente se $y{=}\ \alpha x$. ■

PROPRIEDADES DA NORMA

Sejam $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$;

01. $\|x\| \geq 0$, valendo a igualdade apenas quando $x = 0$;

02. $\|\alpha x\| = |\alpha|\|x\|$;

03. $\|x + y\| \leq \|x\| + \|y\|$ (Desigualdade Triangular).

OBS: $\|x + y\|^2 \leq (\|x\| + \|y\|)^2$ pois,

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

Ainda podemos definir uma norma como sendo qualquer função $\mathbb{R}^n \to \mathbb{R}$, que associe cada vetor $x \in \mathbb{R}^n$ a um número $\|x\|$, satisfazendo as propriedades anteriores. Vamos estudar três tipos de normas em $\mathbb{R}^n$;

01. Norma Euclidiana: $\|x\| = \sqrt{x_1^2 + ... + x_n^2}$;

02. Norma do Máximo: $\|x\|_M{=} \max\{|x_1|, ..., |x_n|\}$;

03. Norma da Soma: $\|x\|_S = |x_1| + ... + |x_n|$.

Lema 1.9. *Se $a, b \geq 0$, então $\sqrt{a + b} \leq \sqrt{a} + \sqrt{b}$.*

Demonstração: Note que

$$\begin{aligned}\sqrt{a+b} \leq \sqrt{a}+\sqrt{b} &\Leftrightarrow a+b \leq (\sqrt{a}+\sqrt{b})^2\\ &\Leftrightarrow a+b \leq a+2\sqrt{a}\sqrt{b}+b\\ &\Leftrightarrow 0 \leq 2\sqrt{a}\sqrt{b}.\end{aligned}$$

■

Proposição 1.10. *Note que para todo $x \in \mathbb{R}^n, n \in \mathbb{N}$; $\|x\|_M \leq \|x\| \leq \|x\|_S \leq n \cdot \|x\|_M$.*

Demonstração:

- $\|x\|_M \leq \|x\|$
  Seja $j \in \{1, ..., n\}$ tal que $|x_j| = \|x\|_M$, então

  $$\|x\|_M = |x_j| = \sqrt{x_j^2} \leq \sqrt{x_1^2 + ... + x_n^2} = \|x\|.$$

  Logo, $\|x\|_M \leq \|x\|$.

- $\|x\| \leq \|x\|_S$
  Note que, pelo Lema 1.9, $\|x\| = \sqrt{x_1^2 + ... + x_n^2} \leq \sqrt{x_1^2}+...+\sqrt{x_n^2}=|x_1|+...+|x_n| = \|x\|_S$.
  Logo, $\|x\| \leq \|x\|_S$.

- $\|x\|_S \leq n \cdot \|x\|_M$

  $$\begin{aligned}\|x\|_S &= |x_1| + ... + |x_n|\\ &\leq |x_j| + ... + |x_j|\\ &= n \cdot \max_{1\leq i\leq n}\{|x_i|\}\\ &= n \cdot \|x\|_M.\end{aligned}$$

  Logo, $\|x\|_S \leq n \cdot \|x\|_M$.

■

Observação 1.11. Quando estivermos usando apenas uma das três normas, vamos apenas indicá-la por $\|x\|$.

Observação 1.12. Para todas as normas vale $|\|x\| - \|y\|| \leq \|x-y\|$, pois, podemos escrever $x \in \mathbb{R}^n$ como $x = (x-y)+y$, logo $\|x\| \leq \|x-y\|+\|y\|$, então $\|x\|-\|y\| \leq \|x-y\|$.
E analogamente $y = (y-x)+x$, $\|y\| \leq \|y-x\| + \|x\|$, $\|y\| - \|x\| \leq \|y-x\|$.
Como $\|y-x\| = \|x-y\|$, pois note que dado $\alpha \in \mathbb{R}$, temos $\|\alpha x\| = |\alpha|\|x\|$. Assim, $\|(-1)x\| = |-1|\|x\| = \|x\|$. Logo, $\|y-x\| = \|-(y-x)\| = \|x-y\|$.

Assim, segue que $\|y\| - \|x\| \leq \|x - y\|$.
Portanto, $|\|x\| - \|y\|| \leq \|x - y\|$.

De posse de uma norma em $\mathbb{R}^n$, podemos ter a distância entre dois pontos.

Definição 1.13. Sejam $x, y \in \mathbb{R}^n$, a distância $d(x, y)$ entre os pontos $x = (x_1, ..., x_n)$ e $y = (y_1, ..., y_n)$ é dada por $d(x, y) = \|x - y\|$.

Proposição 1.14. *A distância d como definida acima, satisfaz as seguintes propriedades. Sejam $x, y, z \in \mathbb{R}^n$, então*

i) $d(x, y) \geq 0$*, vale a igualdade se, e somente se,* $x = y$*;*

ii) $d(x, y) = d(y, x)$*;*

iii) $d(x, z) \leq d(x, y) + d(y, z)$*.*

Observação 1.15. Assim como definimos distância em função da norma euclidiana, podemos definir uma distância d em função das normas do máximo e da soma.

1.2 Bolas e Conjuntos Limitados

Vamos introduzir agora, a noção de bola aberta, que é de extrema importância para o conceito de conjuntos abertos e demais noções topológicas.

Definição 1.16. Sejam $a \in \mathbb{R}^n$ e $r \in \mathbb{R}$ tal que $r > 0$, a bola aberta de centro a e raio r, é o conjunto $B(a; r)$ dos pontos $x \in \mathbb{R}^n$ cuja distância ao ponto a é menor que r.
Simbolicamente: $B(a; r) = \{x \in \mathbb{R}^n; \|x - a\| < r\}$.

Definição 1.17. Sejam $a \in \mathbb{R}^n$ e $r \in \mathbb{R}$ tal que $r > 0$, a bola fechada de centro a e raio r, é o conjunto $B[a; r]$ dos pontos $x \in \mathbb{R}^n$ cuja distância ao ponto a é menor ou igual a r.
Simbolicamente: $B[a; r] = \{x \in \mathbb{R}^n; \|x - a\| \leq r\}$.

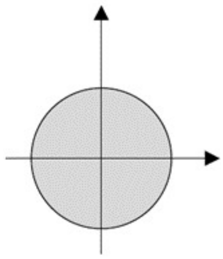
A bola fechada também é chamada de disco n-dimensional.

Definição 1.18. A esfera de centro a e raio r é o conjunto $S[a; r] = \{x \in \mathbb{R}^n; \|x - a\| = r\}$.

Note que $B[a; r] = B(a; r) \cup S[a; r]$, a bola fechada $B[0; 1]$ de centro 0 e raio 1 é chamada de disco unitário de $\mathbb{R}^n$.

A esfera unitária de dimensão n tem uma notação especial; $S^{n-1} = \{x \in \mathbb{R}^n; \|x\| = 1\}$. Portanto, S^{n-1} é a esfera de centro 0 e raio 1.
A representação geométrica dessas bolas vai depender da norma que estivermos adotando.

Figura 1.2: Representação da bola na Norma Euclidiana.



Fonte: LIMA, 2016, p. 6

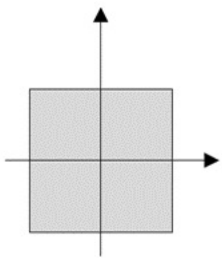
Exemplo 1.19. Norma Euclidiana:

Conjunto dos pontos $z \in \mathbb{R}^2$ tais que $\|z\| \leq 1$, ou equivalentemente $B[0;1] = \{z \in \mathbb{R}^2; \|x\| \leq 1\}$, temos o disco unitário.
Seja $(x, y) \in \mathbb{R}^2$, para que $\|(x, y)\| = \sqrt{x^2 + y^2} = 1$, temos que $x^2+y^2 = 1^2$ que é a equação da circunferência de centro $(0, 0)$ e raio 1, que pode ser vista como $(x-0)^2+(y-0)^2 = 1$.

Exemplo 1.20. Norma do Máximo:
Temos a "esfera unitária".
Seja $(x, y) \in \mathbb{R}^2$, para que $\|(x, y)\| = 1$, $\|(x, y)\| = max\{|x|, |y|\} = 1$.
-Se $x, y \geq 0$, temos que
$\max\{|x|, |y|\}$=$|x|$, então $x = 1$;
$\max\{|x|, |y|\}$=$|y|$, então $y = 1$.
-Se $x, y \leq 0$, temos que
$\max\{|x|, |y|\}$=$|x|$, então $-x = 1$, logo $x = -1$;
$\max\{|x|, |y|\}$=$|y|$, então $-y = 1$, logo $y = -1$.
-Se $x \geq 0$ e $y \leq 0$, temos que
$\max\{|x|, |y|\}$=$|x|$, então $x = 1$;
$\max\{|x|, |y|\}$=$|y|$, então $-y = 1$, logo $y = -1$.
-Se $x \leq 0$ e $y \geq 0$, temos que
$\max\{|x|, |y|\}$=$|x|$, então $-x = 1$, logo $x = -1$;
$max\{|x|, |y|\}$=$|y|$, então $y = 1$.

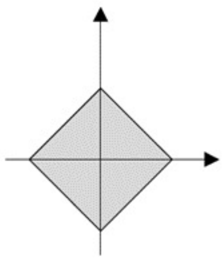
Exemplo 1.21. Norma da Soma:
Temos o "disco unitário".
Seja $(x, y) \in \mathbb{R}^2$, para que $\|(x, y)\| = 1$, $\|(x, y)\| = |x| + |y|$, logo $|x| + |y| = 1$.
- Se $x, y \geq 0$, temos que $x + y = 1$.
- Se $x, y \leq 0$, temos que $-x - y = 1$, então $-(x + y) = 1$, logo $x + y = -1$.
- Se $x \geq 0$ e $y \leq 0$, temos que $x - y = 1$.
- Se $x \leq 0$ e $y \geq 0$, temos que $-x + y = 1$.

Figura 1.3: Representação da bola na Norma do Máximo.



Fonte: LIMA, 2016, p. 6

Figura 1.4: Representação da bola na Norma da Soma.



Fonte: LIMA, 2016, p. 6

As bolas de centro a e raio r em $\mathbb{R}^n$ denotamos por B na norma Euclidiana, B_M na norma do máximo e B_S na norma da soma. Na norma do máximo, denotamos a bola de centro a e raio $\frac{r}{n}$ por B'_M.
Note que as desigualdades $\|x\|_M \leq \|x\| \leq \|x\|_S \leq n \cdot \|x\|_M$, que provamos anteriormente, implicam em $B'_M \subset B_S \subset B \subset B_M$.

De fato, perceba que

- $B'_M \subset B_S$
  Se $x \in B'_M \Rightarrow \|x - a\| \leq \frac{r}{n}$. Assim, note que

$$\begin{aligned}\|x-a\|_S &= |x_1 - a_1| + ... + |x_n - a_n| \\ &\leq \|x-a\|_M + ... + \|x-a\|_M \\ &\leq \frac{r}{n} + ... + \frac{r}{n} \\ &= n \cdot \frac{r}{n} \\ &= r.\end{aligned}$$

  Logo, $x \in B_S$.

- $B_S \subset B$

Se $x \in B_S \Rightarrow \|x-a\|_S \leq r$. Assim, note que

$$\begin{aligned}\|x-a\| &= \sqrt{(x_1-a_1)^2+...+(x_n-a_n)^2}\\ &\leq \sqrt{(x_1-a_1)^2}+...+\sqrt{(x_n-a_n)^2}\\ &= |x_1-a_1|+...+|x_n-a_n|\\ &= \|x-a\|_S\\ &\leq r.\end{aligned}$$

Logo, $x \in B$.

- $B \subset B_M$
  Se $x \in B \Rightarrow \|x-a\| \leq r$. Assim, note que

$$\begin{aligned}\|x-a\|_M &= \max\{|x_1-a_1|,...,|x_n-a_n|\}\\ &= |x_j-a_j|, \text{ para algum } j \in \{1,...,n\}\\ &\leq \sqrt{(x_1-a_1)^2+...+(x_n-a_n)^2}\\ &= \|x-a\|\\ &\leq r.\end{aligned}$$

Logo, $x \in B_M$.

Definição 1.22. Seja $X \subset \mathbb{R}^n$, dizemos que X é limitado quando $X \subset B[a;r]$, para algum $a \in \mathbb{R}^n$ e algum $r > 0$.
Equivalentemente, X é limitado se existe $k > 0$ tal que $\|x\| \leq k$, para todo $x \in X$.

Essa equivalência vem do fato que $B[a;r] \subset B[0;r+\|a\|]$, pois de $B[a;r]$ temos $\|x-a\| \leq r$. Pela desigualdade triangular;

$$\begin{aligned}\|x\| &= \|x-a+a\|\\ &\leq \|x-a\|+\|a\|\\ &\leq r+\|a\|.\end{aligned}$$

Portanto, se $x \in B[a;r]$, então $x \in B[0;r+\|a\|]$.

Estudamos também alguns conceitos referentes às aplicações definidas em espaços Euclidianos.

Definição 1.23. Dizemos que uma aplicação $f : X \to \mathbb{R}^n$ é limitada no conjunto $X \subset \mathbb{R}^m$ quando $f(X) \subset \mathbb{R}^n$ é um conjunto limitado, ou seja, quando existe $c > 0$ tal que $\|f(x)\| \leq c$, $\forall\, x \in X$.

Sejam $a, b \in \mathbb{R}^n$ tais que $a \neq b$, a reta que une esses dois pontos é o conjunto $ab = \{(1-t)a + tb; t \in \mathbb{R}\}$. E o segmento de reta com extremos a, b é o conjunto $[a, b] = \{(1-t)a + tb; 0 \leq t \leq 1\}$.

Definição 1.24. Um conjunto $X \subset \mathbb{R}^n$ é dito convexo quando o segmento de reta que une quaisquer dois de seus pontos está completamente contido em X. Ou ainda, X é convexo quando dados $a, b \in X$; $0 \leq t \leq 1$, então $(1-t)a + tb \in X$.

Exemplo 1.25. Toda bola aberta ou fechada é um conjunto convexo.
Considere a bola fechada $B = B[x_0; r]$. Sejam $a, b \in B$, temos $\|a - x_0\| \leq r$ e $\|b - x_0\| \leq r$. Note que para todo $t \in [0, 1]$, vale $x_0 = (1-t)x_0 + tx_0$, assim;

$$\begin{aligned}
\|(1-t)a + tb - x_0\| &= \|(1-t)a + tb - [(1-t)x_0 + tx_0]\| \\
&= \|(1-t)a + tb - (1-t)x_0 - tx_0)\| \\
&= \|(1-t)(a - x_0) + t(b - x_0)\| \\
&\leq \|(1-t)(a - x_0)\| + \|t(b - x_0)\| \\
&= |1-t|\|a - x_0\| + |t|\|b - x_0\| \\
&= (1-t)\|a - x_0\| + t\|b - x_0\| \\
&\leq (1-t)r + tr \\
&= r - tr + tr \\
&= r.
\end{aligned}$$

Assim, $[(1-t)a + tb] \in B[x_0; r]$, então $[a, b] \subset B[x_0; r]$. Logo, toda bola fechada é um conjunto convexo. Considerando a bola aberta, a demonstração é análoga. Portanto, toda bola aberta é um conjunto convexo.

Exemplo 1.26. O conjunto $X \subset \mathbb{R}^2$ tal que $X = \{(x, y) \in \mathbb{R}^2; y \leq x^2\}$ não é convexo.

Dados $a = (-1, 1)$ e $b = (1, 1)$, temos que $a \in X$ e $b \in X$. Agora considere $\frac{1}{2}a + \frac{1}{2}b = (\frac{-1}{2}, \frac{1}{2}) + (\frac{1}{2}, \frac{1}{2}) = (0, 1)$. Tem-se que $1 > 0$. Logo $(0, 1) \notin X$, portanto X não é convexo.

1.3 Conjuntos Abertos

Como forma de generalização dos intervalos abertos da reta real, temos os conjuntos abertos no espaço Euclidiano.

Definição 1.27. Seja $a \in X \subset \mathbb{R}^n$, dizemos que a é ponto interior do conjunto X quando para algum $r > 0$, tem-se $B(a; r) \subset X$. Ou seja, todos os pontos suficientemente próximos de a também pertencem a X.

Definição 1.28. O conjunto dos pontos interiores a X é chamado de interior de X, e denotado por $int.X$.

Observação 1.29. É imediato que $int.X \subset X$.

Exemplo 1.30. Seja $X = \{(x, y) \in \mathbb{R}^2; y \geq 0\}$ o semi-plano superior fechado. Se $p = (a, b)$ com $b > 0$, então $p \in int.X$.

Considere a bola $B = B(p; b)$ e seja $(x, y) \in B$, então;

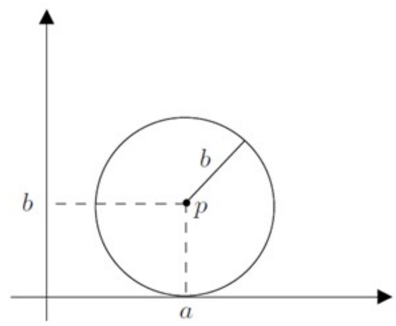
$$b > \|(x, y) - p\| = \|(x, y) - (a, b)\| = \sqrt{(x - a)^2 + (y - b)^2}.$$

Disto temos que $(x - a)^2 + (y - b)^2 < b^2$, em particular, $(y - b)^2 < b^2$. De onde obtemos que

$$y^2 - 2yb < 0 \quad \Rightarrow \quad 0 < y < 2b.$$

Logo, $(x, y) \in B(p; b)$ e $(x, y) \in X$, logo $B(p; b) \subset X$.
Geometricamente,

Figura 1.5: Representação da bola inteiramente contida no semi-plano superior fechado X.



Fonte: LIMA, 2016, p. 8

Exemplo 1.31. Seja $X = \{(x, y) \in \mathbb{R}^2; y \geq 0\}$, os pontos da forma $q = (a, 0)$ pertencem a X, porém não são interiores a X.

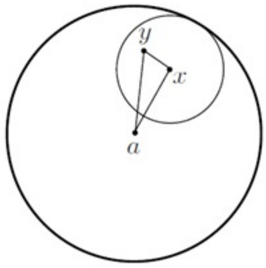
Seja $r > 0$ e vejamos que $B(q; r) \not\subseteq X$. De fato, observe que $p = \left(a, -\frac{r}{2}\right) \in B(q, r)$ pois

$$\|p - q\| = \left\|\left(0, \frac{r}{2}\right)\right\| = \sqrt{\left(\frac{r}{2}\right)^2} = \frac{r}{2} < r.$$

Porém $p \notin X$.

Definição 1.32. Seja $A \subset \mathbb{R}^n$, dizemos que A é um conjunto aberto se todos os seus pontos são interiores, ou seja, se $A = int.A$.

Figura 1.6: Representação de conjunto aberto



Fonte: LIMA, 2016, p. 8

Exemplo 1.33. Toda bola aberta $B = B(a;r)$ é um conjunto aberto.

Seja $x \in B = B(a;r)$, então $\|x-a\| < r \Rightarrow r - \|x-a\| > 0$.
Seja $s = r - \|x-a\|$ e considere a bola $B(x;s)$ e $y \in B(x;s)$, então

$$\begin{aligned}\|y-a\| &= \|y-x+x-a\| \\ &\leq \|y-x\| + \|x-a\| \\ &< s + \|x-a\| = r.\end{aligned}$$

Temos que $y \in B(a;r)$, portanto, $B(x;s) \subset B(a;r)$.

Definição 1.34. Seja $X \subset \mathbb{R}^n$, a fronteira de X é o conjunto formado pelos pontos de $\mathbb{R}^n$ que não são interiores a X, com os pontos de $X^c = \mathbb{R}^n \backslash X$ que não são interiores a X^c.
Notação: $fr.X$.
Ou seja, $x \in fr.X$ quando toda bola de centro x contém pontos de X e pontos de X^c.

Exemplo 1.35. Seja $X = \{(x,y) \in \mathbb{R}^2; y \geq 0\}$, todo ponto de $X^c = \{(x,y) \in \mathbb{R}^2; y < 0\}$ é ponto interior de X^c. Logo, X^c é um conjunto aberto. Portanto, nenhum ponto de X^c pode estar na fronteira de X. Logo, $fr.X = \{(x,0); x \in \mathbb{R}\} = fr.X^c$.

Portanto $y < 0$, assim $(x,y) \in X$. Logo, X^c é aberto. Como essa bola arbitrária só tem pontos de X^c, segue que não há nenhum ponto de X^c em $fr.X$ e como $X^c = \{(x,y) \in \mathbb{R}^2; y < 0\}$, é imediato que $fr.X = \{(x,0); x \in \mathbb{R}\} =$ eixo dos x.

Veremos agora, propriedades importantes a respeito dos conjuntos abertos.

Teorema 1.36. *i) Se A_1, A_2 são abertos em $\mathbb{R}^n$, então $A_1 \cap A_2$ é aberto.*

ii) Se $(A_\lambda)_\lambda \in L$ é uma família arbitrária de conjuntos abertos $A_\lambda \subset \mathbb{R}^n$, então a união $A = \cup_{\lambda \in L} A_\lambda$ é um conjunto aberto.

Demonstração:

i) Seja $a \in (A_1 \cap A_2)$, então $a \in A_1$ e $a \in A_2$. Como A_1 e A_2 são abertos, segue que para algum $\varepsilon_1, \varepsilon_2 > 0$ temos $B(a, \varepsilon_1) \subset A_1$ e $B(a, \varepsilon_2) \subset A_2$. Considere $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$, então $B(a, \varepsilon) \subset A_1$ e $B(a, \varepsilon) \subset A_2$. Logo, $B(a, \varepsilon) \subset (A_1 \cap A_2)$. Assim, todo ponto $a \in (A_1 \cap A_2)$ é ponto interior, logo, $A_1 \cap A_2$ é aberto.

ii) Seja $a \in A = \cup_{\lambda \in L} A_\lambda$, então existe $\lambda \in L$ tal que $a \in A_\lambda$. Como A_λ é uma família arbitrária de conjuntos abertos, segue para algum $\varepsilon > 0$ temos $B(a, \varepsilon) \subset A_\lambda \subset A$. Assim, todo ponto $a \in A = \cup_{\lambda \in L} A_\lambda$ é ponto interior, logo, $\cup_{\lambda \in L} A_\lambda$ é aberto.

■

Segue como consequência de *i)* que dados $A_1, ..., A_n$ conjuntos abertos, então $A = \cap_{i=1}^{n} A_i$ é aberto.

Observação 1.37. A interseção de infinitos conjuntos abertos pode não ser aberta.

Exemplo 1.38. $\cap_{k=1}^{\infty} B\left(a; \frac{1}{k}\right) = \{a\}$.

Como $a \in B\left(a, \frac{1}{k}\right)$, para todo $k \in \mathbb{N}$, basta provar que $B = \cap_{k=1}^{\infty} B\left(a; \frac{1}{k}\right) \subset \{a\}$. Suponha por contradição que existe $x \neq a$ tal que $x \in B$, então $x \in B\left(a; \frac{1}{k}\right)$, para todo $k \in \mathbb{N}$. Como $x \neq a$, $r = \frac{\|x-a\|}{10} > 0$, daí $\frac{1}{k} > \|x - a\| = 10r$; para todo $k \in \mathbb{N}$, logo $k < \frac{1}{10r}$; para todo $k \in \mathbb{N}$, então $\mathbb{N}$ seria limitado superiormente, o que é uma contradição. Portanto, segue que $B = \cap_{k=1}^{\infty} B\left(a; \frac{1}{k}\right) = \{a\}$.
Note que a não é interior a $\{a\}$, pois como não existe $r > 0$ tal que $B(a; r) \subset \{a\}$. Assim, segue que $\{a\}$ não é aberto.

Definição 1.39. Sejam $X \subset \mathbb{R}^n$ e $A \subset X$, dizemos que o subconjunto A é aberto em X, quando cada ponto $a \in A$ é centro de uma bola aberta $B(a; r)$, tal que $B(a; r) \cap X \subset A$. Ou seja, os pontos de X que estão suficientemente próximos de cada $a \in A$, pertencem a A também. A é conhecido como aberto relativo de X.

Observação 1.40. $A \subset X$ é aberto em X se, e somente se, $A = U \cap X$, onde U é aberto em $\mathbb{R}^n$.
($\Rightarrow$) Se $u \in A$, então $u \in X$, visto que $A \subset X$. Considerando $U = \cup_{a \in U} B(a; r)$, pelo Teorema 1.36 a união de abertos é aberto, então U é aberto em $\mathbb{R}^n$. Logo, $a \in U \cap X$. Se $u \in U \cap X$, então $u \in U$ e $u \in X$, como U é um aberto de $\mathbb{R}^n$, $u \in B(a; r_0)$, para algum $a \in A$. Como $B(a; r_0) \cap X \subset A$, então $u \in A$.
($\Leftarrow$) Se $A = U \cap X$, então para todo $a \in A$, temos que $a \in U$ e $a \in X$. Por U ser um aberto de $\mathbb{R}^n$, para algum $r > 0$ existe $B(a; r) \subset U$. Assim, $B(a; r) \cap X \subset A$. Portanto, A é um aberto em X.

Exemplo 1.41. Seja $\left(\frac{1}{2}, 1\right]$, note que $\left(\frac{1}{2}, 1\right] = \left(\frac{1}{2}, 2\right) \cap [0, 1]$. Logo, $\left(\frac{1}{2}, 1\right]$ é aberto em $[0, 1]$.

1.4 Sequências em $\mathbb{R}^n$

De posse do que vem a ser uma norma, bolas e conjuntos abertos em $\mathbb{R}^n$, nos é pertinente abordar as sequências neste espaço e ampliar os resultados que são vistos em $\mathbb{R}$.

Definição 1.42. Uma sequência em $\mathbb{R}^n$ é uma função $x : \mathbb{N} \to \mathbb{R}^n$, que associa a cada número natural k, um ponto $x_k \in \mathbb{R}^n$.
Notação: $(x_1, ..., x_k, ...)$, $(x_k)_{k\in\mathbb{N}}$ ou (x_k).

Assim, se (x_k) é uma sequência, para cada $k \in \mathbb{N}$, temos que $x_k = (x_{k_1}, ..., x_{k_n})$. Perceba que uma sequência em $\mathbb{R}^n$ dá origem a n sequências de números reais.

Dizemos que uma sequência $(x_k)_{k\in\mathbb{N}}$ é limitada quando existe uma bola em $\mathbb{R}^n$ que contém todos os termos x_k. Equivalentemente, $(x_k)_{k\in\mathbb{N}}$ é limitada se existe $c > 0$ tal que $\|x_k\| \leq c$, para todo $k \in \mathbb{N}$.

Observação 1.43. Uma sequência limitada independe da norma que estamos adotando.

Proposição 1.44. *Se (x_k) é uma sequência limitada, então $\forall\ i = 1, ..., n$, as sequências $(x_{k_i})_{k\in\mathbb{N}}$ das i-ésimas coordenadas de (x_k) também são limitadas, pois $|x_{k_i}| \leq \|x_k\|, \forall i = 1, ..., n$.*

Demonstração: Considerando a norma do máximo em $\mathbb{R}^n$, como $(x_{k_i})_{k\in\mathbb{N}}$ é limitada $\forall$ $i = 1, ..., n$, então $\exists\ c_1, ..., c_n > 0$ tais que $|x_{k_1}| \leq c_1$, $|x_{k_2}| \leq c_2$, ..., $|x_{k_n}| \leq c_n$; $\forall\ k \in \mathbb{N}$. Seja $c = max\{c_1, ..., c_n\}$, então $\|x_k\| = max\{|x_{k_1}|, ..., |x_{k_n}|\} \leq c$, $\forall\ k \in \mathbb{N}$. Logo, como cada $(x_{k_i})_{k\in\mathbb{N}}; i = 1, ..., n$, é limitada, segue que a sequência $(x_k)_{k\in\mathbb{N}}$ é limitada. ■

Observação 1.45. A recíproca é verdadeira.
Pois se a sequência $(x_{k_i})_{k\in\mathbb{N}}$ para todo $i = 1, ..., n$ das i-ésimas coordenadas de (x_k) é limitada, então (x_k) também é limitada.

Definição 1.46. Uma subsequência de $(x_k)_{k\in\mathbb{N}}$ é a restrição da sequência a um subconjunto infinito $\mathbb{N}' = \{k_1 < ... < k_m < ...\} \subset \mathbb{N}$.
Notação: $(x_k)_{k\in\mathbb{N}'}$, $(x_{k_m})_{m\in\mathbb{N}}$ ou $(x_{k_1}, ..., x_{k_m}, ...)$.

Definição 1.47. Seja $a \in \mathbb{R}^n$, dizemos que o ponto a é o limite da sequência (x_k) se para todo $\varepsilon > 0$ arbitrário, podemos obter $k_0 \in \mathbb{N}$ tal que $k > k_0$ implica $\|x_k - a\| < \varepsilon$. Ou seja, se $k > k_0$, $x_k \in B(a; \varepsilon)$. Neste caso, Dizemos que uma sequência $(x_k)_{k\in\mathbb{N}}$ é convergente se existe $a \in \mathbb{R}^n$ tal que $\lim x_k = a$.
Notação: $\lim\limits_{k\to\infty} x_k = a$, $\lim x_{k_{k\in\mathbb{N}}} = a$, $\lim x_k = a$ ou $x_k \to a$.

Em consequência desta definição, $\lim x_k = a$ se, e somente se, $\lim \|x_k - a\| = 0$. Afirmar que $\lim x_k = a$, quer dizer que qualquer bola de centro a, contém todos os x_k, exceto possivelmente, um número finito de seus pontos, que são $x_1, x_2, ..., x_{k_0}$.

Proposição 1.48. *Toda sequência convergente é limitada.*

Demonstração: Seja (x_k) uma sequência tal que $x_k \to a$. Dado $\varepsilon = 1$, existe $k_0 \in \mathbb{N}$ tal que $(x_k) \subset B(a; \varepsilon)$ para todo $k \geq k_0$. Então $(x_k) \subset B(a; 1) \cup \{x_1, ..., x_{k_0}\}$, como a união de conjuntos limitados é limitado, obtemos que (x_k) é limitada. ∎

Proposição 1.49. *Toda subsequência de uma sequência convergente é convergente e tem o mesmo limite da sequência.*

Demonstração: Como $x_k \to a$ e $(x_{k_n}) \subset (x_k)$, temos que dado $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ de modo que se $k > k_0$, temos $\|x_k - a\| < \varepsilon$. Sendo $\mathbb{N}' = \{k_1, k_2, ..., k_n, ...\} \subseteq \mathbb{N}$ o subconjunto que define os índices de (x_{k_n}), considere $k_{n_0} \in \mathbb{N}'$ tal que $k_{n_0} \geq k_0$, obtemos então $\|x_{k_n} - a\| < \varepsilon$, $\forall\ k_n \geq k_{n_0}$. Logo, $x_{k_n} \to a$. ∎

Observação 1.50. Embora a definição de limite envolva uma norma, como são válidas as desigualdades $\|x\|_M \leq \|x\| \leq \|x\|_S \leq n \cdot \|x\|_M$, a existência e o valor do limite independe de qual das três normas estamos adotando em $\mathbb{R}^n$.

Teorema 1.51. *A sequência (x_k) em $\mathbb{R}^n$ converge para o ponto $a = (a_1, ..., a_n)$ se, e somente se, $\forall\ i = 1, ..., n$, tivermos que $\lim\limits_{k\to\infty} x_{k_i} = a_i$. Ou seja, cada coordenada de x_k converge para a coordenada correspondente de a.*

Demonstração: $(\Rightarrow)$ Como $|x_{k_i}| \leq \|x_k\|\ \forall i = 1, ..., n$, temos $|x_{k_i} - a_i| \leq \|x_k - a\|$, assim, se $\lim x_k = a$ então, $\lim\limits_{k\to\infty} x_{k_i} = a_i$.
$(\Leftarrow)$ Se $\lim\limits_{k\to\infty} x_{k_i} = a_i$, então dado $\varepsilon > 0$, existem $k_1, ..., k_n \in \mathbb{N}$ tais que $k > k_i \Rightarrow |x_{k_i} - a_i| < \varepsilon; i = 1, ..., n$. Seja $k_0 = max\{k_1, ..., k_n\}$, adotando a norma do máximo, temos que $k > k_0 \Rightarrow \|x_k - a\| < \varepsilon$, logo, $\lim x_k = a$. ∎

Proposição 1.52. *Sejam $(x_k), (y_k) \subset \mathbb{R}^n$, e $(\alpha_k) \subset \mathbb{R}$, tais que $x_k \to a$, $y_k \to b$ e $\alpha_k \to \alpha$. Então;*

i) $x_k + y_k \to a + b$.

ii) $\alpha_k \cdot x_k \to \alpha \cdot a$.

iii) $\langle x_k, y_k\rangle \to \langle a, b\rangle$.

iv) $\|x_k\| \to \|a\|$.

Demonstração: Pelos resultados conhecidos em $\mathbb{R}$,

i) $x_k + y_k = (x_{k_1} + y_{k_1}, ..., x_{k_n} + y_{k_n})$
Assim,

$$\begin{aligned} \lim x_k + y_k &= \lim(x_{k_1} + y_{k_1}, ..., x_{k_n} + y_{k_n}) \\ &= (a_1 + b_1, ..., a_n + b_n) \\ &= (a_1, ..., a_n) + (b_1, ..., b_n) \\ &= a + b. \end{aligned}$$

ii) $\alpha_k \cdot x_k = (\alpha_k x_{k_1}, ..., \alpha_k x_{k_n})$

Assim,

$$\begin{aligned} \lim \alpha_k \cdot x_k &= \lim(\alpha_k x_{k_1}, ..., \alpha_k x_{k_n}) \\ &= (\alpha_k a_1, ..., \alpha_k a_n) \\ &= \alpha(a_1, ..., a_n) \\ &= \alpha a. \end{aligned}$$

iii) Note que;

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

iv) Pelo item anterior, temos $\lim \langle x_k, x_k \rangle = \langle a, a \rangle$. Então,

$$\|x_k\|^2 \to \|a\|^2 \quad \Rightarrow \quad \|x_k\| \to \|a\|.$$

■

A seguir, veremos um dos principais resultados a respeito das sequências em $\mathbb{R}^n$, o Teorema de Bolzano-Weierstrass.

Teorema 1.53. *(Bolzano-Weierstrass) Toda sequência limitada em $\mathbb{R}^n$ possui uma subsequência convergente.*

Demonstração: Seja (x_k) uma sequência limitada em $\mathbb{R}^n$. Note que as primeiras coordenadas dos termos de (x_k) formam uma sequência limitada $(x_{k_1})_{k \in \mathbb{N}}$ de números reais, pois de maneira geral, como (x_k) é limitada $\forall\ i = 1, ..., n$, então a sequência $(x_{k_i})_{k \in \mathbb{N}}$ das i-ésimas coordenadas de (x_k) também é limitada. Como $(x_{k_1})_{k \in \mathbb{N}}$ é limitada em $\mathbb{R}$, o Teorema de Bolzano Weierstrass na reta, garante que $(x_{k_1})_{k \in \mathbb{N}}$ possui uma subsequência convergente. Assim, $\exists\ \mathbb{N}_1 \subset \mathbb{N}$ infinito e $a_1 \in \mathbb{R}$ tal que $\lim\limits_{k \in \mathbb{N}_1} x_{k_1} = a_1$. Além disso, a sequência $(x_{k_2})_{k \in \mathbb{N}_1}$ é limitada em $\mathbb{R}$. Logo $\exists\ \mathbb{N}_2 \subset \mathbb{N}_1 \subset \mathbb{N}$ e $a \in \mathbb{R}$ tal que $\lim\limits_{k \in \mathbb{N}_2} x_{k_2} = a_2$. Prosseguindo dessa forma, de modo que se obtenham n conjuntos tais que $\mathbb{N}_n \subset ... \subset \mathbb{N}_2 \subset \mathbb{N}_1 \subset \mathbb{N}$ e $a_1, ..., a_n \in \mathbb{R}$ tais que $\lim\limits_{k \in \mathbb{N}_i} x_{k_i} = a_i$, $\forall\ i = 1, ..., n$. Considere $a = (a_1, ..., a_n)$, logo, pelo Teorema 1.51, como $\lim\limits_{k \in \mathbb{N}_i} x_{k_i} = a_i$, então $\lim\limits_{k \in \mathbb{N}_n} x_k = a$. ■

Definição 1.54. Uma sequência de pontos $x_k \in \mathbb{R}^n$ é uma sequência de Cauchy quando para todo $\varepsilon > 0$ dado, existe $k_0 \in \mathbb{N}$ tal que $k, l > k_0$ implica $\|x_k - x_l\| < \varepsilon$.

Proposição 1.55. *Toda sequência de Cauchy (x_k) é limitada.*

Demonstração: Dado $\varepsilon = 1$, existe $k_0 \in \mathbb{N}$ tal que $\|x_k - x_l\| < 1$, para todos $k, l \geq k_0$. Em particular, $\|x_k - x_{k_0}\| < 1$, $\forall\ k \geq k_0$. Assim, $x_k \in B(x_{k_0}, 1) \cup \{x_1, ..., x_{k_0}\}$ para todo $k \in \mathbb{N}$. Logo, (x_k) é limitada. ■

Podemos dizer que a sequência (x_k) é de Cauchy quando $\lim\limits_{k,l\to\infty} \|x_k - x_l\| = 0$, ou seja $\lim\limits_{k,l\in\mathbb{N}} \|x_k - x_l\| = 0$. Logo, se $\mathbb{N}' \subset \mathbb{N}$ é um subconjunto infinito, ou seja, se $(x_l)_{l\in\mathbb{N}'}$ é uma subsequência de (x_k), então $\lim\limits_{k\in\mathbb{N}, l\in\mathbb{N}'} \|x_k - x_l\| = 0$.

Teorema 1.56. *(Critério de Cauchy) Uma sequência em $\mathbb{R}^n$ converge se, e somente se, é de Cauchy.*

Demonstração: ($\Rightarrow$) Seja (x_k) uma sequência convergente em $\mathbb{R}^n$. Logo, existe $a \in \mathbb{R}^n$ tal que $\lim\limits_{k,l\to\infty} x_k = a$. Seja $(x_l)_{l\in\mathbb{N}'}$ uma subsequência de (x_k), note que;

$$\begin{aligned}\|x_k - x_l\| &= \|x_k - a + a - x_l\| \\ &\leq \|x_k - a\| + \|a - x_l\| \\ &= \|x_k - a\| + \|x_l - a\|.\end{aligned}$$

Assim, dado $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que se $k \geq k_0$, temos $\|x_k - a\| < \frac{\varepsilon}{2}$. Logo, se $k, l \geq k_0$, temos $\|x_k - x_l\| < \varepsilon$.

($\Leftarrow$) Seja (x_k) uma sequência de Cauchy em $\mathbb{R}^n$. Logo, (x_k) é limitada. Assim, pelo Teorema de Bolzano-Weierstrass, (x_k) possui uma subsequência $(x_l)_{l\in\mathbb{N}'}$ convergente. Daí, existe $a \in \mathbb{R}^n$ tal que $\lim\limits_{l\in\mathbb{N}'} x_l = a$. Portanto, $\lim\limits_{l\in\mathbb{N}'} \|x_l - a\| = 0$ e $\lim\limits_{k\in\mathbb{N}, l\in\mathbb{N}'} \|x_k - x_l\| = 0$.
Note que,

$$\begin{aligned}\|x_k - a\| &= \|x_k - x_l + x_l - a\| \\ &\leq \|x_k - x_l\| + \|x_l - a\|.\end{aligned}$$

Assim, dado $\varepsilon > 0$, existem $l_0, k_0 \in \mathbb{N}$ tais que se $k, l \geq k_0$, temos $\|x_k - x_l\| < \frac{\varepsilon}{2}$ e se $l \geq l_0$, $\|x_l - a\| < \frac{\varepsilon}{2}$. Se $k \geq max\{k_0, l_0\}$, temos $\|x_k - a\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, o que nos diz que (x_k) é convergente. ■

1.5 Conjuntos Fechados

Definição 1.57. Sejam $X \subset \mathbb{R}^n$ e $a \in \mathbb{R}^n$, dizemos que a é ponto aderente ao conjunto X, quando existe uma sequência de pontos $(x_k) \subset X$ tais que $\lim x_k = a$.

Definição 1.58. Seja $X \subset \mathbb{R}^n$, o conjunto de todos os pontos aderentes a X é chamado de fecho de X. Denotamos tal conjunto por $\overline{X}$.

Definição 1.59. Seja $F \subset \mathbb{R}^n$, dizemos que F é um conjunto fechado se $F = \overline{F}$. Ou seja, F é fechado se o limite de toda sequência convergente de pontos de F, ainda é um ponto de F.

Observação 1.60. Todo ponto $b \in X$ é aderente a X pois é limite da sequência constante $(b, b, ...)$. Assim, $X \subset \overline{X}$, para todo $X \subset \mathbb{R}^n$.

Observação 1.61. Note que se $X \subset Y$, então $\overline{X} \subset \overline{Y}$.

Exemplo 1.62. O fecho da bola aberta é a bola fechada.
Com efeito, seja $\|x\| = r$, se $y \in \mathbb{R}^n$ tal que $y \in B(0; r)$, então $\|y\| < r$. Perceba que como $\|x\| = r$, temos que $x \notin B(0; r)$.
Agora, considere (x_k) uma sequência em $\mathbb{R}^n$, onde $(x_k) = \left(1 - \frac{1}{k}\right) x$; $\forall\, k \in \mathbb{N}$. Note que; $\|x_k\| = \left\|1 - \frac{1}{k}\right\| \cdot \|x\| = \left|1 - \frac{1}{k}\right| \cdot r < r$; $\forall\, k \in \mathbb{N}$ e

$$\lim_{k\to\infty} x_k = \lim_{k\to\infty} \left(1 - \frac{1}{k}\right) x = \lim_{k\to\infty} x - \lim_{k\to\infty} \frac{1}{k} \cdot x = x - 0 = x.$$

Assim, $x_k \in B$ e $\lim x_k = x$, então $x \in \overline{B(0; r)}$.
Reciprocamente, seja $x \in \overline{B(0; r)}$, então existe $(x_k) \subset B(0; r)$ tal que $x_k \to x$. Suponha, por contradição, que $x \notin B[0; r]$, então $\|x\| = d > r$. Já vimos que como $x_k \to x$, então $\|x_k\| \to \|x\|$ e como $\|x_k\| \leq r$; $\forall\, k \in \mathbb{N}$, temos $d = \|x\| = \lim \|x_k\| \leq r$, ou seja $d \leq r$, o que é um absurdo.

Teorema 1.63. *a) 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.*

b) Um conjunto $F \subset \mathbb{R}^n$ é fechado se, e somente se, seu complementar F^c é aberto. Ou equivalentemente, $A \subset \mathbb{R}^n$ é aberto se, e somente se, A^c é fechado.

c) O fecho de qualquer conjunto $X \subset \mathbb{R}^n$ é fechado. Ou equivalentemente, $\forall\, X \subset \mathbb{R}^n$ tem-se $\overline{\overline{X}} = \overline{X}$

Demonstração:

a) $(\Rightarrow)$ Sejam $X \subset \mathbb{R}^n$ e $a \in \mathbb{R}^n$ com a aderente a X. Então existe $(x_k) \subset X$ tal que $\lim x_k = a$. Assim, $x_k \in B(a; r)$ qualquer, para todo $k \in \mathbb{N}$ suficientemente grande.
$(\Leftarrow)$ Considere que toda bola de centro a contém todos os pontos de X. Então, para cada $k \in \mathbb{N}$, podemos escolher um ponto $x_k \in X$ que esteja na bola $B\left(a; \frac{1}{k}\right)$, isto é $\|x_k - a\| < \frac{1}{k}$. Assim; $\lim x_k = a$. O que caracteriza a como ponto aderente de X.

b) As afirmações são equivalentes;

01. F é fechado.

02. Se $x \in F^c$, então x não é aderente a F.

03. Se $x \in F^c$, então $\exists\, r > 0$ tal que $B(x; r) \subset F^c$.

04. F^c é aberto.
(1 $\Rightarrow$ 2) Dado F fechado, então $F = \overline{F}$. Seja $x \in F^c$, isto é, $x \notin \overline{F}$. Logo, x não é aderente a F.
(2 $\Rightarrow$ 3) Dado $x \in F^c$ e $x \notin \overline{F}$, então existe $r > 0$ tal que $B(x;r) \cap F = \varnothing$; $B(x;r) \subset F^c$. Ou seja, x é ponto interior de F^c.
(3 $\Rightarrow$ 4) Dado $x \in F^c$ e $r > 0$ tal que $B(x;r) \subset F^c$, então x é ponto interior de F^c. Como x é arbitrário, segue que todos os pontos de F^c são interiores. Logo, F^c é aberto.
(4 $\Rightarrow$ 1) Seja $a \in \overline{X}$, então existe $(x_k) \subset X$ tal que $\lim x_k = a$. Suponha, por contradição, que $a \notin X$. Então $a \in X^c$ e por 04., existe $B(a;r) \subset X^c$ onde $(x_k) \not\subset B(a;r)\ \forall\ k \in \mathbb{N}$. Tome $\varepsilon = \frac{r}{2}$, então existe k_0 tal que $\|x_k - a\|$, para todo $k > k_0$. Assim, $\|x_k - a\| < \frac{r}{2}$. Logo, $x_k \in B(a; \frac{r}{2}) \subset B(a;r)$, o que é um absurdo, pois supomos que não existiam pontos de X em $B(a;r)$. Assim, $a \in \overline{X}$ e $a \in X$, logo X é fechado.

c) Seja $Y = \overline{X}$ e $\overline{Y} = \overline{\overline{X}}$. Já temos $Y \subset \overline{Y}$, basta provar que $\overline{Y} \subset Y$. Seja $x \in \overline{Y}$, então existe $y_k \in Y$ tal que $\lim y_k = x$. Agora, para cada $k \in \mathbb{N}$ existe $(x_{k_r}) \subset X$ tal que $\lim x_{k_r} = y_k$. Assim, $y_k \to x$ e $x_{k_r} \to y_k$. Dado $\varepsilon > 0$, existem $k_0 \in \mathbb{N}$ e $k_{r_0} \in \mathbb{N}$ tais que $\|y_k - x\| < \frac{\varepsilon}{2}$, $\forall\ k \geq k_o$ e $\|x_{k_r} - y_k\| < \frac{\varepsilon}{2}$, $\forall\ k_r \geq k_{r_0}$, $x_{k_r} \in X$. Defina z_k; como $z_1 = x_{1_{r_{0+1}}}$, $z_2 = x_{2_{r_{0+1}}}$, $z_3 = x_{3_{r_{0+1}}}$,..., $z_k = x_{k_{r_{0+1}}}$. Note que;

$$\begin{aligned} \|z_k - x\| &\leq \|z_k - y_k\| + \|x - y_k\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Assim, $z_k \to x$, $(z_k) \subset X$, $x \in \overline{X}$. Logo, $x \in Y$.

■

Observação 1.64. Existem conjuntos $X \subset \mathbb{R}^n$ que não são fechados e nem abertos e outros que são simultaneamente fechados e abertos.

Exemplo 1.65. Seja $X = B(a;r) \cup \{b\}$, com $\|b - a\| = r$.

- X não é aberto;
  Note que $b \in S(a;r)$, visto que $\|b - a\| = r$. Logo X não é aberto, pois possui um ponto b que está na fronteira.

- X não é fechado;
  Seja $z \neq b$ tal que $\|z - a\| = r$, onde $z \in \overline{X}$; $\overline{B(a;r)} = B[a;r]$. Note que $B(a;r) \subset X$, então $\overline{B(a;r)} \subset \overline{X}$, logo $B[a;r] \subset \overline{X}$. Assim, $z \in \overline{X}$ e $z \notin X$. Portanto $X \neq \overline{X}$. Então, X não é fechado. Assim, X não é fechado e nem aberto.

Exemplo 1.66. $\mathbb{R}^n$ é aberto e fechado simultaneamente.

- Seja $x \in \mathbb{R}^n$, temos que $\mathbb{R}^n \subset \overline{\mathbb{R}^n}$ e $\overline{\mathbb{R}^n} \subset \mathbb{R}^n$, logo $\mathbb{R}^n$ é fechado, pois $\mathbb{R}^n = \overline{\mathbb{R}^n}$.

- Dado $x \in \mathbb{R}^n$, $B(x;1) \subset \mathbb{R}^n$. Logo, $\mathbb{R}^n$ é aberto.
  Portanto, $\mathbb{R}^n$ é fechado e aberto.

Proposição 1.67. $\overline{S[a;r]} = S[a;r]$.

Demonstração: Como $B(a;r)$ é aberta e $\mathbb{R}^n \backslash B[a;r]$ é aberto, então $A = B(a;r) \cup (\mathbb{R}^n \backslash B[a;r])$ é aberto. Daí, note que $\mathbb{R}^n - A = S[a;r]$, então $S[a;r]$ é fechada. ∎

Definição 1.68. A distância do ponto $a \in \mathbb{R}^n$ ao conjunto $X \subset \mathbb{R}^n$ é o número $d(a;X) = inf\{\|x-a\|; x \in X\}$.

Decorre da definição de ínfimo que para todo $k \in \mathbb{N}$, existe um ponto $x_k \in X$ tal que $d(a;X) \leq \|x_k - a\| < d(a;X) + \frac{1}{k}$. Em particular, $\lim\limits_{k\to\infty} \|x_k - a\| = d(a;X)$.

Observação 1.69. Se F for fechado, então a distância de um ponto $a \in \mathbb{R}^n$ a F é atingida. Isto é, existe $x_0 \in F$ tal que $d(a,F) = \|a - x_0\|$. Com efeito, como vimos acima, existe $(x_k) \subset F$ tal que $\lim \|x_k - a\| = d(a,F)$, daí $\|x_k - a\|$ é limitada e por $\|x_k\| - \|a\| \leq \|x_k - a\|$, $\forall\ k \in \mathbb{N}$, obtemos que (x_k) é limitada. Logo, pelo Teorema de Bolzano-Weierstrass, existe $x_0 \in F$ e $(x_{k_n}) \subset (x_k)$ tal que $x_{k_n} \to x_0$. Assim, $d(F,a) = \lim \|x_k - a\| = \|x_{k_n} - a\| = \|x_0 - a\|$.

Definição 1.70. Se $X \subset Y \subset \mathbb{R}^n$, dizemos que X é denso em Y, quando $\overline{X} = Y$.

Exemplo 1.71. $B(a;r)$ é denso em $B[a;r]$.

Pois $B(a;r) \subset B[a;r]$ e o fecho de toda bola aberta é a própria bola fechada.

Exemplo 1.72. $\mathbb{Q}^n$ é denso em $\mathbb{R}^n$.

Definição 1.73. Seja $a \in \mathbb{R}^n$, dizemos que a é ponto de acumulação de $X \subset \mathbb{R}^n$, quando toda bola de centro a contém algum ponto de X diferente de a. Ou seja, quando $a \in \overline{X \backslash \{a\}}$.

Observação 1.74. Um ponto de acumulação de X pode pertencer ou não a X.

Observação 1.75. Todo ponto de acumulação é ponto aderente.

Baseada na definição de ponto de acumulação, temos a definição de ponto isolado.

Definição 1.76. Seja $a \in X$, dizemos que a é um ponto isolado de X, quando $\exists\ r > 0$ tal que $B(a;r) \cap X = \{a\}$.

Observação 1.77. Se $a \in X$ não é ponto de acumulação de X, então a é ponto isolado de X.

Definição 1.78. Dizemos que X é um conjunto discreto, quando todos os seus pontos são isolados.

Exemplo 1.79. Todos os pontos de uma bola aberta são pontos de acumulação.

Seja $B(a;r) \subset \mathbb{R}^n$, considere $y \in B(a;r)$ então, $\|y-a\| = d < r$. Vejamos que y é ponto de acumulação. Seja $R > 0$ e note que $B(y;R) \cap B(a;r)$ possui pelo menos um ponto de $B(a;r)$ diferente de y.

- Se $R \leq r - d$;
  Neste caso, $B(y;R) \subset B(a;r) \Rightarrow B(y;R) \cap B(a;r) = B(y;R)$.

- Se $R > r - d$;
  Considere $0 < \varepsilon < r - d$, então $B(y;\varepsilon) \subset B(y;R) \cap B(a;r)$.

Exemplo 1.80. O conjunto $\mathbb{Z}^n$ dos pontos de $\mathbb{R}^n$ com coordenadas inteiras é um conjunto discreto.

Teorema 1.81. *Sejam a um ponto e X um subconjunto de $\mathbb{R}^n$, as seguintes afirmações são equivalentes:*

1) a é um ponto de acumulação de X.

2) a é limite de uma sequência de pontos de $x_k \in X \backslash \{a\}$.

3) Toda bola de centro a contém uma infinidade de pontos de X.

Demonstração: $(1 \Rightarrow 2)$ Seja a um ponto de acumulação de X, então para todo $n \in \mathbb{N}$ podemos encontrar um ponto $x_k \in X$, tal que $x_k \neq a$, em $B(a, \frac{1}{k})$. De fato, dado $\varepsilon > 0$, tome $k_0 \in \mathbb{N}$ tal que $k_0 > \frac{1}{\varepsilon} \Rightarrow \frac{1}{k_0} < \varepsilon$. Daí se $k > k_0$, temos $\|x_k - a\| < \frac{1}{k} < \frac{1}{k_0} < \varepsilon$, então $\lim x_k = a$.
$(2 \Rightarrow 3)$ Seja a tal que $\lim x_k = a$, onde $x_k \in X \backslash \{a\}$, então para todo $k_0 \in \mathbb{N}$ o conjunto $\{x_k; k > k_0\}$ é infinito, pois se fosse finito, teria que existir um x_{k_1} que se repetiria infinitas vezes e isto forneceria uma sequência constante onde $\lim x_{k_i} \neq a$. Logo, toda bola de centro a, contém uma infinidade de pontos de X.
$(3 \Rightarrow 1)$ Supondo que toda bola de centro a contém uma infinidade de pontos de X, segue da definição que a é um ponto de acumulação de X. ■

Teorema 1.82. *Todo subconjunto infinito e limitado $X \subset \mathbb{R}^n$, possui pelo menos um ponto de acumulação.*

Demonstração: Seja $X \subset \mathbb{R}^n$ infinito e limitado, então X admite subconjunto infinito e enumerável $\{x_1, x_2, ..., x_k, ...\}$. Fixando esta enumeração, temos uma sequência (x_k) de termos dois a dois distintos, pertencentes a X. Logo, (x_k) é limitada e por Bolzano-Weierstrass, (x_k) possui uma subsequência (x_{k_m}) convergente. Defina $(y_m) = (x_{k_m})$, com

(y_m) convergente, onde $\lim y_m = a$. Como cada coordenada y_m são distintas, no máximo uma coordenada pode ser igual a a. Se existe, descartamos e temos a como limite de uma sequência de pontos $y_m \in X \backslash \{a\}$, logo a é um ponto de acumulação de X. ■

Teorema 1.83. *i) Se F_1 e F_2 são subconjuntos fechados de $\mathbb{R}^n$, então $F_1 \cup F_2$ é fechado.*

ii) Se $(F_\lambda)_{\lambda \in L}$ é uma família arbitrária de conjuntos fechados, então $F = \cap_{\lambda \in L} F_\lambda$ é fechado.

Demonstração:

$i)$ Dados F_1 e F_2 fechados, então ${F_1}^c$ e ${F_2}^c$ são abertos. Note que ${F_1}^c \cap {F_2}^c = (F_1 \cup F_2)^c$, como a interseção finita de abertos é aberta, então ${F_1}^c \cap {F_2}^c$ é aberto, logo $(F_1 \cup F_2)^c$ é aberto. Portanto, $F_1 \cup F_2$ é fechado.

$ii)$ Seja $(F_\lambda)_{\lambda \in L}$ uma família de fechados. Como para todo $\lambda \in L$, F_λ é fechado, então ${F_\lambda}^c$ é aberto. Como a união arbitrária de abertos é aberta, então $\cup_{\lambda \in L} {F_\lambda}^c$ é aberto. Note que $(\cup_{\lambda \in L} {F_\lambda}^c)^c = \cap_{\lambda \in L} F_\lambda$, logo $\cap_{\lambda \in L} F_\lambda$ é um conjunto fechado.

■

Observação 1.84. Segue de $i)$ que a união finita de conjuntos fechados é um conjunto fechado. Contudo, para a união infinita nem sempre é verdade. Pois, um conjunto qualquer, aberto ou fechado, é a união dos seus pontos cujos pontos são conjuntos fechados.

Observação 1.85. Segue do Teorema 1.81, que $\overline{X}$ é formado acrescentando-lhe seus pontos de acumulação que porventura não pertencem a X.

Definição 1.86. Seja $X \subset \mathbb{R}^n$, dizemos que um subconjunto $F \subset X$ é fechado em X, quando F contém todos os seus pontos aderentes que pertencem a X. F é conhecido como fechado relativo.

Proposição 1.87. *Sejam $X \subset \mathbb{R}^n$ e $F \subset X$, F é fechado em X se, e somente se $F = \overline{F} \cap X$.*

Demonstração: $(\Rightarrow)$ Perceba que $F \subset \overline{F}$ e como $F \subset X \Rightarrow F \subset \overline{F} \cap X$. Seja $a \in \overline{F} \cap X \Rightarrow a \in \overline{F}$ e $a \in X$. Logo existe $(x_k) \subset F$ tal que $x_k \to a$ então a é aderente a F. Pela definição de fechado relativo, $a \in F$, logo $\overline{F} \cap X = F$.
$(\Leftarrow)$ Seja $a \in X$ e $a \in \overline{F} \Rightarrow a \in X \cap \overline{F} = F$. Portanto, F é fechado relativo em X. ■

Proposição 1.88. *Seja $X \subset \mathbb{R}^n$ e $F \subset X$, então F é fechado em X se, e somente se, $X \backslash F$ é aberto em X.*

Demonstração: ($\Rightarrow$) Se F é fechado em X, então $F = \mathcal{F} \cap X$, onde $\mathcal{F} \subseteq \mathbb{R}^n$ é fechado. Daí, $X \backslash F = X \cap F^c = X \cap (\mathcal{F}^c \cup X^c) = (X \cap \mathcal{F}^c) \cup (X \cap X^c) = X \cap \mathcal{F}^c$. Ou seja, $X \backslash F = X \cap \mathcal{F}^c$, onde $\mathcal{F}^c$ é aberto. Pela observação 1.40, temos que $X \backslash F$ é aberto em X.
($\Leftarrow$) Suponha que $X \backslash F = X \cap U$, onde $U \subseteq \mathbb{R}^n$ é aberto em $\mathbb{R}^n$. Assim, $F = X \backslash (X \backslash F) = X \backslash (X \cap U) = X \cap (X \cap U)^c = X \cap (X^c \cup U^c) = (X \cap X^c) \cup (X \cap U^c) = X \cap U^c$. Ou seja, $F = X \cap U^c$, com U^c fechado. Pela Proposição 1.87, vemos que F é fechado em X. ∎

Proposição 1.89. *Seja $X \subset \mathbb{R}^n$ e $A \subset X$, então A é aberto em X se, e somente se, $X \backslash A$ é fechado em X.*

Demonstração: ($\Rightarrow$) Se A é aberto em X, então $A = \mathcal{A} \cap X$, onde $\mathcal{A} \subseteq \mathbb{R}^n$ é aberto. Daí, $X \backslash A = X \cap A^c = X \cap (\mathcal{A}^c \cup X^c) = (X \cap \mathcal{A}^c) \cup (X \cap X^c) = X \cap \mathcal{A}^c$. Ou seja, $X \backslash A = X \cap \mathcal{A}^c$, onde $\mathcal{A}^c$ é um fechado. Pela observação 1.40, temos que $X \backslash A$ é fechado em X.
($\Leftarrow$) Suponha que $X \backslash A = X \cap Y$, onde $Y \subseteq \mathbb{R}^n$ é fechado em $\mathbb{R}^n$. Então $A = X \backslash (X \backslash A) = X \backslash (X \cap Y) = X \cap (X \cap Y)^c = X \cap (X^c \cup Y^c) = (X \cap X^c) \cup (X \cap Y^c) = X \cap Y^c$. Ou seja, $A = X \cap Y^c$, com Y^c aberto. Pela Proposição 1.87, vemos que A é aberto em X. ∎

1.6 Conjuntos Compactos

Definição 1.90. Um conjunto $X \subset \mathbb{R}^n$ chama-se compacto quando é fechado e limitado.

Exemplo 1.91. Toda bola fechada $B[a; r]$ é compacta.

Exemplo 1.92. Nenhuma bola aberta $B(a; r)$ é compacta.

Provamos na seção de conjuntos abertos que toda bola aberta $B(a; r)$ é um conjunto aberto. Além disso, note que $\overline{B(a; r)} = B[a; r] \neq B(a; r)$. Logo, como $B(a; r)$ não é fechada, segue que não é compacta.

Exemplo 1.93. O conjunto $\mathbb{Z}^n$ não é compacto.

Suponha por contradição que $\mathbb{Z}^n$ é limitado. Logo, existe $r > 0$, tal que $\mathbb{Z}^n \subset B(0; r)$. Considere $k \in \mathbb{N}$; $k > r$ e $z = (0, ..., 0, k) \in \mathbb{Z}^n$, mas $\|z - 0\| = \|z\| = k > r$. O que é um absurdo. Portanto, como $\mathbb{Z}^n$ não é limitado, segue que não é compacto.

Exemplo 1.94. Toda esfera $S[a; r]$ é compacta.

Teorema 1.95. *As seguintes afirmações sobre o conjunto $K \subset \mathbb{R}^n$ são equivalentes:*

i) *K é compacto;*

ii) *Toda sequência de pontos $x_k \in K$ possui uma subsequência que converge para um ponto de K.*

Demonstração: $i) \Rightarrow ii)$ Como K é compacto, em particular, K é limitado. Logo, toda sequência de pontos $x_k \in K$ é limitada. Pelo Teorema de Bolzano-Weierstrass, existe uma subsequência $(x_k)_{k\in\mathbb{N}'}$ de (x_k) que converge para um ponto a, ou seja, $\lim\limits_{k\in\mathbb{N}'} x_k = a$. Como K é um conjunto fechado, então $a \in K$.
$ii) \Rightarrow i)$ Como vale $ii)$, então K é limitado, pois se não fosse, para todo $k \in \mathbb{N}$ existiria um ponto $x_k \in K$ tal que $\|x_k\| > k$. Assim a sequência obtida (x_k) não possuiria subsequência limitada, o que impediria a existência de suas subsequências serem convergentes. Perceba também que K é fechado, pois se $\lim x_k = a$, onde $x_k \in K$, $\forall\ k \in \mathbb{N}$, então por $ii)$, uma subsequência de (x_k) convergiria para um ponto de K, mas toda subsequência de (x_k) converge para a. Logo $a \in K$. Portanto, K é compacto. ■

Definição 1.96. Sejam $X, Y \subset \mathbb{R}^n$, a distância entre eles é dada por $d(X,Y) = inf\{\|x - y\|; x \in X, y \in Y\}$.

Mas em caso particular, vale o teorema a seguir.

Corolário 1.97. *Sejam $K \subset \mathbb{R}^n$ compacto e $F \subset \mathbb{R}^n$ fechado, então existem $x_0 \in K$ e $y_0 \in F$ tais que $\|x_0 - y_0\| \leq \|x - y\|$ para quaisquer $x \in K$ e $y \in F$.*

Demonstração: Pela definição de ínfimo, existem sequências $(x_k) \subset K$ e $(y_k) \subset F$ tais que $\|x_k - y_k\| \to d(K,F)$. Como K é compacto, pelo Teorema 1.95, existe uma subsequência $(x_k)_{k\in\mathbb{N}'}$ e $x_0 \in K$ tal que $\lim\limits_{k\in\mathbb{N}'} x_k = x_0$. Note que $\|y_k\| \leq \|y_k - x_k\| + \|x_k\|$, o que torna $(y_k)_{k\in\mathbb{N}'}$ limitada. Por Bolzano-Weierstrass, existe uma subsequência $(y_k)_{k\in\mathbb{N}''}$ e $y_0 \in \mathbb{R}^n$ tal que $\lim\limits_{k\in\mathbb{N}''} y_k = y_0$. Portanto, $\|x_0 - y_0\| = \lim\limits_{k\in\mathbb{N}''} \|x_k - y_k\| = \lim\limits_{k\in\mathbb{N}} \|x_k - y_k\| = d(K,F)$. Além disso, $d(K,F) \leq \|x - y\|$, $\forall\ x \in K, y \in F$. E isso implica em $\|x_0 - y_0\| \leq \|x - y\|$, $\forall\ x \in K$ e $y \in F$. ■

Observação 1.98. Se $F_1 \supset F_2 \supset ... \supset F_k \supset ...$ é uma sequência decrescente de conjuntos fechados não-vazios em $\mathbb{R}^n$, pode ocorrer que $\cap_{k=1}^{\infty} F_k = \varnothing$.

Se tomarmos $F_k = [K, +\infty)$ em $\mathbb{R}$, isto ocorre. Contudo, o mesmo não ocorre quando um dos F_k é limitado. Vejamos o teorema a seguir.

Teorema 1.99. *(Cantor) Seja $K_1 \supset K_2 \supset ... \supset K_k \supset ...$ uma sequência decrescente de compactos não-vazios em $\mathbb{R}^n$. Existe pelo menos um ponto $a \in \mathbb{R}^n$ que pertence a todos os K_k. Ou seja, $\cap_{k=1}^{\infty} K_k \neq \varnothing$.*

Demonstração: Escolhamos um ponto $x_k \in K_k\ \forall\ k \in \mathbb{N}$. Como cada K_k é compacto, a sequência (x_k) é limitada, pois $(x_k) \subset K_1$. Pelo Teorema de Bolzano-Weierstrass, (x_k) possui uma subsequência $(x_r)_{r\in\mathbb{N}'}$ convergindo para $a \in \mathbb{R}^n$. Suponha, por contradição, que exista $n_0 \in \mathbb{N}$ tal que $a \notin K_{n_0}$. Escrevamos $\mathbb{N}' = \{n_1, ..., n_r, ...\}$ e considere $r_0 \in \mathbb{N}$ tal que $n_{r_0} > n_0$, assim $X_{n_r} \in K_{n_0}$ para todo $r > r_0$. Como $x_{n_r} \to a$, obtemos que $a \in K_{n_0}$, o que é um absurdo. ■

Definição 1.100. Uma cobertura do conjunto $X \subset \mathbb{R}^n$ é uma família $(C_\lambda)_{\lambda \in L}$ de subconjuntos $C_\lambda \subset \mathbb{R}^n$ tais que $X \subset \cup_{\lambda \in L} C_\lambda$. Isto significa que para cada $x \in X$, existe um $\lambda \in L$ tal que $x \in C_\lambda$.

Definição 1.101. Dizemos que $\cup_{\lambda \in L} C_\lambda$ é uma cobertura aberta de X, quando os C_λ forem todos abertos.

Definição 1.102. Dizemos que $\cup_{\lambda \in L} C_\lambda$ é uma cobertura finita de X, quando L é um conjunto finito.

Definição 1.103. Uma subcobertura é uma subfamília $(C_\lambda)_{\lambda \in L'}$, $L' \subset L$, tal que ainda se tem $X \subset \cup_{\lambda \in L'} C_\lambda$.

Uma propriedade fundamental dos conjuntos compactos é o fato de que toda cobertura aberta de um conjunto compacto possui uma subcobertura finita.

Definição 1.104. Seja $X \subset \mathbb{R}^n$ um conjunto limitado, o diâmetro de X é o número

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

Observação 1.105. Segue que, se $diam.X = d$ e $x \in X$, então $X \subset B[x; d]$.

Lema 1.106. *Seja $K \subset \mathbb{R}^n$ compacto. Para todo $\varepsilon > 0$ existe uma decomposição $K = K_1 \cup K_2 \cup ... \cup K_k$, onde cada K_i é compacto e tem diâmetro $\leq \varepsilon$.*

Demonstração: Seja $\alpha > 0$, note que $\mathbb{R} = \cup_{m \in \mathbb{Z}} [m\alpha, (m+1)\alpha]$. Analogamente, podemos escrever $\mathbb{R}^n = \cup_{m \in \mathbb{Z}^n} C_m$, onde $C_m = \prod_{i=1}^n [m_i\alpha, (m_i + 1)\alpha]$ e $m = (m_1, ..., m_n)$. Assim, $diam.C = \alpha\sqrt{n}$ para todo cubo C. Dado um subconjunto $X \subset \mathbb{R}^n$, temos que $X = X \cap \mathbb{R}^n = X \cap (\cup_{m \in \mathbb{Z}^n} C_m) = \cup_{m \in \mathbb{Z}^n} (X \cap C_m)$. Em particular, para o compacto $K \subset \mathbb{R}^n$; $K = \cup_{m \in \mathbb{Z}^n} (K \cap C_m) = \cup_{m \in \mathbb{Z}^n} K_m$, onde $K_m = K \cap C_m$, $\forall\ m \in \mathbb{Z}^n$. Claramente K é compacto para todo $m \in \mathbb{Z}^n$. Por K ser limitado, existe apenas uma quantidade finita de K_i's não-vazios. Pois sendo $s > 0$ tal que $K \subset B(0; s)$, considere $t \in \mathbb{N}$ tal que $t\alpha > s$ e seja $m \in \mathbb{Z}^n$ tal que $|m_i| \leq t$, para algum $i \in \{1, ..., n\}$. Com isto afirmamos que $C_m \cap K \neq \varnothing$, pois dado $x \in C_m$, $\|x - 0\| = \|x\| \leq t\alpha > s \Rightarrow x \notin B(0; s) \Rightarrow x \notin K$. Logo, existem no máximo t^n cubos da forma C_m em que $C_m \cap K = \varnothing$. Tome $K_1, ..., K_k$ os conjuntos da forma $K \cap C_m$ em que $K \cap C_m \neq \varnothing$. Logo, $K = \cup_{m \in \mathbb{Z}^n} K_m = \cup_{i=1}^k K_i$. Para que $diam\ K_i < \varepsilon$, escolha $\alpha > 0$ de modo que $\alpha < \frac{\varepsilon}{\sqrt{n}}$, pois;

$$\begin{aligned} diam.K_i &= diam.(K \cap C_m) \\ &\leq diam.C_m = \alpha\sqrt{n} \\ &< \frac{\varepsilon}{\sqrt{n}}\sqrt{n} = \varepsilon. \end{aligned}$$

■

Teorema 1.107. *(Borel-Lebesgue) Toda cobertura aberta* $K \subset \cup A_\lambda$ *de um compacto* $K \subset \mathbb{R}^n$ *admite uma subcobertura finita* $K \subset A_{\lambda_1} \cup \ldots \cup A_{\lambda_k}$.

Demonstração: Seja $K \subset \mathbb{R}^n$ compacto, suponha por absurdo, que $K \subset \cup A_\lambda$ seja uma cobertura aberta que não admite subcobertura finita. Pelo Lema 1.106, podemos expressar K como união finita de compactos de diâmetro < 1. Pelo menos um dos compactos, digamos, K_1 é tal que $K_1 \subset \cup A_\lambda$ não admite subcobertura finita. Agora, escrevendo K_1 como união finita de compactos cujo diâmetro é $< \frac{1}{2}$, pelo menos um deles, digamos, K_2, não pode ser coberto por um número finito de A_λ*'s*. Continuando esse processo, obtemos uma sequência decrescente de conjuntos compactos $K_1 \supset K_2 \supset \ldots \supset K_k \supset \ldots$ com diam $K_k < \frac{1}{k}$ de modo que nenhum dos K_i*'s*, para todo $i = 1, \ldots, k$ está contido em uma união finita de A_λ's. Em particular, todos os K_k são não-vazios, então pelo Teorema de Cantor, existe $a \in \cap_{k=1}^{\infty} K_k$, para algum λ temos $a \in A_\lambda$. Como A_λ é aberto, temos $B\left(a; \frac{1}{k}\right) \subset A_\lambda$; para algum k. Sendo $a \in K_k$ e $diam.K_k < \frac{1}{k}$, temos que $K_k \subset B\left(a; \frac{1}{k}\right)$, onde $K_k \subset A_\lambda$, o que é uma contradição. Pois nenhum K_k poderia está contido numa quantidade finita de A_λ*'s*. ■

Capítulo 2

Aplicações Contínuas

As funções contínuas desempenham um papel de destaque no cotidiano por apresentarem aspectos especiais. Na matemática, a garantia da continuidade de uma determinada função é algo bastante relevante. Neste capítulo, veremos como a noção de função contínua tomando valores reais está intrinsecamente relacionada com aplicações contínuas no espaço Euclidiano. Recorremos aos números positivos ε e δ, para definir as aplicações contínuas e sua relação com os conjuntos abertos, fechados, compactos e com as sequências.

2.1 Limites de Funções

Definição 2.1. Seja $X \subset \mathbb{R}^m$, uma aplicação $f : X \to \mathbb{R}^n$ associa a cada ponto $x \in X$ sua imagem $f(x) = (f_1(x), ...f_n(x))$. As funções reais $f_1, ..., f_n : X \to \mathbb{R}$ são chamadas de funções-coordenadas de f e escrevemos $f = (f_1, ..., f_n)$.

Se $Y \subset \mathbb{R}^n$ é tal que $f(X) \subset Y$, podemos escrever $f : X \to Y$ em vez de $f : X \to \mathbb{R}^n$.

Definição 2.2. Sejam $f : X \to \mathbb{R}^n$ definida no conjunto $X \subset \mathbb{R}^m$ e $a \in \mathbb{R}^m$ um ponto de acumulação de X. Dizemos que $b \in \mathbb{R}^n$ é o limite de $f(x)$ quando x tende para a e escrevemos $\lim\limits_{x \to a} f(x) = b$, se para todo $\varepsilon > 0$ dado, existir $\delta > 0$ de modo que $\|f(x) - b\| < \varepsilon$, sempre que $0 < \|x - a\| < \delta$, onde $x \in X$.

Observação 2.3. i. O ponto a pode pertencer ou não a X.

ii. Quando o ponto de acumulação a pertence a X, a aplicação $f : X \to \mathbb{R}^n$ é contínua no ponto a se, e somente se, $\lim\limits_{x \to a} f(x) = f(a)$.

Teorema 2.4. *(Permanência do Sinal) Sejam a um ponto de acumulação de $X \subset \mathbb{R}^n$ e $f : X \to \mathbb{R}$ uma função real. Se $b = \lim\limits_{x \to a} f(x)$ é um número positivo, então existe $\delta > 0$ tal que $x \in X$ e $0 < \|x - a\| < \delta$ implicam $f(x) > 0$.*

Demonstração: Como b é um número positivo, tome $\varepsilon = b$. Assim, pela definição de limite, existe $\delta > 0$ tal que $x \in X$ e $0 < |x - a| < \delta$ implicam que $|f(x) - b| < \varepsilon = b$, ou seja, $-b < f(x) - b < b$, $0 < f(x) < 2b$. Portanto, $f(x) > 0$. ■

Quando X é um intervalo da reta, tem sentido a noção de limite lateral de uma aplicação $f : I \to \mathbb{R}^n$, ou seja, de um caminho, num ponto $a \in I$.

Definição 2.5. Se a não é o extremo superior de I, dizemos que $b \in \mathbb{R}^n$ é o limite à direita de $f(t)$ quando t tende para a, e escrevemos $\lim\limits_{t \to a^+} f(t) = b$, quando para todo $\varepsilon > 0$ dado, existe $\delta > 0$ tal que $a < t < a + \delta$; $t \in I$, implica $\|f(x) - b\| < \varepsilon$.

Definição 2.6. Se a não é o extremo inferior de I, dizemos que $b \in \mathbb{R}^n$ é o limite à esquerda de $f(t)$ quando t tende para a, e escrevemos $\lim\limits_{t \to a^-} f(t) = b$ quando para todo $\varepsilon > 0$ dado, existe $\delta > 0$ tal que $a - \delta < t < a$ implica $t \in I$ e $\|f(x) - b\| < \varepsilon$.

De modo semelhante à continuidade de uma aplicação, a existência e o valor do limite se exprimem em termos das funções-coordenadas.

Teorema 2.7. *Sejam a um ponto de acumulação do conjunto X, $f : X \to \mathbb{R}^n$ uma aplicação e $f_1, ..., f_n : X \to \mathbb{R}$ as funções-coordenadas de f. Temos $\lim\limits_{x \to a} f(x) = b$ se, e somente se $\lim\limits_{x \to a} f_i(x) = b_i$, para todo $i = 1, ..., n$.*

Demonstração: $(\Rightarrow)$ Dado $\varepsilon > 0$, queremos exibir $\delta > 0$ tal que $|f_i(x) - b_i| < \varepsilon$, para todo $0 < \|x - a\| < \delta_i$. No entanto, como $\lim\limits_{x \to a} f(x) = b$, existe $\delta > 0$ tal que $\|f(x) - b\| < \varepsilon$, para todo $0 < \|x - a\| < \delta$. Daí, tomando $\delta_i = \delta$, se $\|x - a\| < \delta = \delta_i$, temos;

$$\begin{aligned} |f_i(x) - b_i| &= [(f_i(x) - b_i)^2]^{\frac{1}{2}} \\ &\leq [(f_1(x) - b_1)^2 + ... + (f_n(x) - b_n)^2]^{\frac{1}{2}} \\ &= \|f(x) - b\|_S < \varepsilon. \end{aligned}$$

Assim, $\lim\limits_{x \to a} f_i(x) = b_i$, para todo $i = 1, ..., n$.
$(\Leftarrow)$ Como $\lim\limits_{x \to a} f_i(x) = b_i$ para todo $i = 1, ..., n$, dado $\varepsilon > 0$ existe $\delta_i > 0$ de modo que $|f_i(x) - b_i| < \frac{\varepsilon}{n}$, para todo $0 < \|x - a\| < \delta_i$. Assim;

$$\begin{aligned} \|f(x) - b\| &= [(f_1(x) - b_1)^2 + ... + (f_n(x) - b_n)^2]^{\frac{1}{2}} \\ &\leq \|f(x) - b\|_S \\ &= |f_1(x) - b_1| + ... + |f_n(x) - b_n|. \end{aligned}$$

Tomando $\delta = min\{\delta_1, ..., \delta_n\} > 0$, temos que se $\|x - a\| < \delta$, vale que $\|f(x) - b\| < \frac{\varepsilon}{n} + ... + \frac{\varepsilon}{n} = n \cdot (\frac{\varepsilon}{n}) = \varepsilon$. Assim, $\lim\limits_{x \to a} f(x) = b$. ■

O Teorema a seguir associa o limite de aplicações com o limite de sequências.

Teorema 2.8. *Seja a um ponto de acumulação do conjunto $X \subset \mathbb{R}^m$. A fim de que se tenha $\lim\limits_{x\to a} f(x) = b$ é necessário e suficiente que para toda sequência de pontos $x_k \in X\backslash\{a\}$ com $\lim\limits_{k\to\infty} x_k = a$, temos $\lim\limits_{k\to\infty} f(x_k) = b$.*

Demonstração: $(\Rightarrow)$ Dado $\varepsilon > 0$, arbitrário, existe $\delta > 0$ tal que $x \in X$, $0 < \|x-a\| < \delta$ implica em $\|f(x) - b\| < \varepsilon$. Seja $(x_k) \subset X\backslash\{a\}$ tal que $x_k \to a$. Note que existe $k_0 \in \mathbb{N}$ tal que $k > k_0$ implica em $0 < \|x_k - a\| < \delta$. Portanto $\|f(x_k) - b\| < \varepsilon$.
$(\Leftarrow)$ Suponha por contradição que $\lim\limits_{x\to a} f(x) \neq b$. Assim, existe $\varepsilon > 0$ tal que para todo $k \in \mathbb{N}$, podemos encontrar $x_k \in X$ tal que $0 < \|x_k - a\| < \frac{1}{k}$ mas $\|f(x_k) - b\| \geq \varepsilon$. Então teríamos $x_k \in X\backslash\{a\}$, $\lim\limits_{k\to\infty} x_k = a$ mas com $\lim\limits_{k\to\infty} f(x_k) \neq b$. O que é um absurdo. ■

Teorema 2.9. *Sejam a um ponto de acumulação de $X \subset \mathbb{R}^m$, $b \in Y$, $f : X \to Y$ uma aplicação tal que $\lim\limits_{x\to a} f(x) = b$ e $g : Y \to \mathbb{R}^p$ contínua no ponto b. Então $\lim\limits_{x\to a} g(f(x)) = g(b)$.*

Demonstração: Dado $\varepsilon > 0$, como g é contínua no ponto $f(a)$, existe $\lambda > 0$ tal que se $y \in Y$, $\|y-f(a)\| < \lambda \Rightarrow \|g(y)-g(f(a))\| < \varepsilon$. Como f é contínua no ponto a, para $\lambda > 0$, existe $\delta > 0$ tal que se $x \in X$, $\|x-a\| < \delta \Rightarrow \|f(x)-f(a)\| < \lambda \Rightarrow \|g(f(x))-g(f(a))\| < \varepsilon$, assim, $g \circ f$ é contínua no ponto a. ■

Teorema 2.10. *Sejam $f, g : X \to \mathbb{R}^n$ e $\alpha : X \to \mathbb{R}$ definidas no conjunto $X \subset \mathbb{R}^m$ e a um ponto de acumulação de X. Se existem $\lim\limits_{x\to a} f(x) = b$, $\lim\limits_{x\to a} g(x) = c$ e $\lim\limits_{x\to a} \alpha(x) = \lambda$, então existem os limites e valem as seguintes igualdades;*

i) $\lim\limits_{x\to a}[f(x) + g(x)] = b + c$.

ii) $\lim\limits_{x\to a} \alpha(x) \cdot f(x) = \lambda \cdot b$.

iii) $\lim\limits_{x\to a}\langle f(x), g(x)\rangle = \langle b, c\rangle$.

iv) $\lim\limits_{x\to a} \|f(x)\| = \|b\|$.

Demonstração:

i) A aplicação $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, definida por $s(x, y) = x + y$, é contínua. Observando que $f(x) + g(x) = s(f(x), g(x))$, resulta do Teorema 2.9 que

$$\lim_{x\to a}[f(x) + g(x)] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = b + c.$$

ii) A aplicação $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, definida por $p(x, y) = xy$, é contínua. Observando que $\alpha(x)g(x) = p(\alpha(x), f(x))$, resulta do Teorema 2.9 que

$$\lim_{x\to a}[\alpha(x)g(x)] = \lim_{x\to a} \alpha(x) \cdot \lim_{x\to a} f(x) = \lambda \cdot b.$$

iii) Note que;

$$
\begin{aligned}
|\langle f(x), g(x)\rangle - \langle b, c\rangle| &= |\langle f(x), g(x)\rangle - \langle b, g(x)\rangle + \langle b, g(x)\rangle - \langle b, c\rangle| \\
&\leq |\langle f(x) - b, g(x)\rangle| + |\langle b, g(x) - c\rangle|
\end{aligned}
$$

$$
\leq \|f(x) - b\| \cdot \|g(x)\| + \|b\| \cdot \|g(x) - c\|. \tag{2.1}
$$

Agora, note que dado $\varepsilon > 0$, existem $\delta_1, \delta_2 > 0$ tais que $\|f(x) - b\| < \varepsilon_1$, se $0 < \|x - a\| < \delta_1$ e $\|g(x) - c\| < \varepsilon_2$, se $0 < \|x - a\| < \delta_2$. Se $0 < \|x - a\| < \delta = min\{\delta_1, \delta_2\}$, temos por (2.1),

$$
\begin{aligned}
|\langle f(x), g(x)\rangle - \langle b, c\rangle| &< \varepsilon_1 \|g(x)\| + \|b\|\varepsilon_2 \\
&< \varepsilon_1(\|g(x)\| + 1) + (\|b\| + 1)\varepsilon_2 \\
&= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}
$$

Assim, $\lim\limits_{x\to a}\langle f(x), g(x)\rangle = \langle b, c\rangle$.

iv) Como $\langle f, g\rangle(x) = \langle f(x), g(x)\rangle$ e $\lim\limits_{x\to a}\langle f(x), g(x)\rangle = \langle b, c\rangle$, se $g = f$ e $c = b$, temos:

$$
\begin{aligned}
\lim_{x\to a}\langle f(x), f(x)\rangle &= \langle b, b\rangle \\
\lim_{x\to a}\|f(x)\|^2 &= \|b\|^2.
\end{aligned}
$$

Agora, defina;

$$
\begin{aligned}
\varphi : \mathbb{R}^+ &\longrightarrow \mathbb{R} \\
t &\longmapsto \sqrt{t}.
\end{aligned}
$$

Como φ é contínua, então $\lim\limits_{t\to t_0}\varphi(t) = \varphi(t_0)$. Assim, considere;

$$
\begin{aligned}
\sigma : X &\longrightarrow \mathbb{R} \\
x &\longmapsto \sigma(x) = \varphi_0(\langle f, f\rangle)(x) = \varphi(\langle f(x), f(x)\rangle).
\end{aligned}
$$

Pelo Teorema 2.9, segue que

$$
\begin{aligned}
\lim_{x\to a}\sigma(x) &= \sigma(a) \\
\lim_{x\to a}\sqrt{\|f(x)\|^2} &= \sqrt{\|b\|^2} \\
\lim_{x\to a}\|f(x)\| &= \|b\|.
\end{aligned}
$$

$\blacksquare$

Observação 2.11. Considerando $\alpha : X \to \mathbb{R}$, perceba que se $\lim\limits_{x\to a} \alpha(x) = 0$ e $f : X \to \mathbb{R}^n$ é limitada na vizinhança de a (isto é, existem $\delta > 0$ e $M > 0$ tais que $x \in X$ e $0 < \|x-a\| < \delta$ implicam $\|f(x)\| \leq M$), então $\lim\limits_{x\to a} \alpha(x).f(x) = 0$, mesmo que não exista o $\lim\limits_{x\to a} f(x)$.

Exemplo 2.12. Seja $f(x,y) = \frac{xy}{x^2+y^2} = \frac{x}{\sqrt{x^2+y^2}} \cdot \frac{y}{\sqrt{x^2+y^2}}$ e $g : \mathbb{R}^2\backslash\{0\} \to \mathbb{R}$ definida por $g(x,y) = \frac{x^2y}{(x^2+y^2)}$. Note que $\lim\limits_{(x,y)\to(0,0)} g(x,y) = 0$, pois podemos escrever $g(x,y) = \alpha(x,y) \cdot f(x,y)$ de modo que $\lim\limits_{(x,y)\to(0,0)} \alpha(x,y) = 0$. Além disso,

$$|f(x,y)| = \left|\frac{x}{\sqrt{x^2+y^2}}\right| \cdot \left|\frac{y}{\sqrt{x^2+y^2}}\right| = \frac{|x|}{\sqrt{x^2+y^2}} \cdot \frac{|y|}{\sqrt{x^2+y^2}} = \frac{\sqrt{x^2}}{\sqrt{x^2+y^2}} \cdot \frac{\sqrt{y^2}}{\sqrt{x^2+y^2}} \leq 1.$$

Teorema 2.13. *(Permanência da Desigualdade) Sejam $f, g : X \to \mathbb{R}$ definidas no conjunto $X \subset \mathbb{R}^m$ e a um ponto de acumulação de X. Se $f(x) \leq g(x)$ para todo $x \in X$ e existem $\lim\limits_{x\to a} f(x)$ e $\lim\limits_{x\to a} g(x)$, então tem-se $\lim\limits_{x\to a} f(x) \leq \lim\limits_{x\to a} g(x)$.*

Demonstração: Suponha, por contradição, que $\lim\limits_{x\to a} f(x) > \lim\limits_{x\to a} g(x)$ e defina;

$$\begin{aligned} h : X &\longrightarrow \mathbb{R} \\ x &\longmapsto h(x) = f(x) - g(x). \end{aligned}$$

Assim, $\lim\limits_{x\to a} h(x) = \lim\limits_{x\to a} f(x) - \lim\limits_{x\to a} g(x) > 0$. Portanto, pelo Teorema 2.4, existe $\delta > 0$ tal que se $x \in X \cap B(a;\delta)$, temos $h(x) > 0$, o que implica que $f(x) - g(x) > 0$, ou seja $f(x) > g(x)$. O que é um absurdo. ■

2.2 Funções Contínuas

Definição 2.14. Seja $a \in X$, dizemos que f é contínua no ponto a quando, para cada $\varepsilon > 0$ arbitrariamente dado, pode-se obter $\delta > 0$ tal que $x \in X$, $\|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \varepsilon$. Em outros termos, para cada bola $B(f(a);\varepsilon)$ dada, existe uma bola $B(a;\delta)$ tal que $f(B(a;\delta) \cap X) \subset B(f(a);\varepsilon)$. Dizemos que $f : X \to \mathbb{R}^n$ é uma aplicação contínua no conjunto X, quando f é contínua em todos os pontos $a \in X$.

Teorema 2.15. *Sejam $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, $f : X \to \mathbb{R}^n$ com $f(X) \subset Y$ e $g : Y \to \mathbb{R}^p$. Se f é contínua no ponto $a \in X$ e g é contínua no ponto $f(a) \in Y$, então $g \circ f : X \to \mathbb{R}^p$ é contínua no ponto a. Ou seja, a composta de duas aplicações contínuas é contínua.*

Demonstração: A demonstração segue a mesma ideia do Teorema 2.9. ■

Teorema 2.16. *i) A aplicação $f : X \to \mathbb{R}^n$ é contínua no ponto $a \in X$ se, e somente se, para toda sequência de pontos $x_k \in X$ com $\lim x_k = a$, tem-se $\lim f(x_k) = f(a)$.*

ii) *A aplicação $f : X \to \mathbb{R}^n$ é contínua no ponto $a \in X$ se, e somente se, suas funções-coordenadas $f_1, ..., f_n : X \to \mathbb{R}$ são contínuas nesse ponto.*

Demonstração:

i) ($\Rightarrow$) Seja $f : X \to \mathbb{R}^n$ contínua no ponto a. Dada a sequência de pontos $x_k \in X$ com $\lim x_k = a$, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $f(B(a;\delta)) \subset B(f(a);\varepsilon)$. Em relação a δ, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow x_k \in B(a;\delta)$, logo $k > k_0 \Rightarrow f(x_k) \in B(f(a);\varepsilon)$. Logo, $\lim f(x_k) = f(a)$.
($\Leftarrow$) Suponha, por contradição, que f não é contínua em a. Assim, existe $\varepsilon > 0$ tal que para todo $k \in \mathbb{N}$ existe $x_k \in X$ de modo que $\|x_k - a\| < \frac{1}{k}$ e $\|f(x_k) - f(a)\| \geq \varepsilon$. Assim, $x_k \to a$, mas $f(x_k) \nrightarrow f(a)$, o que é um absurdo.

ii) Consequência direta do Teorema 2.7.

■

Corolário 2.17. *Seja $X \subset \mathbb{R}^m$. Se as aplicações $f, g : X \to \mathbb{R}^n$ e $\alpha : X \to \mathbb{R}$ são contínuas no ponto $a \in X$, então também são contínuas nesse ponto, as aplicações $f + g : X \to \mathbb{R}^n$; $\langle f, g\rangle : X \to \mathbb{R}$, $\|f\| : X \to \mathbb{R}$ e $\alpha f : X \to \mathbb{R}^n$, definidas por; $(f + g)(x) = f(x) + g(x)$; $\langle f, g\rangle(x) = \langle f(x), g(x)\rangle$, $\|f\|(x) = \|f(x)\|$ e $(\alpha f)(x) = \alpha(x)f(x)$.*

Demonstração: Esta é uma consequência do Teorema 2.16, juntamente com a Proposição 1.52. ■

Teorema 2.18. *A imagem $f(K)$ do conjunto compacto $K \subset X$ pela aplicação contínua $f : X \to \mathbb{R}^n$ é também um conjunto compacto.*

Demonstração: Seja (y_k) uma sequência de pontos em $f(K)$, para todo $k \in \mathbb{N}$, existe $x_k \in K$ tal que $f(x_k) = y_k$. Como K é compacto, pelo Teorema de Bolzano-Weierstrass, existe uma subsequência de (x_k), digamos $(x_k)_{k\in\mathbb{N}'}$ que converge para o ponto $a \in K$. Como f é uma aplicação contínua, em particular, f é contínua no ponto a. Aplicando o Teorema 2.16, temos que $\lim\limits_{k\in\mathbb{N}'} f(x_k) = f(a)$. Portanto, toda sequência de pontos $y_k = f(x_k) \in f(K)$ possui uma subsequência $(y_k)_{k\in\mathbb{N}'}$ que converge para um ponto $f(a) \in f(K)$. Logo, $f(K)$ é compacto. ■

Corolário 2.19. *(Weierstrass) Seja $K \subset \mathbb{R}^m$ compacto. Se $f : K \to \mathbb{R}$ é uma função real contínua, então existem $x_0, x_1 \in K$ tais que $f(x_0) \leq f(x) \leq f(x_1)$ para todo $x \in K$. Em outras palavras, toda função real contínua num conjunto compacto K, atinge seus valores mínimo e máximo em pontos de K.*

Demonstração: Como K é compacto, pelo Teorema 2.18, $f(K)$ também é compacto. Assim, os números $y_0 = inf\ f(K)$ e $y_1 = sup\ f(K)$ pertencem a $f(K)$, isto é, $y_0 = f(x_0)$ e $y_1 = f(x_1)$, onde x_0 e $x_1 \in K$. Portanto, $f(x_0) \leq f(x) \leq f(x_1)$, para todo $x \in K$. ■

Teorema 2.20. *Seja $X \subset \mathbb{R}^m$. A aplicação $f : X \to \mathbb{R}^n$ é contínua se, e somente se, a imagem inversa $f^{-1}(A)$ de todo conjunto aberto $A \subset \mathbb{R}^n$ é um subconjunto aberto em X.*

Demonstração: ($\Rightarrow$) Seja f contínua e $A \subset \mathbb{R}^n$ aberto, então para todo $x \in f^{-1}(A)$, existe $\varepsilon > 0$ tal que $B(f(x);\varepsilon) \subset A$. Pela continuidade de f, x é centro de uma bola aberta B_x, tal que $f(B_x \cap X) \subset B(f(x);\varepsilon) \subset A$. Assim, $x \in B_x \cap X \subset f^{-1}(A)$. Considere $U = \cup B_x$, daí, para todo $x \in f^{-1}(A)$, temos que $f^{-1}(A) \subset (U \cap X) \subset f^{-1}(A)$. Logo, $f^{-1}(A) = U \cap X$.
($\Leftarrow$) Suponha que para todo $A \subset \mathbb{R}^n$ aberto, $f^{-1}(A)$ é aberto em X, ou seja, $f^{-1}(A) = M \cap X$, onde M é aberto em $\mathbb{R}^m$. Assim, dados $x \in X$ e $\varepsilon > 0$, tomamos $A = B(f(x);\varepsilon)$ e obtemos $M \subset \mathbb{R}^m$ aberto tal que $M \cap X = f^{-1}(A)$. Certamente $x \in M$, então existe $\delta > 0$ tal que $B(x;\delta) \subset M$ e assim, $f(B(x;\delta) \cap X) \subset B(A)$. Logo, f é contínua em todos os pontos $x \in X$. ■

Teorema 2.21. *Seja $X \subset \mathbb{R}^m$. A aplicação $f : X \to \mathbb{R}^n$ é contínua se, e somente se, a imagem inversa de todo conjunto fechado $F \subset \mathbb{R}^n$ é um subconjunto $f^{-1}(F)$ fechado em X.*

Demonstração: ($\Rightarrow$) Seja $F \subseteq \mathbb{R}^n$ fechado, então $A = \mathbb{R}^n \backslash F$ é aberto. Pelo Teorema 2.20, $f^{-1}(A)$ é aberto. Uma vez que $f^{-1}(A) = f^{-1}(\mathbb{R}^n \backslash F) = f^{-1}(\mathbb{R}^n) \backslash f^{-1}(F) = X \backslash f^{-1}(F)$. Como $f^{-1}(A)$ é aberto em X, então $f^{-1}(F)$ é fechado.
($\Leftarrow$) Seja $A \subset \mathbb{R}^n$ aberto e considere $F = \mathbb{R}^n \backslash A$, como $f^{-1}(F)$ é fechado em X, temos que $f^{-1}(\mathbb{R}^n \backslash A)$ é fechado em X, como por hipótese $f^{-1}(\mathbb{R}^n \backslash A) = X \backslash f^{-1}(A)$ e $f^{-1}(A)$ é aberto em X, vemos pelo Teorema 2.20 que a aplicação f é contínua. ■

Corolário 2.22. *Seja $X \subset \mathbb{R}^m$ aberto(respect. fechado). A fim de que $f : X \to \mathbb{R}^n$ seja contínua, é necessário e suficiente que a imagem inversa por f de todo subconjunto aberto(respect. fechado) em $\mathbb{R}^n$, seja um conjunto aberto(respect. fechado) em $\mathbb{R}^m$.*

Corolário 2.23. *Sejam $f, g : X \to \mathbb{R}$ contínuas no conjunto $X \subset \mathbb{R}^m$. O conjunto $A = \{x \in X; f(x) < g(x)\}$ é aberto em X, enquanto que os conjuntos $F = \{x \in X; f(x) \leq g(x)\}$ e $G = \{x \in X; f(x) = g(x)\}$ são fechados em X.*

Demonstração: Defina $h : X \to \mathbb{R}$ tal que $h(x) = g(x) - f(x)$. Como f e g são contínuas, h é contínua. Agora considere os $x \in X$ tais que $h(x) > 0$. Mais precisamente, $A = h^{-1}((0,+\infty))$; $F = h^{-1}([0,+\infty))$ e $G = h^{-1}(\{0\})$. Como $(0,+\infty)$ é aberto em $\mathbb{R}$, $[0,+\infty)$ e $\{0\}$ são fechados em $\mathbb{R}$, temos que A é aberto e F e G são fechados. ■

Em particular, tomando g constante, vemos que o conjunto dos pontos $x \in X$ tais que $f(x) < c$ é aberto em X, enquanto que as soluções $x \in X$ da inequação $f(x) \leq c$ ou da equação $f(x) = c$, formam conjuntos fechados em X.

Teorema 2.24. *Sejam $\varphi : K \to \mathbb{R}^n$ contínua no compacto $K \subset \mathbb{R}^m$ e $L = \varphi(K)$ a imagem compacta de φ. A fim de que uma aplicação $f : L \to \mathbb{R}^p$ seja contínua, é necessário e suficiente que a composta $f \circ \varphi : K \to \mathbb{R}^p$ seja contínua.*

Figura 2.1: Representação da continuidade de uma aplicação partindo de uma imagem compacta.

$$\begin{tikzcd} K \arrow[r, "f \circ \varphi"] \arrow[d, "\varphi"'] & \mathbb{R}^p \\ L \arrow[ur, "f"'] & \end{tikzcd}$$

Fonte: LIMA, 2016, p. 23

Demonstração: Se φ e f são contínuas, então $f \circ \varphi$ é contínua, pelo Teorema 2.15. Reciprocamente, suponha que $f \circ \varphi$ é contínua, então para todo $F \subset \mathbb{R}^p$ fechado, a imagem inversa $(f \circ \varphi)^{-1}(F) = \varphi^{-1}[f^{-1}(F)]$ é um subconjunto fechado de K, logo, é compacto, pois F é compacto. Assim, pelo Teorema 2.18, e por φ ser sobrejetiva, $f^{-1}(F) = \varphi[\varphi^{-1}(f^{-1}(F))]$ é compacto, logo, fechado em $\mathbb{R}^m$. Segue que f é contínua pelo Corolário 2.22. ■

Observação 2.25. Quando se tem uma aplicação arbitrária $\varphi : K \to L$ entre dois conjuntos, para todo $Z \subset L$, vale a inclusão $\varphi[\varphi^{-1}(Z)] \subset Z$. Quando $\varphi : K \to L$ é sobrejetiva, tem-se $\varphi[\varphi^{-1}(Z)] = Z$.

2.3 Continuidade Uniforme

As funções de adição e multiplicação de números reais $s, p : \mathbb{R}^2 \to \mathbb{R}$, definidas por $s(x, y) = x + y$ e $p(x, y) = xy$, são contínuas. Agora vamos analisar a continuidade dessas funções no ponto $(a, b) \in \mathbb{R}^2$. Adotaremos a norma do máximo em $\mathbb{R}^2$, onde temos $(x, y) \in B((a, b); \delta)$ se, e somente se, $\|x - a\| < \delta$ e $\|y - b\| < \delta$.

- Adição: Como a adição é uma função contínua, dado $\varepsilon > 0$, tome $\delta = \frac{\varepsilon}{2}$. Note que $\|x - a\| < \frac{\varepsilon}{2}$ e $\|y - b\| < \frac{\varepsilon}{2}$. Ou seja, $(x, y) \in B((a, b); \delta)$, então;

$$\begin{aligned} \|s(x, y) - s(a, b)\| &= \|x + y - (a + b)\| \\ &= \|x + y - a - b\| \\ &\leq \|x - a\| + \|y - b\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

- Multiplicação: Como $xy-ab=(x-a)(y-b)+(x-a)b+a(y-b)$ e pela multiplicação ser uma função contínua, dado $\varepsilon>0$, considerando $\delta>0$ menor do que cada um dos números: $\sqrt{\frac{\varepsilon}{3}}$, $\frac{\varepsilon}{3(|a|+1)}$ e $\frac{\varepsilon}{3(|b|+1)}$, vemos que se $\|x-a\|<\delta$ e $\|y-b\|<\delta$, ou seja, $(x,y)\in B((a,b);\delta)$, então;

$$\begin{aligned}\|p(x,y)-p(a,b)\| &= |xy-ab|\\ &\le |x-a||y-b|+|x-a||b|+|a||y-b|\\ &< \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.\end{aligned}$$

Percebe-se a seguinte diferença; na adição, δ depende apenas de ε e não do ponto (a,b) onde é testada a continuidade. Mas na multiplicação, δ não depende apenas de ε, depende do ponto (a,b) também. Note ainda que se um dos números a ou b aumentar na multiplicação, para o mesmo ε, devemos considerar δ cada vez menor.

Definição 2.26. Uma aplicação $f:X\to\mathbb{R}^n$ é dita uniformemente contínua no conjunto $X\subset\mathbb{R}^m$ quando, para todo $\varepsilon>0$, for possível obter $\delta>0$ tal que $\|x-y\|<\delta$ implica $\|f(x)-f(y)\|<\varepsilon$, para quaisquer $x,y\in X$.

Assim, a adição é uniformemente contínua, enquanto que a multiplicação não é.

Teorema 2.27. *A fim de que $f:X\to\mathbb{R}^n$ seja uniformemente contínua no conjunto $X\subset\mathbb{R}^m$ é necessário e suficiente que, para todo par de sequências de pontos $x_k,y_k\in X$ com $\lim\|x_k-y_k\|=0$, se tenha $\lim\|f(x_k)-f(y_k)\|=0$.*

Demonstração: $(\Rightarrow)$ Seja $f:X\to\mathbb{R}^n$ uma aplicação uniformemente contínua em $X\subset\mathbb{R}^n$. Então, para todo $\varepsilon>0$, tem-se $\delta>0$ tal que $\|x-y\|<\delta$ implica $\|f(x)-f(y)\|<\varepsilon$, para todo $x,y\in X$. Considere $(x_k),(y_k)\subset X$ com $\lim\limits_{k\to\infty}\|x_k-y_k\|=0$. Por definição de limite, para todo $\delta>0$ obtemos $k_0\in\mathbb{N}$ tal que $k>k_0$ implica $\|x_k-y_k\|<\delta$. Assim, segue que $\|f(x_k)-f(y_k)\|<\varepsilon$, para todo $k>k_0$. Ou seja, $\lim\limits_{k\to\infty}\|f(x_k)-f(y_k)\|=0$.
$(\Leftarrow)$ Suponha, por contradição, que f não é uniformemente contínua. Logo, existe $\varepsilon>0$ tal que para todo $\delta>0$, encontramos $x_\delta,y_\delta\in X$ com $\|x_\delta-y_\delta\|<\delta$ e $\|f(x_\delta)-f(y_\delta)\|\ge\varepsilon$. Tome $\delta\in\{1,\frac{1}{2},...,\frac{1}{k},...\}$. Assim, existem $(x_k),(y_k)\subseteq X$ tais que $0\le\|x_k-y_k\|<\frac{1}{k}$ e $\|f(x_k)-f(y_k)\|\ge\varepsilon$. Portanto, $0\le\lim\limits_{k\to\infty}\|x_k-y_k\|\le\lim\limits_{k\to\infty}\dfrac{1}{k}=0$. Ou seja, $\lim\limits_{k\to\infty}\|x_k-y_k\|=0$. Se $\lim\limits_{k\to\infty}\|f(x_k)-f(y_k)\|=0$, então $0=\lim\limits_{k\to\infty}\|f(x_k)-f(y_k)\|\ge\varepsilon$. O que é um absurdo, pois $\varepsilon>0$. ■

Teorema 2.28. *Toda aplicação contínua $f:X\to\mathbb{R}^n$, definida num conjunto compacto $X\subset\mathbb{R}^m$, é uniformemente contínua.*

Demonstração: Se f não fosse uma aplicação uniformemente contínua, existiriam $\varepsilon>0$ e duas sequências $(x_k),(y_k)$ em X satisfazendo $\lim\limits_{k\to\infty}(y_k-x_k)=0$ e $\|f(y_k)-f(x_k)\|\ge\varepsilon$

para todo $k \in \mathbb{N}$. Como X é compacto, existem $a \in X$ e $N_1 \subset \mathbb{N}$, tal que $\lim\limits_{\substack{k\to\infty \\ k\in N_1}} x_k = a$. Assim, como $y_k = (y_k - x_k) + x_k$, vale também que $\lim\limits_{\substack{k\to\infty \\ k\in N_1}} y_k = a$. Sendo f contínua no ponto a, temos que

$$\begin{aligned} \lim_{\substack{k\to\infty \\ k\in N_1}} (f(y_k) - f(x_k)) &= \lim_{\substack{k\to\infty \\ k\in N_1}} f(y_k) - \lim_{\substack{k\to\infty \\ k\in N_1}} f(x_k) \\ &= f(a) - f(a) = 0. \end{aligned}$$

Contradizendo que $\|f(y_k) - f(x_k)\| \geq \varepsilon$, para todo $k \in \mathbb{N}$. ■

Definição 2.29. Uma aplicação $f : X \to \mathbb{R}^n$, definida no conjunto $X \subset \mathbb{R}^m$, chama-se Lipschitziana quando existe $c > 0$ tal que $\|f(x) - f(y)\| \leq c\|x - y\|$, para quaisquer $x, y \in X$.

Observação 2.30. O número c é chamado constante de Lipschitz de f.

Exemplo 2.31. Toda aplicação Lipschitziana é uniformemente contínua.

Seja $f : X \to \mathbb{R}^n$ uma aplicação Lipschitziana cuja constante de Lipschitz é $c > 0$. Dado $\varepsilon > 0$, considere $\delta = \frac{\varepsilon}{c} > 0$. Assim, para todos $x, y \in X$ com $\|x - y\| < \delta = \frac{\varepsilon}{c}$, temos $\|f(x) - f(y)\| \leq c\|x - y\| < c \cdot \frac{\varepsilon}{c} < \varepsilon$.

Perceba que a recíproca não é válida, pois, dado $\varepsilon > 0$, considere $\delta = \frac{\varepsilon}{c}$. A função $f : [0, 1] \to \mathbb{R}$ definida por $f(x) = \sqrt{x}$, é uniformemente contínua mas não é Lipschitziana, visto que, $|\sqrt{x} - \sqrt{y}| = \frac{1}{\sqrt{x}+\sqrt{y}}|x - y|$ e que, com $x, y \in [0, 1]$ pode-se tomar $\sqrt{x} + \sqrt{y}$ tão pequeno (logo $(\sqrt{x} + \sqrt{y})^{-1}$ tão grande) quanto se queira.

Definição 2.32. O número $\|T\| = sup\{\|T \cdot x\|; x \in S^{m-1}\}$ chama-se a norma da transformação linear $T : \mathbb{R}^m \to \mathbb{R}^n$.

Exemplo 2.33. Toda transformação linear $T : \mathbb{R}^m \to \mathbb{R}^n$ é contínua.

De fato, para cada $i = 1, 2, ..., n$, a i-ésima função coordenada de T é a função contínua $(x_1, ..., x_m) \to a_{i_1}x_1 + ... + a_{i_m}x_m$, onde $[a_{ij}]$ é a matriz de T. A esfera unitária $S^{m-1} = \{x \in \mathbb{R}^m; \|x\| = 1\}$ é compacta. Logo T é limitada em S^{m-1}. E para todo vetor $v \in \mathbb{R}^m$, tem-se $\|T \cdot v\| \leq \|T\| \cdot \|v\|$. Isto é óbvio quando $v = 0$. Se $v \neq 0$, então $\frac{v}{\|v\|} \in S^{m-1}$, logo;

$$\|T \cdot v\| = \|v\| \cdot \left\| T \cdot \left(\frac{v}{\|v\|} \right) \right\| \leq \|T\| \cdot \|v\|.$$

Para $x, y \in \mathbb{R}^m$ quaisquer, tem-se $\|Tx - Ty\| = \|T(x - y)\| \leq \|T\| \cdot \|x - y\|$. Logo, a transformação linear T é uma aplicação Lipschitziana, com constante de Lipschitz $\|T\|$. Pelo exemplo 2.31, T é contínua.

Exemplo 2.34. Dado $A \subset \mathbb{R}^n$ não-vazio, seja $f : \mathbb{R}^n \to \mathbb{R}$ definida por $f(x) = d(x, A)$. Afirmamos que $|d(x, A) - d(y, A)| \leq \|x - y\|$, para quaisquer $x, y \in \mathbb{R}^n$.

Neste caso, $f(x) = d(x, A)$ é Lipschitziana com constante $c = 1$, portanto, uniformemente contínua de acordo com o exemplo 2.31. Provemos a afirmação: dados $x, y \in \mathbb{R}^n$, existem $\overline{a}, \overline{b} \subset \overline{A}$ tais que $d(x, A) = \|x - \overline{a}\|$ e $d(y, A) = \|y - \overline{b}\|$. Temos $\overline{b} = \lim y_k$, com $y_k \in A$. Como $\|x - \overline{a}\| \leq \|x - y_k\|$ para todo $k \in \mathbb{N}$, segue que $\|x - \overline{a}\| \leq \|x - \overline{b}\|$. Sem perda de generalidade, podemos supor que $d(x, A) \geq d(y, A)$, logo;

$$\begin{aligned}
|d(x, A) - d(y, A)| &= d(x, A) - d(y, A) \\
&= \|x - \overline{a}\| - \|y - \overline{b}\| \\
&\leq \|x - \overline{b}\| - \|y - \overline{b}\| \\
&\leq \|x - y\|, \text{ pois } \|x - \overline{b}\| \leq \|x - y\| + \|y - \overline{b}\|.
\end{aligned}$$

Definição 2.35. A aplicação Lipschitziana $f : X \to \mathbb{R}^n$ é chamada contração, se $\|f(x) - f(y)\| \leq c\|x - y\|$ com $0 < c < 1$, para quaisquer $x, y \in X$. No caso em que $c = 1$, f é chamada de contração fraca.

2.4 Homeomorfismos

Definição 2.36. Um homeomorfismo do conjunto $X \subset \mathbb{R}^m$ sobre um conjunto $Y \subset \mathbb{R}^n$, é uma bijeção contínua $f : X \to Y$ cuja inversa $f^{-1} : Y \to X$ também é contínua.

Definição 2.37. O gráfico de uma aplicação $f : X \to \mathbb{R}^n$, definida no conjunto $X \subset \mathbb{R}^m$ é o conjunto $G = \{(x, f(x)); x \in X\} \subset \mathbb{R}^m \times \mathbb{R}^n$.

Observação 2.38. Se $f : X \to \mathbb{R}^n$ é uma aplicação contínua, então seu gráfico G é homeomorfo a seu domínio $X \subset \mathbb{R}^m$. Mais especificamente $\varphi : X \to G$, dada por $\varphi(x) = (x, f(x))$, é um homeomorfismo, cujo inverso $\varphi^{-1}(x, f(x)) = x$ que é a restrição a G da projeção de $\mathbb{R}^m \times \mathbb{R}^n$ sobre $\mathbb{R}^m$.

Exemplo 2.39. A aplicação $f : [0, 2\pi) \to S^1$, onde $f(t) = (cos(t), sen(t))$, é uma bijeção contínua mas não é um homeomorfismo.

Sabendo que S^1 é a circunferência de centro 0 e raio 1, note que a inversa de f, $f^{-1} : S^1 \to [0, 2\pi)$ aplica S^1, que é compacto, sobre o intervalo $[0, 2\pi)$ não fechado, logo, não compacto. Portanto f^{-1} não é contínua. De modo particular, perceba que f^{-1} é descontínua no ponto $a = (1, 0) = f(0) \in S^1$. Com efeito, se pusermos para cada $k \in \mathbb{N}$, $t_k = \left(1 - \frac{1}{k}\right) 2\pi$ e $z_k = (cos(t_k), sen(t_k))$, teremos $\lim\limits_{k\to\infty} z_k = (cos(2\pi), sen(2\pi)) = (1, 0) = a$, contudo, $\lim\limits_{k\to\infty} f^{-1}(z_k) = \lim\limits_{k\to\infty} t_k = 2\pi$. Assim, não vale $\lim\limits_{k\to\infty} f^{-1}(z_k) = f^{-1}(a) = 0$.

Se $f : X \longrightarrow Y$ é um homeomorfismo, então dizemos que X é homeomorfo a Y ou que Y é homeomorfo a X.

Exemplo 2.40. A bola aberta $B = B(0; 1) \subset \mathbb{R}^n$ é homeomorfa ao espaço $\mathbb{R}^n$.

De fato, as aplicações $f : \mathbb{R}^n \to B$ e $g : B \to \mathbb{R}^n$ definidas por $f(x) = \frac{x}{1+\|x\|}$ e $g = \frac{y}{1-\|y\|}$ são contínuas e como $\left\|\frac{x}{1+\|x\|}\right\| = \frac{\|x\|}{1+\|x\|} < \frac{\|x\|}{\|x\|} = 1$. Segue que,

$$(g \circ f)(x) = g(f(x)) = g\left(\frac{x}{1+\|x\|}\right) = \frac{\frac{x}{1+\|x\|}}{1 - \left\|\frac{x}{1+\|x\|}\right\|} = \frac{\frac{x}{1+\|x\|}}{1 - \frac{\|x\|}{1+\|x\|}} = \frac{\frac{x}{1+\|x\|}}{\frac{1+\|x\|-\|x\|}{1+\|x\|}} = x,$$

para quaisquer $x \in \mathbb{R}^n$ e analogamente $(f \circ g)(x) = f(g(x)) = f\left(\frac{y}{1+\|y\|}\right) = y$, para quaisquer $y \in B$. Assim, $g = f^{-1}$.

Exemplo 2.41. Sejam $S^n = \{x \in \mathbb{R}^{n+1}; \langle x, x\rangle = 1\}$ a esfera unitária n-dimensional e $N = (0, ..., 0, 1) \in S^n$ seu polo norte. A projeção estereográfica $\xi : S^n \backslash \{N\} \to \mathbb{R}^n$ é um homeomorfismo.

Note que para todo $x \in S^n \backslash \{N\}$, $\xi(x)$ é o ponto em que a semi-reta $\overrightarrow{N_x}$ corta o hiperplano $x_{n+1} = 0$, o qual identificamos com $\mathbb{R}^n$. Além disso, os pontos da semi-reta $\overrightarrow{N_x}$ são da forma $N + t(x - N)$ com $t > 0$, e um ponto está no hiperplano $\mathbb{R}^n$ quando sua última coordenada $1 + t(x_{n+1} - 1)$ é igual a zero, isto é, quando $f(x) = \frac{1}{1-x_{n+1}}$. Logo, $\xi(x) = \frac{x'}{1-x_{n+1}}$ onde $x' = (x_1, ..., x_n)$ para $x = (x_1, ..., x_{n+1})$. Portanto, $\xi : S^n \backslash N \to \mathbb{R}^n$ é contínua. Agora considere $\varphi : \mathbb{R}^n \to S^n \backslash \{N\}$ dada por $\varphi(y) = x$, onde $x' = \frac{2y}{\|y\|^2+1}$ e $x_{n+1} = \frac{\|y\|^2-1}{\|y\|^2+1}$. Note que

$$\xi(\varphi(y)) = \xi(x) = \frac{\frac{2y}{\|y\|^2+1}}{1 - \frac{\|y\|^2-1}{\|y\|^2+1}} = \frac{\frac{2y}{\|y\|^2+1}}{\frac{\|y\|^2+1-\|y\|^2+1}{\|y\|^2+1}} = \frac{2y(\|y\|^2+1)}{(\|y\|^2+1)2} = y, \text{ para todo } y \in \mathbb{R}^n.$$

Pela unicidade da inversa de ξ, segue que $\varphi = \xi^{-1}$. Assim, $\varphi(\xi(x)) = x$ e consequentemente, ξ é um homeomorfismo.

Teorema 2.42. *Se $K \subset \mathbb{R}^m$ é compacto, então toda aplicação contínua injetiva $f : K \to \mathbb{R}^n$ é um homeomorfismo sobre sua imagem compacta $L = f(K)$.*

Demonstração. Conforme o Teorema 2.21, basta-nos mostrar que $f^{-1} : L \to K$ é contínua. Seja $F \subseteq K$ um conjunto fechado qualquer, então $(f^{-1})^{-1}(F) = f(F)$. Pela compacidade de K e $F \subseteq K$ ser fechado, então F é compacto. Como f é uma aplicação contínua, $f(F)$ é um compacto, em particular, $f(F)$ é fechado. Assim, $f^{-1} : L \to K$ é contínua.

2.5 Conjuntos Conexos

Aqui vamos estudar uma relevante propriedade a respeito de conjuntos, que é preservada pelas funções contínuas, a conexidade.

Definição 2.43. Uma cisão do conjunto $X \subset \mathbb{R}^n$ é uma decomposição $X = A \cup B$ onde $\overline{A} \cap B = A \cap \overline{B} = \varnothing$, isto é, nenhum ponto de A é aderente a B e nenhum ponto de B é aderente a A.

Observação 2.44. $A \cap B \subseteq \overline{A} \cap B = \varnothing$. Isto é, a união $X = A \cup B$ é disjunta.

Exemplo 2.45. Sejam $X \subseteq \mathbb{R}^n$ não vazio, $A = X$ e $B = \varnothing$, então $X = A \cup B$ é uma cisão para X e essa cisão é chamada de cisão trivial.

Exemplo 2.46. $\mathbb{R}\backslash\{0\} = (-\infty, 0) \cup (0, +\infty)$ é uma cisão não-trivial para o conjunto $\mathbb{R}\backslash\{0\}$, pois $\overline{(-\infty, 0)} = (-\infty, 0]$ e $\overline{(0, +\infty)} = [0, +\infty)$. Assim, $(-\infty, 0] \cap (0, +\infty) = (-\infty, 0) \cap [0, +\infty) = \varnothing$.

Exemplo 2.47. Sejam $A = (-\infty, 0]$ e $B = (0, +\infty)$, a decomposição $\mathbb{R} = A \cup B$ não é uma cisão, pois como $A = \overline{A}$ e $\overline{B} = [0, +\infty)$, temos que $\overline{A} \cap B = (-\infty, 0] \cap (0, +\infty) = \varnothing$ e $A \cap \overline{B} = (-\infty, 0] \cap [0, +\infty) = \{0\}$, portanto $\overline{A} \cap B \neq A \cap \overline{B}$.

Proposição 2.48. *Se $X = A \cup B$ é uma cisão, então A e B são simultaneamente abertos e fechados em X.*

Demonstração: Como $X = A \cup B$ é uma cisão, então $\overline{A} = \overline{A} \cap X = \overline{A} \cap (A \cup B) = (\overline{A} \cap A) \cup (\overline{A} \cap B) = A \cup \varnothing = A$. Como $A = \overline{A} \cap X$, A é fechado em X. Analogamente para B, e concluímos que B é fechado em X. Por outro lado, como $X = A \cup B$, temos que $X\backslash B = A$, como B é fechado em X, seu complementar A é aberto em X. De maneira análoga, concluímos que B é aberto em X, pois $X\backslash A = B$ e como A é fechado em X, segue que $X\backslash A = B$ é aberto em X. ■

Reciprocamente,

Proposição 2.49. *Se $A \subset X$ é aberto e fechado em X e $B = X\backslash A$, então a decomposição $X = A \cup B$ é uma cisão.*

Demonstração: Como A é fechado em X e $B = X\backslash A$, nenhum ponto de X aderente a A pode pertencer a B, pois se isso acontecesse, digamos que exista um $c \in (\overline{A} \cap B)$, mas como $A = \overline{A} \cap X$, então $c \in (A \cap B) = \varnothing$, portanto $\overline{A} \cap B = \varnothing$, uma contradição. Analogamente, como B é fechado em X e $B = X\backslash A$, nenhum ponto de X aderente a B pode pertencer a A, pois se um certo $d \in (\overline{B} \cap A)$, note que, como $B = \overline{B} \cap X$, então $d \in (A \cap B) = \varnothing$, portanto $\overline{B} \cap A = \varnothing$. ■

Observação 2.50.

- Se $X \subset \mathbb{R}^n$ é aberto (fechado), uma cisão $X = A \cup B$ é uma expressão de X como união de dois abertos (fechados) disjuntos.
- Se X é compacto e $X = A \cup B$ é uma cisão, então A e B são compactos.

Proposição 2.51. *Se $X = A \cup B$ é uma cisão, então para todo $Z \subset X$, $Z = (A \cap Z) \cup (B \cap Z)$ é uma cisão.*

Demonstração: Como $X = A \cup B$ e $Z \subset X$, temos que $(A \cap Z) \subset A$ e $(B \cap Z) \subset B$. Assim, $\overline{(A \cap Z)} \cap (B \cap Z) \subset \overline{A} \cap B = \varnothing$. Por outro lado, temos que $(A \cap Z) \cap \overline{(B \cap Z)} \subset A \cap \overline{B} = \varnothing$. Portanto, $Z = (A \cap Z) \cup (B \cap Z)$ é uma cisão. ■

Definição 2.52. Um conjunto $X \subset \mathbb{R}^n$ chama-se conexo quando só admite a cisão trivial. Caso contrário, diz-se que X é desconexo.

Teorema 2.53. *Os únicos subconjuntos conexos de $\mathbb{R}$ são os intervalos.*

Demonstração: Suponha, por contradição, que exista $X \subset \mathbb{R}$ conexo tal que X não é um intervalo. Assim, existem $a, b \in X$ e $c \notin X$ tais que $a < c < b$. Considerando $A = \{x \in X; x < c\}$ e $B = \{x \in X; x > c\}$, temos que $X = A \cup B$ é uma cisão, contudo, note que nenhum dos conjuntos que expressam X por meio da união é vazio, visto que $a \in A$ e $b \in B$. Logo, a cisão $X = A \cup B$ não é trivial, portanto X não é conexo. ■

Teorema 2.54. *i) A imagem do conjunto conexo $X \subset \mathbb{R}^m$ por uma aplicação contínua $f : X \to \mathbb{R}^n$ é um conjunto conexo.*

ii) A união $X = \cup_{\lambda \in L} X_\lambda$ de uma família qualquer de conjuntos conexos $X_\lambda \subset \mathbb{R}^n$ que têm um ponto a em comum é um conjunto conexo.

iii) 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 um conjunto conexo se, e somente se, X e Y são conexos.

iv) Se $X \subset \mathbb{R}^n$ é conexo e $X \subset Y \subset \overline{X}$, então Y é conexo. Em particular, o fecho de um conjunto conexo é conexo.

Demonstração:

i) Se $f(X) = A \cup B$ é uma cisão da imagem de X, então A e B são simultaneamente fechados e abertos em $f(X)$, além de A e B serem disjuntos. Assim, $f^{-1}(A)$ e $f^{-1}(B)$ também são disjuntos abertos e fechados em X, portanto $X = f^{-1}(A) \cup f^{-1}(B)$ é uma cisão trivial, visto que X é conexo. Além disso, $A = f(f^{-1}(A))$ e $B = f(f^{-1}(B))$, pois $A, B \subset f(X)$. Logo, A ou B é vazio, o que torna $f(X) = A \cup B$ uma cisão trivial. Portanto, $f(X)$ é conexo.

ii) Como existe $a \in X_\lambda$ para todo $\lambda \in L$, então $X = \cap_{\lambda \in L} X_\lambda \neq \varnothing$. Seja $X = A \cup B$ uma cisão. Então para todo $\lambda \in L$, $X_\lambda = X_\lambda \cap X = X_\lambda \cap (A \cup B) = (X_\lambda \cap A) \cup (X_\lambda \cap B)$. Note ainda que $\overline{(X_\lambda \cap A)} \cup (X_\lambda \cap B) \subseteq \overline{A} \cap B = \varnothing$, pois $X = A \cup B$ é uma cisão. Analogamente temos que $(X_\lambda \cap A) \cup \overline{(X_\lambda \cap B)} = \varnothing$. Assim, $X_\lambda = (X_\lambda \cap A) \cup (X_\lambda \cap B)$ é uma cisão e como X_λ é conexo, ou $(X_\lambda \cap A) = \varnothing$ e $(X_\lambda \cap B) = X_\lambda$, ou $(X_\lambda \cap A) = X_\lambda$ e $(X_\lambda \cap B) = \varnothing$. Por outro lado, $a \in X = A \cup B$ e $A \cap B = \varnothing$, portanto, ou $a \in A$ ou $a \in B$.

- Se $a \in A$, $a \in (X_\lambda \cap A) \neq \varnothing$, então $(X_\lambda \cap A) = X_\lambda$ e $(X_\lambda \cap B) = \varnothing$. Logo, $B = B \cap X = B \cap (\cup_{\lambda \in L} X_\lambda) = \cup_{\lambda \in L}(B \cap X_\lambda) = \varnothing$. Portanto $A = X$.
- Se $a \in B$, concluímos de modo análogo que $A = \varnothing$ e $B = X$. Portanto, X é conexo.

iii) ($\Rightarrow$) Considere as projeções;

$$\begin{aligned} p : X \times Y &\longrightarrow X \\ (x,y) &\longmapsto p(x,y) = x \end{aligned} \quad e \quad \begin{aligned} q : X \times Y &\longrightarrow X \\ (x,y) &\longmapsto q(x,y) = y. \end{aligned}$$

Como $X \times Y$ é conexo, $p(x,y) = x$ e $q(x,y) = y$ são contínuas, pelo item *i)*, X e Y são conexos.
($\Leftarrow$) Fixe $(a,b) \in X \times Y$ de forma arbitrária e para cada $z = (x,y) \in X \times Y$ considere $C_z = (X \times \{b\}) \cup (\{x\} \times Y) \subset X \times Y$. Perceba que $(a,b) \in C_z$ para todo $z \in X \times Y$, definindo a função;

$$\begin{aligned} \varphi : X &\longrightarrow X \times \{b\} \\ x &\longmapsto \varphi(x) = (x,b), \end{aligned}$$

temos que φ é contínua e portanto $X \times \{b\}$ é conexo, da mesma forma concluímos que $\{x\} \times Y$ é conexo. Além disso, $(x,b) \in X \times \{b\}$ e $(x,b) \in \{x\} \times Y$. Assim, pelo item *ii)*, C_z é conexo para todo $z \in X \times Y$. Como $(a,b) \in C_z$ para todo $z \in X \times Y$, pelo item *ii)*, novamente, $X \times Y = \cup_{z \in X \times Y} C_z$ é conexo.

iv) Seja $X = A \cup B$ uma cisão. Então $X = (A \cap X) \cup (B \cap X)$ também é uma cisão, pela Proposição 2.51, já que $X \subset Y$. Como X é conexo, então ou $(A \cap X)$ ou $(B \cap X)$ é vazio. Digamos que $(A \cap X) = \varnothing$. Por $X \subset Y$ e $Y = A \cup B$, temos que $X \subset B$, logo $\overline{X} \subset \overline{B}$. Assim $Y \subset \overline{B}$, pois $Y \subset \overline{X}$. Daí, $A = A \cap Y \subset A \cap \overline{B} = \varnothing$, ou seja, $A = \varnothing$ e, consequentemente, $B = Y$. Para $B \cap X = \varnothing$, de maneira análoga, concluímos que $B = \varnothing$ e $A = Y$. Logo, toda cisão $Y = A \cup B$ é trivial, o que torna Y conexo.

■

Corolário 2.55. *Se $X_1, ..., X_k$ são conexos, então $X_1 \times ... \times X_k$ é conexo. Em particular, $\mathbb{R}^n = \mathbb{R} \times ... \times \mathbb{R}$ é conexo.*

Demonstração: Com efeito, $X_1 \times X_2 \times X_3 = (X_1 \times X_2) \times X_3$ e assim por diante, conforme o item *iii)* do Teorema 2.54. ■

Corolário 2.56. *Se $X \subset \mathbb{R}^n$ é conexo, então a imagem de toda função real contínua $f : X \to \mathbb{R}$ é um intervalo.*

Demonstração: Com efeito, pelo item *i)* do Teorema 2.54 $f(X)$ é um conexo e pelo Teorema 2.53 , todo subconjunto conexo de $\mathbb{R}$ é um intervalo. ■

Este corolário é conhecido como o Teorema do Valor Intermediário, pois pode também ser enunciado assim: Sejam $X \subset \mathbb{R}^n$ conexo e $f : X \to \mathbb{R}$ contínua. Se $a, b \in X$ são tais que $f(a) < f(b)$ então, para cada d com $f(a) < d < f(b)$, existe $c \in X$ tal que $f(c) = d$.

Teorema 2.57. *(Da Alfândega) Seja $X \subset \mathbb{R}^n$ um conjunto arbitrário. Se um conjunto conexo $C \subset \mathbb{R}^n$ contém um ponto $a \in X$ e um ponto $b \notin X$, então C contém um ponto $c \in fr.X$.*

Demonstração: Definamos a seguinte função auxiliar $f : C \to \mathbb{R}$, definida por $f(x) = d(x, X) - d(x, \mathbb{R}^n \setminus X)$. Note que f é contínua e como C é conexo, pelo Teorema 2.54, item *i*), $f(C)$ é um conexo de $\mathbb{R}^n$. Além disso, note que $f(a) = -d(a, \mathbb{R}^n \setminus X) \leq 0$ e $f(b) = d(b, X) \geq 0$. Assim, pelo Teorema do Valor Intermediário, existe $c \in C$ tal que $f(c) = 0$, ou seja, $d(c, X) - d(c, \mathbb{R}^n \setminus X) = 0$, assim, $d(c, X) = d(c, \mathbb{R}^n \setminus X)$. Mas, ou $c \in X$ ou $c \in \mathbb{R}^n \setminus X$, e em ambos os casos, teríamos $d(c, X) = 0 = d(c, \mathbb{R}^n \setminus X)$. Portanto, $c \in C \cap fr.X$. ■

Exemplo 2.58. A esfera S^n é um conjunto conexo para todo $n \in \mathbb{N}$.

Com efeito, retirando o polo Norte $N = (0, ..., 0, 1)$, vimos no exemplo 2.41 que $X = S^n \setminus \{N\}$ é conexo por ser homeomorfo a $\mathbb{R}^n$. Como $S^n = \overline{X}$, pelo item *iv*) do Teorema 2.54, a esfera S^n é conexa.

Exemplo 2.59. Para toda função real contínua $f : S^1 \to \mathbb{R}$ existe pelo menos um ponto $z \in S^1$ tal que $f(z) = f(-z)$. Isto é uma consequência do Teorema do Valor Intermediário.

De fato, considere a seguinte função contínua $\varphi : S^1 \to \mathbb{R}$ dada por $\varphi(z) = f(z) - f(-z)$. Note que $\varphi(-z) = f(-z) - f(z)$, assim, $\varphi(-z) = -\varphi(z)$. Logo, ou $\varphi(z) = 0$ para todo z, ou existe $a \in S^1$ com $\varphi(-a) < 0 < \varphi(a)$, então $\varphi(z) = 0$ para algum $z \in S^1$, pelo Teorema do Valor Intermediário, pois S^1 é conexo.

Existe uma noção bem geométrica que fornece uma condição suficiente para a conexidade de um conjunto, que é a conexidade por caminhos.

Definição 2.60. Um caminho num conjunto $X \subset \mathbb{R}^n$ é uma aplicação contínua $\gamma : I \to X$, definida num intervalo I.

Exemplo 2.61. Dados $x, y \in \mathbb{R}^n$, o caminho $\gamma : [0, 1] \to \mathbb{R}^n$, definido por $\gamma(t) = (1 - t)x + ty$, chama-se o caminho retilíneo que liga x a y. Por vezes, nos referimos a ele como o caminho $[x, y]$.

Dizemos que os pontos $a, b \in X$ podem ser ligados por um caminho em X, quando existe um caminho $\gamma : I \to X$ tal que $a = \gamma(\alpha)$, $b = \gamma(\beta)$, com $\alpha < \beta \in I$.

Exemplo 2.62. Se $X \subset \mathbb{R}^n$ é convexo, dois pontos quaisquer $a, b \in X$ podem ser ligados por um caminho em X, a saber, o caminho retilíneo $[a, b]$.

Observação 2.63. Se $a, b \in X$ podem ser ligados por um caminho $\gamma : [\alpha, \beta] \to X$, então existe um caminho $\varphi : [0, 1] \to X$ tal que $\varphi(0) = a$ e $\varphi(1) = b$. Basta considerarmos $\varphi(t) = \gamma((1 - t)\alpha + t\beta)$, onde $a = \varphi(\alpha)$ e $b = \varphi(\beta)$.

Definição 2.64. Se $\alpha, \beta : [0,1] \to X$ são caminhos em X, com $\alpha(1) = \beta(0)$, então definimos o caminho justaposto $\gamma = \alpha \vee \beta : [0,1] \to X$, pondo $\gamma(t) = \alpha(2t)$ se $0 \leq t \leq \frac{1}{2}$ e $\gamma(t) = \beta(2t-1)$ se $\frac{1}{2} \leq t \leq 1$. Como $\gamma \mid_{[0,\frac{1}{2}]}$ e $\gamma \mid_{[\frac{1}{2},1]}$ são contínuas, segue que γ é contínua.

Observação 2.65. Intuitivamente, o caminho h percorre a trajetória de α (com "velocidade" dobrada) até $t = \frac{1}{2}$ e depois, para $t \geq \frac{1}{2}$, descreve (ainda com "velocidade" dobrada) o percurso de β.

Note que, dados a, b, c pontos do conjunto $X \subset \mathbb{R}^n$. Se a, b podem ser ligados por um caminho em X e b, c também podem ser ligados por um caminho em X, então existe um caminho em X, ligando a a c. Basta tomar caminhos $\alpha, \beta : [0,1] \to X$ com $\alpha(0) = a$, $\alpha(1) = b$, $\beta(0) = b$, $\beta(1) = c$ e considerarmos $\gamma : \alpha \vee \beta$. Então $\gamma(0) = a$ e $\gamma(1) = c$.

Definição 2.66. Um conjunto $X \subset \mathbb{R}^n$ é dito conexo por caminhos, quando dois pontos quaisquer $a, b \in X$ podem ser ligados por um caminho em X.

Exemplo 2.67. Todo conjunto convexo $X \subset \mathbb{R}^n$ é conexo por caminhos. Em particular, toda bola aberta ou fechada no espaço Euclidiano é conexa por caminhos.

Exemplo 2.68. A esfera $S^n = \{x \in \mathbb{R}^{n+1}; \langle x, x\rangle = 1\}$ é conexa por caminhos.

Dados $a, b \in S^n$, se a e b não são antípodas, ou seja, se $b \neq -a$, então $\gamma : [0,1] \to S^n$, definida por

$$\gamma(t) = \frac{(1-t)a + tb}{\|(1-t)a + tb\|},$$

é contínua pois o denominador é sempre positivo, com $\gamma(0) = a$ e $\gamma(1) = b$. Se, contudo, $b = -a$, tomamos um ponto $c \in S^n \backslash \{a, b\}$, ligamos a com c e c com b. O caminho justaposto ligará o ponto a ao seu antípoda b.

Proposição 2.69. *Se $X \subseteq \mathbb{R}^n$ é conexo por caminhos, então X é conexo.*

Demonstração: Fixando $a \in X$, para cada $x \in X$ existe um caminho $\varphi_x : [0,1] \to X$ tal que $\varphi_x(0) = a$ e $\varphi_x(1) = x$. Como $[0,1]$ é conexo e φ_x é contínua, para todo $x \in X$, segue que $\varphi_x([0,1])$ é conexo, pelo item $i)$ do Teorema 2.54. Além disso, $a \in \varphi([0,1])$ para todo $x \in X$, então pelo item $ii)$ do Teorema 2.54, segue que $\cup_{x \in X} \varphi([0,1]) = X$ é conexo. ■

Contudo, a recíproca da proposição anterior é falsa, pois considere $X \subset \mathbb{R}^2$ em que X é formado pela união do gráfico de $f(x) = cos(\frac{1}{x})$ onde $0 < x \leq 1$, com o ponto $p = \{(0,0)\}$. Note que por f ser contínua e $(0,1]$ ser conexo, $G(f)$ é conexo em $\mathbb{R}^2$, pois $G(f) = \{(x, f(x)); x \in (0,1]\}$. Além disso, $G(f) \subseteq X \subseteq \overline{G(f)}$. Assim, pelo item $iv)$ do Teorema 2.54, X é conexo. Porém, X não é conexo por caminhos, visto que, não existe $\varphi : [0,1] \to X$ contínua tal que $\varphi(0) = p$ e $\varphi(1) = (1, cos(1))$.

No entanto, existe um caso particular importante, no qual a conexidade implica em conexidade por caminhos; quando o conjunto $X \subset \mathbb{R}^n$ é aberto.

Definição 2.70. $\lambda : [0,1] \to X$ é um caminho poligonal em X, quando f é a justaposição de um número finito de caminhos retilíneos.

Teorema 2.71. *Seja $X \subset \mathbb{R}^n$ um aberto. X é conexo por caminhos se, e somente se, X é conexo.*

Demonstração: Pela Proposição 2.69, em particular se $X \subset \mathbb{R}^n$ é aberto, conexo por caminhos, então X é conexo. Basta-nos provar que se X é conexo, então X é conexo por caminhos. Fixe $x_0 \in X$ arbitrário e defina

$$V = \{x \in X; \text{ existe } \varphi : [0,1] \to X \text{ contínuo tal que } \varphi(0) = x_0 \text{ e } \varphi(1) = x\}.$$

Como X é aberto, existe $r_0 > 0$ de modo que $B(x_0, r_0) \subset X$. Como $B(x_0, r_0)$ é convexa, segue que $B(x_0, r_0) \subset V$. Assim, $V \neq \varnothing$. Além disso, note que $X = V \cup (X \backslash V)$ é uma cisão para X, pois V é aberto, e como X é conexo, segue que $V = X$ e $X \backslash V = \varnothing$. Pela definição do conjunto V, concluímos que X é conexo por caminhos. ■

Corolário 2.72. *Se $A \subset \mathbb{R}^n$ é aberto e conexo, dois pontos quaisquer de A podem ser ligados por um caminho poligonal contido em A.*

Adiante, veremos que todo conjunto $X \subset \mathbb{R}^n$ se exprime como união disjunta de subconjuntos conexos máximos, chamados componentes conexas de X.

Definição 2.73. Sejam $x \in X \subset \mathbb{R}^n$. A componente conexa do ponto x no conjunto X é a união C_x de todos os subconjuntos conexos de X que contêm o ponto x.

Exemplo 2.74. Seja $X = \mathbb{Q} \subset \mathbb{R}$, a componente conexa de qualquer ponto $x \in X$ é $\{x\}$.

Exemplo 2.75. Seja $X \subset \mathbb{R}^n$ conexo, para todo $x \in X$ temos $C_x = X$.

Exemplo 2.76. Seja $X = \mathbb{R} \backslash \{0\}$, a componente conexa de 1 em X é $(0, +\infty)$, enquanto que a componente conexa de -1 em X é $(-\infty, 0)$.

Observação 2.77. Dados $x \in X \subset \mathbb{R}^n$, perceba que a componente conexa C_x é um conjunto conexo, conforme o item $ii)$ do Teorema 2.54.

Proposição 2.78. *Seja $x \in X \subset \mathbb{R}^n$, a componente conexa C_x é o maior subconjunto conexo de X contendo o ponto x.*

Demonstração: Se $C \subset X$ é conexo e contém x, então C é um dos conjuntos cuja união é C_x, logo, $C \subset C_x$. Além disso, se $C \subset C_x$ é conexo e tem algum ponto em comum com C_x, então $C \subset C_x$, pois $C \cup C_x$ é conexo contendo x, portanto $C \cup C_x \subset C_x$, daí $C \subset C_x$. Em particular, nenhum subconjunto conexo de X pode conter C_x propriamente. ■

Observação 2.79. Sejam x, y dois pontos de X. Suas componentes conexas C_x e C_y ou coincidem ou são disjuntas, pois se $z \in C_x \cap C_y$, então $C_x \subset C_y$ e $C_y \subset C_x$. Assim, a relação "x e y pertencem a mesma componente conexa em X" é uma equivalência no conjunto X. As classes de equivalência são as componentes conexas dos pontos de X.

Proposição 2.80. *Toda componente conexa C_x é um conjunto fechado em X.*

Demonstração: Com efeito, sendo $C_x \subset \overline{C_x} \cap X \subset \overline{C_x}$, pelo item *iv*) do Teorema 2.54, segue que $\overline{C_x} \cap X$ é um subconjunto conexo de X, contendo C_x. Assim, $\overline{C_x} \cap X = C_x$. Portanto C_x é fechado em X. ∎

Capítulo 3

Funções Diferenciáveis

Neste capítulo, abordaremos o conceito de derivada, que detém uma notável relevância no ramo da Análise Matemática. Veremos as definições de derivada, diferenciabilidade, função de classe C^1 e como elas estão relacionadas, destacando as principais diferenças e similaridades com o conceito de derivada de uma função real de variável real, bem como algumas de suas propriedades para podermos apresentar e trabalhar com o Teorema da Função Inversa.

3.1 Derivadas Parciais

A derivada nos permite deduzir o comportamento de uma aplicação em uma determinada vizinhança de um ponto. Iniciamos nossa apresentação introduzindo um conceito preliminar de derivada nos espaços Euclidianos.

Definição 3.1. Seja $f : U \to \mathbb{R}$ uma função definida no aberto $U \subset \mathbb{R}^n$. Para cada $i = 1, ..., n$, a i-ésima derivada parcial de f no ponto $a = (a_1, ..., a_n) \in U$ é o número;

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t} = \lim_{t \to 0} \frac{f(a_1, ..., a_i + t, ..., a_n) - f(a)}{t},$$

caso este limite exista. Denotamos por $\frac{\partial f}{\partial x_i}(a)$, a derivada de f em relação a sua i-ésima variável, seja qual for a notação que se atribua a ela.

Observação 3.2. i. A definição anterior diz que $\frac{\partial f}{\partial x_i}(a) = (f \circ \lambda)'(0)$ é a derivada no ponto $t = 0$, da função real $f \circ \lambda : (-\delta, \delta) \to \mathbb{R}$, onde $\lambda(t) = a + te_i$.

ii. Quando $n = 3$, escrevemos (x, y, z) e $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$ são as derivadas parciais de f em relação a primeira, segunda e terceira variáveis, respectivamente.

Exemplo 3.3. Seja $f : \mathbb{R}^2 \to \mathbb{R}$ defina por $f(x, y) = \frac{xy}{x^2+y^2}$ se $x^2 + y^2 \neq 0$ e $f(0, 0) = 0$.

Como $f(0,y)=0$ para todo y e $f(x,0)=0$ para todo x, segue que

$$\frac{\partial f}{\partial x}(0,0)=\lim_{t\to 0}\frac{(f(0,0)+te_1)-f(0,0)}{t}=\lim_{t\to 0}\frac{f(t,0)-f(0,0)}{t}=\frac{0-0}{t}=0$$

e

$$\frac{\partial f}{\partial y}(0,0)=\lim_{t\to 0}\frac{(f(0,0)+te_2)-f(0,0)}{t}=\lim_{t\to 0}\frac{f(0,t)-f(0,0)}{t}=\frac{0-0}{t}=0.$$

Contudo, a função f é descontínua na origem $(0,0)$, pois, defina $v_k=\frac{1}{k}(\sqrt{2},\sqrt{2})$ para cada $k\in\mathbb{N}$. Assim, $(v_k)\subset\mathbb{R}^2$ e v_k converge para $(0,0)$. Porém;

$$\lim_{k\to\infty}f(v_k)=\lim_{k\to\infty}f\left(\frac{\sqrt{2}}{k},\frac{\sqrt{2}}{k}\right)=\lim_{k\to\infty}\frac{\frac{\sqrt{2}}{k}\cdot\frac{\sqrt{2}}{k}}{\left(\frac{\sqrt{2}}{k}\right)^2+\left(\frac{\sqrt{2}}{k}\right)^2}=\lim_{k\to\infty}\frac{\frac{2}{k^2}}{\frac{4}{k^2}}=\lim_{k\to\infty}\frac{1}{2}=\frac{1}{2}.$$

Portanto, como $\frac{1}{2}\neq f(0,0)$, segue que f não é contínua em $(0,0)$.

Perceba que o exemplo anterior aponta que a existência das n derivadas parciais no ponto a não assegura a continuidade da função f neste ponto.

Para cada $i=1,...,n$, a função $\lambda(t)=f(a+te_i)$ é essencialmente a restrição de f ao segmento $(a-\delta e_i,a+\delta e_i)$ da reta que passa pelo ponto a e é paralela ao i-ésimo eixo coordenado de $\mathbb{R}^n$

A derivada parcial $\frac{\partial f}{\partial x_i}(a)=(f\circ\lambda)'(0)$ dá informação apenas sobre o comportamento de f ao longo desse segmento. Em particular, a existência das n derivadas parciais de f no ponto a implica que a restrição de f a cada um dos n segmentos paralelos aos eixos, que se cortam no ponto a, é contínua, embora não garanta a continuidade de $f:U\to\mathbb{R}$ em a.

Observação 3.4. Assim como para funções de uma variável, se $\frac{\partial f(a)}{\partial x_i}$ existe e é positiva em todos os pontos do segmento de reta $[a-\delta e_i,a+\delta e_i]$, paralelo ao i-ésimo eixo coordenado, então f é crescente ao longo desse segmento: $s<t\Rightarrow f(a+se_i)<f(a+te_i)$, desde que $|s|\leq\delta$ e $|t|\leq\delta$.

Embora a noção de derivada parcial também faça sentido para aplicações $f:U\to\mathbb{R}^n$, com $U\subset\mathbb{R}^m$ aberto. Se $a\in U$, põe-se para cada $i=1,...,m$: $\frac{\partial f}{\partial x_i}(a)=\lim\limits_{t\to 0}\dfrac{f(a+te_i)-f(a)}{t}$. Claramente, $\frac{\partial f}{\partial x_i}$ é um vetor de $\mathbb{R}^n$. Se $f=(f_1,...,f_n)$ então $\frac{\partial f}{\partial x_i}(a)=\left(\frac{\partial f_1}{\partial x_i}(a),...,\frac{\partial f_n}{\partial x_i}(a)\right)$. Daremos prioridade às funções com valores reais, pois, para elas, tem sentido o vetor gradiente o qual contribui para entendermos como f cresce ou decresce de modo mais direto.

3.2 Funções de Classe C^1

Definição 3.5. Seja $f : U \to \mathbb{R}$ uma função que possui as n derivadas parciais em todos os pontos do aberto $U \subset \mathbb{R}^n$. Então ficam definidas n funções;

$$\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n} : U \to \mathbb{R}, \text{ onde } \frac{\partial f}{\partial x_i} : x \mapsto \frac{\partial f}{\partial x_i}(x).$$

Se estas funções forem contínuas em U, diremos que f é uma função de classe C^1 e escrevemos $f \in C^1$.

Definição 3.6. Uma aplicação $f : U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^n$ é dita de classe C^1 quando cada uma de suas funções-coordenadas $f_1, ..., f_n : U \to \mathbb{R}$ é de classe C^1.

Muitas propriedades importantes das funções de classe C^1 resultam de serem diferenciáveis no sentido da definição a seguir.

Definição 3.7. Uma função $f : U \to \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$ é diferenciável no ponto $a \in U$ quando cumpre as seguintes condições

i. Existem as derivadas parciais $\frac{\partial f}{\partial x_1}(a), ..., \frac{\partial f}{\partial x_n}(a)$;

ii. Para todo $v = (v_1, ..., v_n)$ tal que $a + v \in U$, tem-se

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

Observação 3.8.

- Na definição anterior e a seguir, escrevemos $\frac{\partial f}{\partial x_i}$ em vez de $\frac{\partial f}{\partial x_i}(a)$ por simplicidade.

- A essência da definição da diferenciabilidade está na condição $\lim\limits_{v \to 0} \frac{r(v)}{\|v\|} = 0$, pois a igualdade que define o "resto" $r(v)$ pode ser escrita para qualquer função que possua as n derivadas parciais com $\lim\limits_{v \to 0} \frac{r(v)}{\|v\|} = 0$, temos que $\lim\limits_{v \to 0} r(v) = 0$ pois $r(v) = \left(\frac{r(v)}{\|v\|}\right) \cdot \|v\|$. Assim, segue que $\lim\limits_{v \to 0}(f(a + v) - f(a)) = 0$.

Proposição 3.9. *Se $f : U \to \mathbb{R}$ é diferenciável em $a \in U$, então f é contínua em a.*

Demonstração: Basta mostrar que $\lim\limits_{\|v\| \to 0} f(a + v) = f(a)$. Como f é diferenciável em a, existem $\frac{\partial f}{\partial x_i}(a)$, para todo $i \in \{1, ..., n\}$ e existe $r(v)$ tal que

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

Daí, observe que $f(a+v) = f(a)+\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\cdot v_i + \frac{r(v)}{\|v\|}\cdot\|v\|$, assim, temos que $\lim_{\|v\|\to 0} f(a+v) = \lim_{\|v\|\to 0}\left(f(a)+\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)\cdot v_i+\frac{r(v)}{\|v\|}\cdot\|v\|\right) = f(a).$ ■

Observação 3.10. Note que a Definição 3.7, de diferenciabilidade, generaliza o conceito de diferenciabilidade visto em Análise Real, pois se $f : I \subset \mathbb{R} \to \mathbb{R}$ é diferenciável em $a \in I$, temos $r(v) = f(a+v) - f(a) - f'(a)\cdot v$, para todo $v \in \mathbb{R}$ tal que $(a+v) \in I$, satisfaz

$$\begin{aligned}\frac{r(v)}{|v|} &= \frac{f(a+v)-f(a)}{|v|} - \frac{f'(a)\cdot v}{|v|}\\ &= \frac{v}{|v|}\cdot\left(\frac{f(a+v)-f(a)}{v} - f'(a)\right).\end{aligned}$$

Assim, $\lim_{|v|\to 0}\frac{r(v)}{|v|} = \pm(f'(a)-f'(a)) = 0.$

Teorema 3.11. *Toda função $f : U \to \mathbb{R}$ de classe C^1 é diferenciável.*

Demonstração: Por simplicidade vamos supor $U \subset \mathbb{R}^2$. O caso geral, a menos de uma notação mais elaborada, é feito de maneira análoga. Fixemos $a = (a_1, a_2) \in U$ e tomamos $v = (v_1, v_2)$ tal que $a + v = (a_1 + v_1, a_2 + v_2) \in B$. Seja

$$r(v) = f(a_1+v_1, a_2+v_2) - f(a_1,a_2) - \sum_{i=1}^{2}\frac{\partial f}{\partial x_i}(a)\cdot v_i$$

$$= f(a_1+v_1, a_2+v_2) - f(a_1, a_2+v_2) + f(a_1, a_2+v_2) - f(a_1,a_2) - \frac{\partial f}{\partial x_1}(a)\cdot v_1 - \frac{\partial f}{\partial x_2}(a)\cdot v_2. \quad (3.1)$$

Agora defina,

$$\begin{aligned} g : [0, v_1] &\longrightarrow \mathbb{R} \\ t &\longmapsto g(t) = f(a_1+t, a_2+v_2)\end{aligned} \quad e \quad \begin{aligned} h : [0, v_2] &\longrightarrow \mathbb{R} \\ t &\longmapsto h(t) = f(a_1, a_2+t)\end{aligned}$$

que são funções diferenciáveis e pelo Teorema do Valor Médio na reta, temos que existe $\theta_1 \in (0, v_1)$ e $\theta_2 \in (0, v_2)$, tais que $g(v_1) - g(0) = g'(\theta_1)\cdot(v_1 - 0)$, isso implica em

$$f(a_1+v_1, a_2+v_2) - f(a_1, a_2+v_2) = \frac{\partial f}{\partial x_1}(a_1+\theta_1, a_2+v_2)\cdot v_1. \quad (3.2)$$

E $h(v_2) - h(0) = h'(\theta_2)\cdot(v_2 - 0)$, o que implica em

$$f(a_1, a_2+v_2) - f(a_1, a_2) = \frac{\partial f}{\partial x_2}(a_1, a_2+\theta_2)\cdot v_2. \quad (3.3)$$

Utilizando $(3.1), (3.2)$ $e(3.3)$, obtemos

$$\begin{aligned} r(v) &= \frac{\partial f}{\partial x_1}(a_1+\theta_1, a_2+v_2)v_1 + \frac{\partial f}{\partial x_2}(a_1, a_2+\theta_2)v_2 - \frac{\partial f}{\partial x_1}(a_1,a_2)v_1 - \frac{\partial f}{\partial x_2}(a_1,a_2)v_2 \\ &= v_1\left[\frac{\partial f}{\partial x_1}(a_1+\theta_1, a_2+v_2) - \frac{\partial f}{\partial x_1}(a_1,a_2)\right] + v_2\left[\frac{\partial f}{\partial x_2}(a_1, a_2+\theta_2) - \frac{\partial f}{\partial x_2}(a_1,a_2)\right]. \end{aligned}$$

Com isto, observe que

$$\begin{aligned} \lim_{\|v\|\to 0}\left|\frac{r(v)}{\|v\|}\right| \leq \lim_{\|v\|\to 0} &\left[\left|\frac{v_1}{\|v\|}\right|\left|\frac{\partial f}{\partial x_1}(a_1+\theta_1, a_2+v_2) - \frac{\partial f}{\partial x_1}(a_1,a_2)\right|\right] \\ &+ \lim_{\|v\|\to 0}\left[\left|\frac{v_2}{\|v\|}\right|\left|\frac{\partial f}{\partial x_2}(a_1, a_2+\theta_2) - \frac{\partial f}{\partial x_2}(a_1,a_2)\right|\right]. \end{aligned}$$

Como $\frac{\partial f}{\partial x_i}$ são contínuas para $i = 1, 2$ e por $\left|\frac{v_1}{\|v\|}\right| < 1$, para $i = 1, 2$, segue que $\displaystyle\lim_{\|v\|\to 0}\left|\frac{r(v)}{\|v\|}\right| = 0$, assim $\displaystyle\lim_{\|v\|\to 0}\frac{r(v)}{\|v\|} = 0$. Logo, f é diferenciável. ■

Como vimos na Proposição 3.9, toda função diferenciável é contínua, e no Teorema 3.11, toda função de classe C^1 é diferenciável, segue por transitividade que toda função de classe C^1 é contínua.

Teorema 3.12. *Sejam $U \subseteq \mathbb{R}^m$ e $V \subseteq \mathbb{R}^n$ abertos, $g : V \to \mathbb{R}$ uma função diferenciável em $b \in V$ e $f : U \to V$ uma função cuja $\frac{\partial f_j}{\partial x_i}(a)$ existem para todo $j = 1, ..., n$ e $i = 1, ..., m$. Então, se $b = f(a)$, temos que $h = g \circ f$ possui $\frac{\partial h}{\partial x_i}(a)$ para todo $i = 1, ..., m$ e além disso;*

$$\frac{\partial h}{\partial x_i}(a) = \sum_{j=1}^{n}\frac{\partial g}{\partial y_j}(b)\cdot\frac{\partial f_j}{\partial x_i}(a), \forall i = 1, ..., m.$$

Demonstração: Seja $v \in \mathbb{R}^m$ tal que $(a + v) \in U$, assim;

$$\begin{aligned} h(a+v) - h(a) &= g(f(a+v)) - g(f(a)) \\ &= g(f(a) + f(a+v) - f(a)) - g(f(a)) \\ &= g(b+w) - g(b). \end{aligned} \tag{3.4}$$

onde $w = f(a + v) - f(a)$. Por g ser diferenciável, existe $\rho(w)$ tal que $\displaystyle\lim_{\|w\|\to 0}\rho(w) = 0$ e

$$g(b+w) - g(b) = \sum_{j=1}^{n}\frac{\partial g}{\partial y_j}(b)w_j + \rho(w)\cdot\|w\|. \tag{3.5}$$

Por (3.4) e (3.5), segue que

$$h(a+v)-h(a)=\sum_{j=1}^{n}\frac{\partial g}{\partial y_j}(b)\cdot(f_j(a+v)-f_j(a))+\rho(w)\cdot\|f(a+v)-f(a)\|.$$

Daí, sendo $v=te_i$ e fazendo $t\to 0$ concluímos

$$\begin{aligned}\frac{\partial h}{\partial x_i}(a)&=\lim_{t\to 0}\left[\sum_{j=1}^{n}\frac{\partial g}{\partial y_j}(b)\left(\frac{f_j(a+te_i)-f_j(a)}{t}\right)+\rho(w)\left\|\frac{f(a+te_i)-f(a)}{t}\right\|\right]\\&=\sum_{j=1}^{n}\frac{\partial g}{\partial y_j}(b)\cdot\frac{\partial f_j}{\partial x_i}(a).\end{aligned}$$

■

Definição 3.13. O gradiente de uma função diferenciável $f:U\to\mathbb{R}$ no ponto $a\in U$ aberto, é o vetor $\nabla f(a)$ =grad $f(a)=\left(\frac{\partial f}{\partial x_1}(a),...,\frac{\partial f}{\partial x_n}(a)\right)$.

Definição 3.14. Seja v um vetor de $\mathbb{R}^n$, a derivada direcional de f no ponto a, na direção de v é $\frac{\partial f}{\partial v}(a)=\lim\limits_{t\to 0}\dfrac{f(a+tv)-f(a)}{t}$, caso este limite exista.

Estas definições permitem enunciar os corolários a seguir da Regra da Cadeia. O primeiro deles mostra que, quando f é diferenciável no ponto a, a derivada direcional $\frac{\partial f}{\partial v}(a)$ existe em relação a qualquer vetor v, além disso, é possível obter uma expressão para essa derivada em termos das derivadas parciais de f e das coordenadas de v e, finalmente, mostra que, na definição de $\frac{\partial f}{\partial v}(a)$, em vez do caminho retilíneo $t\mapsto a+tv$, pode-se usar qualquer caminho $\lambda:(-\delta,\delta)\to U$ desde que se tenha $\lambda(0)=a$ e $\lambda'(0)=v$.

Corolário 3.15. *Seja $f:U\to\mathbb{R}$ diferenciável no aberto $U\subset\mathbb{R}^n$, com $a\in U$. Dado o vetor $v=(v_1,...,v_n)$, se $\lambda:(-\delta,\delta)\to U$ é qualquer caminho diferenciável tal que $\lambda(0)=a$ e $\lambda'(0)=v$, tem-se;*

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

Demonstração: Perceba que para $\lambda(t)=(\lambda_1(t),...,\lambda_n(t))$, temos $v_i=\lambda_i'(0)$. Além disso, $\frac{\partial f}{\partial v}(a)=(f\circ\lambda)'(0)$ com $\lambda(t)=a+tv$, já que $\lambda'(0)=v$. Portanto, segue que,

$$\begin{aligned}(f\circ\lambda)'(0)&=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)\cdot\frac{d\lambda_i}{dt}(0)\\&=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)\cdot v_i.\end{aligned}$$

■

Teorema 3.16. *(Teorema do Valor Médio) Dada $f : U \to \mathbb{R}$ diferenciável no aberto $U \subset \mathbb{R}^n$, se o segmento de reta $[a, a+v]$ estiver contido em U, então existe $\theta \in (0,1)$ tal que,*

$$\begin{aligned} f(a+v) - f(a) &= \frac{\partial f}{\partial v}(a+\theta v) \\ &= \langle \nabla f(a+\theta v), v \rangle \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a+\theta v) \cdot v_i, \ \textit{com } v = (v_1, ..., v_n). \end{aligned}$$

Demonstração: Com efeito, considerando o caminho retilíneo $\lambda : [0,1] \to U$, dado por $\lambda(t) = a + tv$, vemos que $f(a+v) - f(a) = (f \circ \lambda)(1) - (f \circ \lambda)(0)$. Assim, pelo Teorema do Valor Médio para funções de uma variável real, existe $\theta \in (0,1)$ tal que $(f \circ \lambda)(1) - (f \circ \lambda)(0) = (f \circ \lambda)'(\theta)$. Pela Regra da Cadeia,

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

■

Corolário 3.17. *(Desigualdade do Valor Médio) Seja $f : U \to \mathbb{R}$ diferenciável no aberto $U \subset \mathbb{R}^n$. Se o segmento de reta $[a, a+v]$ estiver contido em U e existir $M > 0$ tal que $\|\nabla f(a+tv)\| \leq M$ para todo $t \in [0,1]$, então $\|f(a+v) - f(a)\| \leq M \cdot \|v\|$.*

Demonstração: Pela desigualdade de Schwarz;

$$\begin{aligned} \|f(a+v) - f(a)\| &= \|\langle \nabla f(a+\theta v), v \rangle\| \\ &\leq \|\nabla f(a+\theta v)\| \cdot \|v\| \\ &\leq M \cdot \|v\|. \end{aligned}$$

■

Observação 3.18. Em particular, se U é convexo, f é diferenciável e $\|\nabla f(x)\| \leq M$ para todo $x \in U$, então $\|f(y) - f(x)\| \leq M\|x - y\|$ quaisquer que sejam $x, y \in U$.

Corolário 3.19. *Seja $f : U \to \mathbb{R}$ diferenciável no aberto $U \subset \mathbb{R}^n$. Se U é conexo e $\nabla f(x) = 0$ para todo $x \in U$, então f é constante.*

Demonstração: Sejam $y, z \in U$, queremos mostrar que $f(y) = f(z)$. Perceba que existe $\lambda : [0,1] \to U$ diferenciável, tal que $\lambda(o) = y$ e $\lambda(1) = z$. Pelo Teorema do Valor Médio, existe $\theta \in (0,1)$, tal que, considerando $v = y - z$ temos

$$f(y) - f(z) = f(z + y - z) - f(z) = \langle \nabla f(z+\theta v), v \rangle = \langle 0, v \rangle = 0.$$

Assim, $f(y) = f(z)$. ■

Definição 3.20. Seja $f : U \to \mathbb{R}$ de classe C^1, o conjunto $f^{-1}(c) = \{x \in U; f(x) = c\}$ é, para todo $c \in \mathbb{R}$, chamado o conjunto de nível c da função f. Quando $U \subset \mathbb{R}^2$, o conjunto é chamado de curva ou linha de nível c de f, que é definida por $f(x, y) = c$. Semelhantemente, quando $U \subset \mathbb{R}^3$, o conjunto $f^{-1}(c)$, definido pela equação $f(x, y, z) = c$, costuma ser chamado de superfície de nível c da função f.

É importante notar que para certas funções escolhidas de forma especial, tais conjuntos podem auxiliar na descrição de comportamento do gradiente ∇f.

Teorema 3.21. *Seja $f : U \to \mathbb{R}$ diferenciável.*

i) Se $\nabla f(a) \neq 0$, então f é crescente na direção de $\nabla f(a)$.

ii) Dentre todas as direções em que f é crescente, a direção do gradiente é a que f "cresce" mais rápido.

iii) Se $c = f(a)$, então $\nabla f(a)$ é ortogonal à v, para todo $v \in f^{-1}(c)$.

Demonstração:

i) Se $w = \nabla f(a)$, então $\frac{\partial f}{\partial w}(a) = \langle \nabla f(a), w \rangle = \|\nabla f(a)\|^2 > 0$.

ii) Seja qualquer $v \in \mathbb{R}^n$, de modo que $\|v\| = \|\nabla f(a)\|$, então;

$$\begin{aligned}
\frac{\partial f}{\partial v}(a) &= \langle \nabla f(a), v \rangle \\
&\leq |\langle \nabla f(a), v \rangle| \\
&\leq \|\nabla f(a)\| \cdot \|v\| \\
&= \|\nabla f(a)\|^2 \\
&= \frac{\partial f}{\partial w}(a), \text{ com } w = \nabla f(a).
\end{aligned}$$

iii) Seja $\lambda : (-\varepsilon, \varepsilon) \to f^{-1}(c)$ um caminho diferenciável e considere $h := f \circ \lambda$, então $h(t) = f(\lambda(t)) = c$, para todo $t \in (\varepsilon, \varepsilon)$, o que implica em $(f \circ \lambda)' = 0$. Portanto $\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\lambda(t)) \cdot \lambda'(t) = 0$. Em particular, considere $\lambda : (-\varepsilon, \varepsilon) \to f^{-1}(c)$ tal que $\lambda(0) = a$ e $\lambda'(0) = v$, então $\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot v_i$ o que implica em $\langle \nabla f(a), v \rangle = 0$.

■

Exemplo 3.22. Sejam $f, g, h : \mathbb{R}^2 \to \mathbb{R}$ definidas por $f(x, y) = ax + by$ com $a^2 + b^2 \neq 0$, $g(x, y) = x^2 + y^2$ e $h(x, y) = x^2 - y^2$.

a) A linha de nível c de f é a reta definida por $ax + by = c$, o vetor $\nabla f(x, y) = (a, b)$ é constante em qualquer ponto $(x, y) \in \mathbb{R}^2$. Logo, as linhas de nível de f são retas paralelas entre si e todas perpendiculares ao vetor $v = (a, b)$.

b) O conjunto de nível c da função $g(x, y) = x^2 + y^2$ é vazio se $c < 0$ e quando $c = 0$, reduz-se ao ponto $0 \in \mathbb{R}^2$. Para $c > 0$, a linha de nível c é a circunferêcia de equação $x^2 + y^2 = c$, cujo centro é a origem e raio $\sqrt{c}$. Note que o $\nabla g(x, y) = (2x, 2y)$, um vetor colinear com o raio.

c) A linha de nível 0 da função $h(x, y) = x^2 - y^2$ é o par de retas perpendiculares definidas pela equação $x^2 - y^2 = 0$, o que equivale a "$x + y = 0$ ou $x - y = 0$". Se $c > 0$, $x^2 - y^2 = c$ define uma hipérbole cujo eixo é o das abscissas. Se $c < 0$, a hipérbole tem como eixo, o eixo das ordenadas. Perceba que $\nabla h(x, y) = (2x, -2y)$. Ao atribuirmos valores a x e y, vemos que este vetor é perpendicular à curva de nível que passa em (x, y) e aponta na direção de crescimento de h.

Definição 3.23. Sejam $f : U \to \mathbb{R}$ diferenciável e $a \in U$, quando $\nabla f(a) = 0$, dizemos que a é um ponto crítico de f.

Definição 3.24. A imagem $f(a) \in \mathbb{R}^m$ de um ponto crítico é chamado um valor crítico de f.

Com isso, notamos que no exemplo anterior f não possui ponto crítico, enquanto que as funções g e h têm a origem como único ponto crítico.

Observação 3.25. Quando há ponto crítico, ocorre uma quebra de regularidade na disposição das curvas de nível quando se atinge um nível em que está o ponto crítico.

3.3 O Teorema de Schwarz

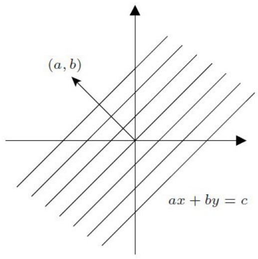
Definição 3.26. Seja $f : U \to \mathbb{R}$ uma função que possui as derivadas parciais $\frac{\partial f}{\partial x_1}(x), ...,$ $\frac{\partial f}{\partial x_n}(x)$ em todo ponto x do aberto $U \subset \mathbb{R}^n$. A j-ésima derivada parcial da função $\frac{\partial f}{\partial x_i} : U \to \mathbb{R}$ no ponto $x \in U$ é indicada por

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

Quando essas derivadas parciais de segunda ordem existirem em cada ponto $x \in U$, teremos n^2 funções $\frac{\partial^2 f}{\partial x_j \partial x_i} : U \to \mathbb{R}$. Quando forem contínuas, diremos que f é de classe C^2, ou seja, $f \in C^2$.

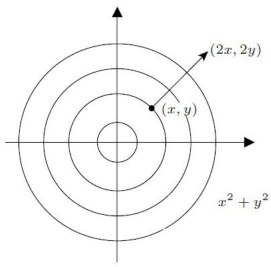
Observação 3.27. Em geral, apenas a existência das derivadas parciais de segunda ordem em todos os pontos em que f está definida, não assegura que $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}$.

Figura 3.1: Figura do exemplo 3.22 a).



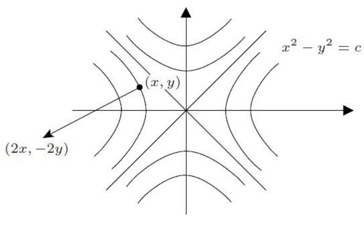
Fonte: LIMA, 2016, p. 64

Figura 3.2: Figura do exemplo 3.22 b).



Fonte: LIMA, 2016, p. 64

Figura 3.3: Figura do exemplo 3.22 c).



Fonte: LIMA, 2016, p. 64

Exemplo 3.28. Seja $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = \frac{xy(x^2-y^2)}{x^2+y^2}$ quando $x^2+y^2 \neq 0$ e $f(0,0) = 0$. Note que se $y \neq 0$, $f(0,y) = 0$ e

$$\begin{aligned}
\frac{\partial f}{\partial x}(0,y) &= \lim_{t\to 0} \frac{f((0,y)+te_i) - f(0,y)}{t} \\
&= \lim_{t\to 0} \frac{f(t,y) - f(0,y)}{t} \\
&= \lim_{t\to 0} \frac{1}{t} \cdot \left[\frac{ty(t^2-y^2)}{t^2+y^2}\right] \\
&= \lim_{t\to 0} \frac{yt^2 - y^3}{t^2+y^2} \\
&= \frac{-y^3}{y^2} \\
&= -y.
\end{aligned}$$

Com isto, observamos que

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0) = \lim_{t\to 0} \frac{\frac{\partial f}{\partial x}(0,h) - \frac{\partial f}{\partial x}(0,0)}{h} = \lim_{t\to 0} \frac{-h}{h} = -1. \tag{3.6}$$

Pois, $\frac{\partial f}{\partial x}(0,0) = \lim\limits_{t\to 0} \dfrac{f(t,0) - f(0,0)}{t} = \lim\limits_{t\to 0} \dfrac{0}{t} = 0$. Por outro lado, se $x \neq 0$, $f(x,0) = 0$ e $\frac{\partial f}{\partial y}(x,0) = \lim\limits_{h\to 0} \dfrac{f(x,h) - f(x,0)}{h} = \lim\limits_{h\to 0} \dfrac{1}{h} \cdot \left[\dfrac{xh(x^2-h^2)}{x^2+h^2}\right] = \dfrac{x^3}{x^2} = x$. Com isso, temos;

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) = \lim_{t\to 0} \frac{\frac{\partial f}{\partial y}(t,0) - \frac{\partial f}{\partial y}(0,0)}{t} = 1. \tag{3.7}$$

Das igualdades (3.6) e (3.7), concluímos que $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$.

Definição 3.29. Seja $f : U \to \mathbb{R}$, se existem as derivadas parciais de segunda ordem em todo ponto $x \in U$, os números $h_{ij}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ formam uma matriz $h(x) = [h_{ij}]$, chamada a matriz hessiana da função f.

O Teorema de Schwarz que veremos mais adiante, afirma que se f é de classe C^2, então a matriz hessiana de f é simétrica.
A demonstração do Teorema de Schwarz se baseia em um resultado, atribuído a Leibniz, segundo o qual é permitido derivar sob o sinal de integral, desde que o resultado da derivação seja uma função contínua. Por sua vez, a demonstração do Teorema de Leibniz utiliza o seguinte Lema:

Lema 3.30. *Sejam $X \subset \mathbb{R}^m$ um conjunto arbitrário e $K \subset \mathbb{R}^n$ compacto. Fixemos $x_0 \in X$. Se $f : X \times K \to \mathbb{R}^p$ é contínua, então para todo $\varepsilon > 0$ dado, pode-se obter $\delta > 0$ tal que $x \in X$ e $\|x - x_0\| < \delta$ implicam em $\|f(x,t) - f(x_0,t)\| < \varepsilon$, seja qual for $t \in K$.*

Demonstração: Supondo por contradição, deve existir $\varepsilon > 0$ e sequências de pontos $x_k \in X$ e $t_k \in K$ tais que $\|x_k - x_0\| < \frac{1}{k}$ e $\|f(x_k, t_k) - f(x_0, t_k)\| \geq \varepsilon$. Passando a uma subsequência, se necessário, podemos admitir que $\lim t_k = t_0 \in K$. Como $\lim\limits_{k\to\infty} x_k = x_0$, a continuidade de f nos garante

$$\begin{aligned}\varepsilon &\leq \lim_{k\to\infty} \|f(x_k, t_k) - f(x_0, t_k)\| \\ &= \|f(x_0, t_0) - f(x_0, t_0)\| \\ &= 0.\end{aligned}$$

O que é um absurdo. ■

Teorema 3.31. *(Derivação sob o sinal de integral) Sejam $U \subseteq \mathbb{R}^n$ aberto e $f : U \times [a, b] \to \mathbb{R}$ uma função tal que $\frac{\partial f}{\partial x_i} : U \times [a, b] \to \mathbb{R}$ existem e são contínuas para todo $i = 1, ..., n$ e f é contínua. Daí, defina a função $\varphi : U \to \mathbb{R}$, dada por $\varphi(x) = \displaystyle\int_a^b f(x, t)dt$. Então, $\frac{\partial \varphi}{\partial x_i} : U \to \mathbb{R}$ existe e vale que $\displaystyle\frac{\partial \varphi}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(x, t)dt$.*

Demonstração: Sejam $x_0 \in U$ fixado arbitrariamente e $s \in \mathbb{R}\backslash\{0\}$ tal que $[x_0, x_0 + se_i] \subset U$. Daí, note que

$$\begin{aligned}&\frac{\varphi(x_0 + se_i) - \varphi(x_0)}{s} - \int_a^b \frac{\partial f}{\partial x_i}(x_0, t)dt = \\ &\qquad \frac{1}{s}\left[\int_a^b f(x_0 + se_i, t) - f(x_0, t)dt\right] - \int_a^b \frac{\partial f}{\partial x_i}(x_0, t)dt.\end{aligned}$$

Pelo Teorema do Valor Médio, existe $\theta \in (0, s)$ tal que

$$\begin{aligned}\frac{\varphi(x_0 + se_i) - \varphi(x_0)}{s} - \int_a^b \frac{\partial f}{\partial x_i}(x_0, t)dt &= \frac{1}{s}\int_a^b \left[\frac{\partial f}{\partial x_i}(x_0 + \theta e_i, t).s\right] dt - \int_a^b \frac{\partial f}{\partial x_i}(x_0, t)dt \\ &= \int_a^b \left[\frac{\partial f}{\partial x_i}(x_0 + \theta e_i, t) - \frac{\partial f}{\partial x_i}(x_0, t)\right] dt. \qquad (3.8)\end{aligned}$$

Pelo Lema 3.30, dado $\varepsilon > 0$ é possível exibir $\delta > 0$ tal que se $\|\theta e_i\| < \delta$, então;

$$\left|\frac{\partial f}{\partial x_i}(x_0 + \theta e_i, t) - \frac{\partial f}{\partial x_i}(x_0, t)\right| < \frac{\varepsilon}{b - a}.$$

Retornando a igualdade (3.8) com $|s| < \delta$, temos que

$$\begin{aligned}\left|\frac{\varphi(x_0 + se_i) - \varphi(x_0)}{s} - \int_a^b \frac{\partial f}{\partial x_i}(x_0, t)dt\right| &\leq \int_a^b \left|\frac{\partial f}{\partial x_i}(x_0 + \theta e_i, t) - \frac{\partial f}{\partial x_i}(x_0, t)\right| dt \\ &< \int_a^b \frac{\varepsilon}{b - a}dt = \varepsilon\end{aligned}$$

Então, $\frac{\partial \varphi}{\partial x_i}(x_0) = \displaystyle\int_a^b \frac{\partial f}{\partial x_i}(x_0, t)dt.$ ■

Teorema 3.32. *(Teorema de Schwarz) Se $f : U \to \mathbb{R}$ é de classe C^2 no aberto $U \subset \mathbb{R}^n$, então para quaisquer $i, j = 1, ..., n$ e $x \in U$, tem-se $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$.*

Demonstração: Suponha, sem perda de generalidade, que $U = I \times J$ seja um retângulo em $\mathbb{R}^2$. Fixando $b \in J$, o Teorema Fundamental do Cálculo garante que para todo $(x, y) \in U$, tem-se $f(x, y) = f(x, b) + \displaystyle\int_b^y \frac{\partial f}{\partial y}(x, t)dt$. Como $\frac{\partial^2 f}{\partial x \partial y}$ é contínua, pelo Teorema 3.31, podemos derivar sob o sinal de integral. Assim, $\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial x}(x, b) + \displaystyle\int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t)dt$. Derivando em relação a y, temos;

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(x, y)\right) = \frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}(x, b) + \int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t)dt\right]$$

$$= \frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}(x, b)\right] + \frac{\partial}{\partial y}\left[\int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t)dt\right]. \tag{3.9}$$

Agora, tome $g(y) = \frac{\partial f}{\partial x}(x, b)$, então;

$$g'(y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(x, b)\right) = \lim_{t \to 0} \frac{g(y + te_i) - g(y)}{t} = \lim_{t \to 0} = \frac{0 - 0}{t} = 0.$$

Além disso, seja H tal que $H'(t) = \frac{\partial^2 f}{\partial x \partial y}(x, t)$, então;

$$\int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t)dt = \int_b^y H'(t)dt = H(y) - H(b).$$

Assim,

$$\frac{\partial}{\partial y}\left(\int_b^y \frac{\partial^2 f}{\partial x \partial y}(x, t)dt\right) = H'(y) - 0 \Rightarrow H'(y) = \frac{\partial^2 f}{\partial x \partial y}(x, y).$$

Portanto, retornando a equação (3.9), temos $\frac{\partial^2 f}{\partial y \partial x}(x, y) = 0 + \frac{\partial^2 f}{\partial x \partial y}(x, y)$, ou seja,

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y).$$

■

Observação 3.33. De maneira geral, para cada inteiro $k \geq 1$, podemos considerar as derivadas parciais de ordem k de uma função $f : U \to \mathbb{R}$, definida no aberto $U \subset \mathbb{R}^n$. Por exemplo, para $1 \leq i, j, k \leq n$, $\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(a)$ significa $\frac{\partial}{\partial x_i}\left(\frac{\partial^2}{\partial x_j \partial x_k}\right)(a)$.

Como toda permutação dos índices $i_1, ..., i_k$ pode ser obtida por meio de repetidas inversões de índices adjacentes, segue do Teorema de Schwarz 3.32 que a derivada de ordem k, $\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} ... \partial x_{i_k}}(a)$ não depende da ordem em que são feitas as derivações, desde que existam todas as derivadas de ordem k de f e sejam contínuas.

Definição 3.34. Seja $f : U \to \mathbb{R}$ uma função que possui em cada ponto de U, todas as derivadas parciais de ordem k, as quais são funções contínuas em U. Dizemos que f é uma função de classe C^k, ou seja, $f \in C^k$.

Observação 3.35. Quando $f \in C^k$ para todo $k = 1, 2, ...$, dizemos que f é uma função de classe C^∞.

Veremos uma aplicação bastante interessante do Teorema de Schwarz envolvendo a função harmônica. Para demonstrar esta aplicação, precisaremos definir o que é uma função harmônica, e usaremos o Teste da Segunda derivada como resultado auxiliar.

Definição 3.36. Sejam $B[0; 1] \subset \mathbb{R}^2$ e $u : B[0; 1] \to \mathbb{R}$ uma função contínua tal que $u|_{B(0;1)}$ é de classe C^2. Dizemos que u é harmônica se,

$$\frac{\partial^2 u}{\partial x^2}(x, y) + \frac{\partial^2 u}{\partial y^2}(x, y) = 0,$$

para todo $(x, y) \in B(0; 1)$.

A seguir vamos enunciar um resultado que será usado para demonstrar uma propriedade a respeito das funções harmônicas, tal resultado nada mais é do que o teste da segunda derivada, vamos omitir sua demonstração aqui, contudo o leitor interessado pode encontrá-la na referência [4], Cálculo, Vol.2, James Stewart.

Lema 3.37 (Teste da Segunda Derivada)**.** *Seja $(a, b) \in int.X$ um ponto crítico de $f : X \to \mathbb{R}$, $f \in C^2$ e $D = D(a, b) = \frac{\partial^2 f}{\partial x^2}(a, b) \cdot \frac{\partial^2 f}{\partial y^2}(a, b) - \left[\frac{\partial^2 f}{\partial x \partial y}(a, b)\right]^2$. Se $D < 0$, então $f(a, b)$ não é mínimo local nem máximo local.*

Teorema 3.38. *Se u é harmônica e $(x_0, y_0) \in B(0; 1)$ é tal que $u(x_0, y_0) \geq u(x, y)$ para todo $(x, y) \in B[0; 1]$, então as derivadas de segunda ordem de u aplicadas em (x_0, y_0), são todas nulas.*

Demonstração: Seja $(x_0, y_0) \in B(0; 1)$ tal que (x_0, y_0) é um ponto de máximo para u. Desta forma, pela Proposição 3.48 (x_0, y_0) é um ponto crítico, portanto $\nabla u(x_0, y_0) = (0, 0)$ e $detH_u(x_0, y_0) \geq 0$, onde

$$H_u(x, y) = \begin{bmatrix} \dfrac{\partial^2 u}{\partial x^2} & \dfrac{\partial^2 u}{\partial x \partial y} \\ \dfrac{\partial^2 u}{\partial y \partial x} & \dfrac{\partial^2 u}{\partial y^2} \end{bmatrix}.$$

Pelo Teorema de Schwarz e por u ser harmônica, temos;

$$\begin{aligned}
0 \leq detH_u(x_0,y_0) &= \begin{vmatrix} \dfrac{\partial^2 u}{\partial x^2}(x_0,y_0) & \dfrac{\partial^2 u}{\partial x \partial y}(x_0,y_0) \\ \dfrac{\partial^2 u}{\partial x \partial y}(x_0,y_0) & -\dfrac{\partial^2 u}{\partial x^2}(x_0,y_0) \end{vmatrix} \\
&= -\left(\frac{\partial^2 u}{\partial x^2}(x_0,y_0)\right)^2 - \left(\frac{\partial^2 u}{\partial x \partial y}(x_0,y_0)\right)^2 \\
&\leq 0.
\end{aligned}$$

Portanto, $\frac{\partial^2 u}{\partial x^2}(x_0,y_0) = 0$ e $\frac{\partial^2 u}{\partial x \partial y}(x_0,y_0) = 0$. ■

3.4 A Fórmula de Taylor

A Fórmula de Taylor que veremos adiante, em sua versão restrita aos termos de até segunda ordem, desempenha um papel fundamental para o estudo do comportamento de uma função de classe C^2 na proximidade de um ponto crítico, e ela é baseada no seguinte Lema,

Lema 3.39. *Seja $r : B \to \mathbb{R}$ de classe C^2 na bola aberta $B \subset \mathbb{R}^n$, de centro 0. Se $r(0) = \frac{\partial r}{\partial x_i}(0) = \frac{\partial^2 r}{\partial x_i \partial x_j}(0) = 0$ para quaisquer $i, j = 1, ..., n$, então $\displaystyle\lim_{v \to 0} \frac{r(v)}{\|v\|^2} = 0$.*

Demonstração: Sendo $r : B \to \mathbb{R}$ uma função de classe C^1, logo diferenciável, que se anula juntamente com todas as suas derivadas $\frac{\partial r}{\partial x_i}$, no ponto $v = 0$, segue da diferenciabilidade de r que $\displaystyle\lim_{v \to 0} \frac{r(v)}{\|v\|} = 0$. Assim, pelo Teorema do Valor Médio 3.16 , para cada $v = (v_1, ..., v_n) \in B$, existe θ com $0 < \theta < 1$, tal que $r(v) = \sum_{i=1}^n \frac{\partial r}{\partial x_i}(\theta v) \cdot v_i$. Logo

$$\frac{r(v)}{\|v\|^2} = \frac{\sum_{i=1}^n \frac{\partial r}{\partial x_i}(\theta v)}{\|v\|} \cdot \frac{v_i}{\|v\|},$$

que podemos reescrever como

$$\frac{r(v)}{\|v\|^2} = \frac{\sum_{i=1}^n \frac{\partial r}{\partial x_i}(\theta v)}{\|\theta v\|} \cdot \frac{\theta v_i}{\|v\|}. \tag{3.10}$$

Note ainda que $0 = \frac{\partial r}{\partial v}(0) = \displaystyle\lim_{\|v\| \to 0} \frac{r(v) - r(0)}{\|v\|} = \lim_{\|v\| \to 0} \frac{r(v)}{\|v\|}$.
Também $\left|\frac{\theta \cdot v_i}{\|v\|}\right| = |\theta| \cdot \frac{|v_i|}{\|v\|} \leq \frac{|v_i|}{\|v\|} \leq 1$. Portanto, existe $\gamma \in (0,1)$ tal que

$$\frac{\partial r}{\partial x_i}(\theta v) = \frac{\partial r}{\partial x_i}(\theta v) - \frac{\partial r}{\partial x_i}(0) = \sum_{j=1}^n \frac{\partial}{\partial x_j}\left(\frac{\partial r}{\partial x_i}\right)(\theta \cdot \gamma \cdot v) \cdot \theta \cdot v_j. \tag{3.11}$$

Assim, por (3.10) e (3.11), temos

$$\begin{aligned}
\lim_{v\to 0}\frac{r(v)}{\|v\|^2} &= \lim_{v\to 0}\frac{\sum_{i,j=1}^{n}\frac{\partial^2 r}{\partial x_j\partial x_i}(\theta\cdot\gamma\cdot v)\cdot\theta\cdot v_j}{\|\theta v\|}\cdot\frac{\theta v_i}{\|v\|}\\
&= \lim_{v\to 0}\sum_{i,j=1}^{n}\frac{\partial^2 r}{\partial x_j\partial x_i}(\theta\cdot\gamma\cdot v)\cdot\frac{\theta\cdot v_j}{\|\theta v\|}\cdot\frac{\theta\cdot v_i}{\|v\|}\\
&= 0.
\end{aligned}$$

■

Teorema 3.40. *(A Fórmula de Taylor) Seja $f : U \to \mathbb{R}$ de classe C^2 no aberto $U \subset \mathbb{R}^n$. Fixando $a \in U$, para todo $v = (v_1, ..., v_n) \in \mathbb{R}^n$ tal que $a + v \in U$, temos*

$$f(a+v) - f(a) = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}v_i + \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}\cdot v_i\cdot v_j + r(v),$$

onde as derivadas são calculadas no ponto a. Então $\lim_{v\to 0}\frac{r(v)}{\|v\|^2} = 0$.

Demonstração: Para concluirmos que $\lim_{v\to 0}\frac{r(v)}{\|v\|^2} = 0$, pelo Lema 3.39, basta-nos mostrar que

$$r(v) = f(a+v) - f(a) - \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}v_i - \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}\cdot v_i\cdot v_j,$$

no ponto $v = 0$, se anula juntamente com suas derivadas parciais de primeira e segunda ordem. Assim;

$$r(0) = f(a+0) - f(a) - \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)\cdot 0 - \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(a)\cdot 0\cdot 0 = 0.$$

Agora, note que $\frac{\partial r}{\partial x_i}(v) = \frac{\partial f}{\partial x_i}(a+v) - \frac{\partial f}{\partial x_i}(a) - \frac{1}{2}\cdot 2\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(a)\cdot v_j$, então $\frac{\partial r}{\partial x_i}(0) = 0$. Além disso, $\frac{\partial^2 r}{\partial x_j\partial x_i}(v) = \frac{\partial^2 f}{\partial x_j\partial x_i}(a+v) - \frac{\partial^2 f}{\partial x_i\partial x_j}(a)$, então $\frac{\partial^2 f}{\partial x_j\partial x_i}(0) = \frac{\partial^2 f}{\partial x_j\partial x_i}(a) - \frac{\partial^2 f}{\partial x_i\partial x_j}(a) = 0$.
■

3.5 Pontos Críticos

Definição 3.41. Uma forma quadrática $H : \mathbb{R}^n \to \mathbb{R}$ é uma função cujo valor no vetor $v = (v_1, ..., v_n)$ é $\sum_{i,j=1}^{n} h_{ij}v_iv_j$, onde $[h_{ij}]$ é uma matriz simétrica $n \times n$.

O valor da forma quadrática H no vetor v é indicado pela notação $H \cdot v^2$. Assim, $H \cdot v^2 = \sum_{i,j=1}^{n} h_{ij}v_iv_j$, quando $v \in \mathbb{R}^n$.

Observação 3.42. Se $t \in \mathbb{R}$, então $H(t \cdot v)^2 = t^2(H \cdot v^2)$.

Uma forma quadrática pode ser;

- Positiva, se $H \cdot v^2 > 0$ para todo $v \neq 0$ em $\mathbb{R}^n$;
- Não-Negativa, se $H \cdot v^2 \geq 0$ para todo $v \in \mathbb{R}^n$;
- Negativa, se $H \cdot v^2 < 0$ para todo $v \neq 0$ em $\mathbb{R}^n$;
- Não-Positiva, se $H \cdot v^2 \leq 0$ para todo $v \in \mathbb{R}^n$;
- Indefinida, quando existem $v, w \in \mathbb{R}^n$ tais que $H \cdot v^2 > 0$ e $H \cdot w^2 < 0$.

Exemplo 3.43. A forma quadrática $H : \mathbb{R}^n \to \mathbb{R}$, onde $H \cdot v^2 = \langle v, v\rangle$, é positiva.

Pois seja $v = (v_1, ..., v_n) \in \mathbb{R}^n \backslash \{0\}$, então existe $j \in \{1, ..., n\}$ tal que $v_j \neq 0$. Assim, $H \cdot v^2 = v_1^2 + ... + v_n^2 \geq v_j^2 > 0$. Além disso, note que $\sum_{i,j=1}^n h_{ij} v_i v_j = H \cdot v^2 = v_1 v_1 + ... + v_n v_n = h_{11} \cdot v_1 v_1 + ... + h_{nn} \cdot v_n v_n$. Logo, se $i \neq j$, $h_{ij} v_i v_j = 0$, para todo $v \in \mathbb{R}^n$, portanto $h_{ij} = 0$. Como

$$Hv^2 = [v_1 \ v_2 \ ... \ v_n] \begin{bmatrix} h_{11} & h_{12} & ... & h_{1n} \\ h_{21} & h_{22} & ... & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & ... & h_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \text{ temos que}$$

$$Hv^2 = [v_1 \ v_2 \ ... \ v_n] \begin{bmatrix} h_{11} & 0 & ... & 0 \\ 0 & h_{22} & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & h_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = [h_{11} v_1^2 + h_{22} v_2^2 + ... + h_{nn} v_n^2].$$

Assim, $h_{11} v_1^2 + h_{22} v_2^2 + ... + h_{nn} v_n^2 = v_1^2 + ... + v_n^2$ se, e somente se, a matriz de H é a identidade.

Exemplo 3.44. Para todo $k \in \{1, 2, ..., n\}$, $H \cdot v^2 = v_1^2 + ... + v_k^2$ é uma forma quadrática não-negativa em $\mathbb{R}^n$.

Visto que ela é positiva, logo não-negativa.

Exemplo 3.45. Se $H \cdot v^2 = v_1^2 + ... + v_k^2 - v_{k+1}^2 - ... - v_n^2$ com $0 < k < n$, teremos uma forma quadrática indefinida.

Pois, sejam $v = (0, ..., 0, 1, ..., 1)$ e $w = (1, ..., 1, 0, ..., 0)$, note que $H \cdot v^2 = 0^2 + ... + 0^2 - 1^2 - ... - 1^2 < 0$ e $H \cdot w^2 = 1^2 + ... + 1^2 - 0^2 - ... - 0^2 > 0$.

Lembrando que um operador $H_u : \mathbb{R}^n \to \mathbb{R}^n$ é auto-adjunto quando $\langle H_u \cdot w, v\rangle = \langle w, H_u \cdot v\rangle$ para todo $w, v \in \mathbb{R}^n$. E é invertível quando $det\ H_u \neq 0$.

Seja $H : \mathbb{R}^n \to \mathbb{R}$ uma forma quadrática cuja matriz é $[h_{ij}]$. Se chamarmos de $H_u : \mathbb{R}^n \to \mathbb{R}^n$ o operador linear cuja matriz na base canônica de $\mathbb{R}^n$ também é $[h_{ij}]$, vemos de imediato que $H \cdot v^2 = \langle H_u \cdot v, v\rangle$ para todo $v \in \mathbb{R}^n$. Como a matriz $[h_{ij}]$ do operador linear H_u na base canônica é simétrica, H_u é auto-adjunto.

Reciprocamente, para qualquer operador auto-adjunto $H_u : \mathbb{R}^n \to \mathbb{R}^n$, a função $H : \mathbb{R}^n \to \mathbb{R}$, dada por $H \cdot v^2 = \langle H_u \cdot v, v\rangle$, é uma forma quadrática. Quando H é definida, o operador H_u é invertível pois $\langle H_u \cdot v, v\rangle \neq 0$ para todo $v \neq 0$, assim $H_u \cdot v \neq 0$ para todo $v \neq 0$.

Definição 3.46. Dada a função $f : U \to \mathbb{R}$, de classe C^2 no aberto $U \subset \mathbb{R}^n$, a forma quadrática hessiana $H(x) = (Hf)(x)$ de f no ponto $x \in U$ é aquela cuja matriz é;

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

Assim, para todo $v \in \mathbb{R}^n$, tem-se;

$$H(x) \cdot v^2 = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) v_i v_j.$$

A forma hessiana é usada para determinar a natureza dos pontos críticos da função f.

Definição 3.47. Um ponto $a \in U$ é dito um ponto de máximo (mínimo) local da função $f : U \to \mathbb{R}$ quando existe $\delta > 0$ tal que $f(x) \leq f(a)$ para todo $x \in U \cap B(a; \delta)(f(x) \geq f(a)$ para todo $x \in U \cap B(a; \delta))$.

Proposição 3.48. *Um ponto a, de máximo(ou de mínimo) local de uma função diferenciável f, é um ponto crítico de f.*

Demonstração: De fato, para todo $i = 1, ..., n$, se $\delta > 0$ é suficientemente pequeno então a função $\varphi_i : (-\delta, \delta) \to \mathbb{R}$, dada por $\varphi(t) = f(a + te_i)$ está bem definida e possui um máximo (ou mínimo) local no ponto $t = 0$, pois $\varphi(0) = f(a)$ e a depender da f, teremos para todo $x \in U \cap B(a; \delta)$; $f(a) \geq f(x)$ ou $f(a) \leq f(x)$. Assim, $\varphi_i'(0) = 0$ e por outro lado $\varphi_i'(0) = \lim_{h \to 0} \frac{\varphi_i(h) - \varphi_i(0)}{h} = \lim_{h \to 0} \frac{f(a + te_i) - f(a)}{h} = \frac{\partial f(a)}{\partial x_i}(a)$ para todo $i = 1, ..., n$. Logo, $\nabla f(a) = 0$. ■

Exemplo 3.49. A origem $0 \in \mathbb{R}^2$ é ponto crítico das três funções $f, g, h : \mathbb{R}^2 \to \mathbb{R}$, definidas por $f(x, y) = x^2 + y^2$, $g(x, y) = -x^2 - y^2$ e $h(x, y) = x^2 - y^2$.

De fato, note que $\nabla f(0, 0) = \nabla g(0, 0) = \nabla h(0, 0) = 0$ e mais; $f(0, 0) \leq f(x, y)$ para todo $(x, y) \in B(0, \delta) \cap \mathbb{R}^2$. Logo, a origem é ponto de mínimo local para f. Além disso, $g(0,0) \geq g(x,y)$ para todo $(x,y) \in B(0,\delta) \cap \mathbb{R}^2$. Portanto, a origem é ponto de máximo local para g. Já para h, a origem não é ponto de máximo e nem de mínimo, pois em qualquer disco centrado na origem, h assume valores maiores e menores do que $h(0,0) = 0$. Com isso, vemos que a recíproca da proposição anterior não é válida.

Teorema 3.50. *Sejam $f : U \to \mathbb{R}$ de classe C^2 e $a \in U$ um ponto crítico de f, então se a forma quadrática hessiana de f em a for;*

i) Positiva, então a é um mínimo local de f.

ii) Negativa, então a é um máximo local de f.

iii) Indefinida, então a não é máximo e nem mínimo local de f.

Demonstração:

i) Se $v \in \mathbb{R}^n$ é tal que $(a+v) \in U$, então pelo Teorema da Fórmula de Taylor, temos

$$f(a+v) - f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) v_i + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \cdot v_i \cdot v_j + r(v), \text{ onde} \lim_{v \to 0} \frac{r(v)}{\|v\|^2} = 0.$$

Como $a \in U$ é um ponto crítico de f, então

$$f(a+v) - f(a) = \frac{1}{2} H \cdot v^2 + r(v), \tag{3.12}$$

onde H é a forma quadrática hessiana de f em a. Uma vez que H é contínua e S^{n-1} é compacta, existe $v_0 \in S^{n-1}$ tal que $H \cdot v_0^2 = 2c \leq H \cdot v^2$, para todo $v \in S^{n-1}$. Como H é forma positiva, temos que $c > 0$. Retornando à equação (3.12) observamos que

$$\begin{aligned} f(a+v) - f(a) =& \frac{\|v\|^2}{2} \cdot H \cdot \left(\frac{v}{\|v\|}\right)^2 + \|v\|^2 \cdot \left(\frac{r(v)}{\|v\|^2}\right) \\ \geq & \frac{\|v\|^2}{2} \cdot (2c) + \|v\|^2 \cdot \left(\frac{r(v)}{\|v\|^2}\right) \\ =& \|v\|^2 \cdot \left(c + \frac{r(v)}{\|v\|^2}\right). \end{aligned} \tag{3.13}$$

Mas, lembre que $\lim\limits_{v \to 0} \dfrac{r(v)}{\|v\|^2} = 0$, assim, dado $\varepsilon = \frac{c}{2} > 0$, existe $\delta > 0$ tal que se $\|v\| < \delta$ implica que $\left|\frac{r(v)}{\|v\|^2}\right| < \frac{c}{2}$, então $\frac{-c}{2} < \frac{r(v)}{\|v\|} < \frac{c}{2}$. Voltando para a equação (3.13), segue que para $\|v\| < \delta$ implica que $f(a+v) - f(a) > \|v\|^2 \left(c - \frac{c}{2}\right) = \|v\|^2 \cdot \frac{c}{2} > 0$ para todo $v \neq 0$. Ou seja, $f(a+v) \geq f(a)$ para todo $v \in B_\delta(0)$. Logo, a é um mínimo local de f.

ii) Similar ao item anterior.

iii) Por H ser indefinida, existem $v, w \in \mathbb{R}^n$ tais que $H \cdot v^2 > 0$ e $H \cdot w^2 < 0$. Considerando $t \in \mathbb{R}$ suficientemente pequeno de modo que $(a+tv), (a+tw) \in U$, por (3.12) segue que,

$$\begin{aligned} f(a+tv) - f(a) &= \frac{1}{2}H(tv)^2 + r(tv) \\ &= \frac{t^2}{2}H \cdot v^2 + t^2\|v\|^2 \cdot \frac{r(tv)}{t^2 \cdot \|v\|^2} \\ &= \frac{t^2}{2}\left(H \cdot v^2 + 2\|v\|^2 \cdot \frac{r(tv)}{t^2 \cdot \|v\|^2}\right) \\ &> 0, \ \text{para } t \text{ suficientemente pequeno.} \end{aligned}$$

Então $f(a+tv) > f(a)$. Por outro lado;

$$\begin{aligned} f(a+tw) - f(a) &= \frac{t^2}{2}H \cdot w^2 + t^2\|w\|^2 \cdot \frac{r(tw)}{t^2 \cdot \|w\|^2} \\ &= \frac{t^2}{2}\left(H \cdot w^2 + 2\|w\|^2 \cdot \frac{r(tw)}{t^2 \cdot \|w\|^2}\right) \\ &< 0, \ \text{para } t \text{ suficientemente pequeno.} \end{aligned}$$

Então $f(a+tw) > f(a)$. Portanto, a nem é mínimo, nem máximo local de f.

■

Corolário 3.51. *Se a função $f : U \to \mathbb{R}$, de classe C^2, possui um mínimo(máximo) local no ponto $a \in U$, então a forma quadrática hessiana de f é não-negativa(não-positiva) nesse ponto.*

Demonstração: Supondo que f possui um mínimo local em a, se para algum $v_0 \in \mathbb{R}^n$ tivéssemos $H \cdot v_0^2 < 0$, pela demostração do Teorema 3.50, segue que $f(a+tv_0) < f(a)$ para todo t suficientemente pequeno, assim a não seria um ponto de mínimo local. Portanto, $H.v_0^2 \geq 0$, ou seja, não-negativa.

Supondo que f possui um máximo local em a, se para algum $w_0 \in \mathbb{R}^n$ tivéssemos $H \cdot w_0^2 > 0$, então $f(a + tw_o) > f(a)$ para todo t suficientemente pequeno, assim a não seria um ponto de máximo local. Portanto, $H \cdot w_0^2 \leq 0$, ou seja, não-positiva. ■

Pela demonstração anterior, vemos que quando a forma quadrática hessiana é positiva (negativa) no ponto a, então a é um ponto de mínimo (máximo) local estrito, ou seja, numa pequena bola de centro a não há outros pontos x onde $f(x) = f(a)$.

Exemplo 3.52. A origem é um ponto de mínimo estrito da função $f(x, y) = x^2 + y^2$, mas todos os pontos $(x, 0)$ do eixo das abcissas são pontos de mínimo não-estritos da função $g(x, y) = y^2$, em que $\mathbb{R}^2$ é o domínio de f e g.

Considerando $v = (\alpha, \beta)$, perceba que a forma hessiana de f na origem de $\mathbb{R}^2$ é

$$H \cdot v^2 = [\alpha\ \beta] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = 2\alpha^2 + 2\beta^2.$$

Enquanto que a da g é

$$K \cdot v^2 = [\alpha\ \beta] \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = 2\beta^2.$$

Portanto, vemos que H é positiva e K é não-negativa, pois embora $v \neq 0 \in \mathbb{R}^n$, nada impede que apenas β seja 0.

Exemplo 3.53. A forma quadrática hessiana da função $h(x, y) = x^2 - y^2$ é indefinida.

Pois considerando $v = (2, 1)$, temos que $H \cdot v^2 = [2\ 1] \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \cdot 2^2 - 2 \cdot 1^2 = 8 - 2 = 6 > 0$. E considerando $w = (1, 2)$, temos que $H \cdot v^2 = [1\ 2] \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2 \cdot 1^2 - 2 \cdot 2^2 = 2 - 8 = -6 < 0$. Assim, a forma quadrática hessiana de h é indefinida e pelo item *iii*) do Teorema 3.50, como $(0, 0)$ é ponto crítico de h, $(0, 0)$ não é ponto de máximo e nem de mínimo local (ponto de sela).

É natural agora, surgir o questionamento sobre a validade da recíproca do último Corolário 3.51. Com este fim, considere a função $f : \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x, y) = x^2 + y^3$. Perceba que $\nabla f(0) = (2 \cdot 0, 3 \cdot 0) = (0, 0)$, logo a origem de $\mathbb{R}^2$ é ponto crítico de f. Adotando $v = (\alpha, \beta)$, a forma quadrática hessiana é $H \cdot v^2 = [\alpha\ \beta] \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = 2\alpha^2 + 3\beta^2$. Perceba que H é não-negativa, contudo, $(0, 0)$ não é um ponto de mínimo local de f, pois $\left(0, \frac{-1}{n}\right) \to (0, 0)$ mas $f\left(0, \frac{1}{n}\right) = \frac{-1}{n^3} < f(0, 0)$ para todo $n \in \mathbb{N}$. Logo, a recíproca do Corolário 3.51 não é verdadeira.

3.6 A Derivada Como Uma Transformação Linear

Definição 3.54. Uma aplicação $f : U \to \mathbb{R}^n$, definida no aberto $U \subset \mathbb{R}^m$ é diferenciável no ponto $a \in U$ quando cada uma das suas funções-coordenada $f_1, ..., f_n : U \to \mathbb{R}$ é diferenciável neste ponto.

Assim, para todo $v \in \mathbb{R}^m$ tal que $a + v \in U$ e para cada $i = 1, ..., n$, tem-se;

$$f_i(a + v) - f_i(a) = \sum_{j=1}^{m} \frac{\partial f_i}{\partial x_j}(a) \cdot v_j + r_i(v), \text{ com } \lim_{v \to 0} \frac{r_i(v)}{\|v\|} = 0.$$

Definição 3.55. A matriz $Jf(a) = \left[\frac{\partial f_i}{\partial x_j}(a)\right] \in M_{(n \times m)}$ chama-se matriz jacobiana de f

no ponto a.

Definição 3.56. A transformação linear $f'(a) : \mathbb{R}^m \to \mathbb{R}^n$, cuja matriz em relação às bases canônicas de $\mathbb{R}^m$ e $\mathbb{R}^n$ é $Jf(a)$, chama-se a derivada da aplicação f no ponto a.

Pela definição de matriz de uma transformação linear, para todo $v \in \mathbb{R}^m$, temos

$$f'(a).v = (\beta_1, ..., \beta_n), \text{ onde } \beta_i = \sum_{j=1}^{m} \frac{\partial f_i}{\partial x_j}(a) \cdot v_j = \frac{\partial f_i}{\partial v}(a).$$

Definindo a derivada direcional da aplicação f, no ponto a, na direção do vetor v como $\frac{\partial f}{\partial v}(a) = \lim_{t\to 0} \dfrac{f(a+tv) - f(a)}{t}$, temos que

$$\frac{\partial f_i}{\partial v}(a) = \left(\frac{\partial f_1}{\partial v}(a), ..., \frac{\partial f_n}{\partial v}(a)\right) = f'(a) \cdot v.$$

Pela Regra da Cadeia, embora $\frac{\partial f}{\partial v}(a)$ tenha sido dada como $\frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0)$, onde $\lambda(t) = a + tv$, de modo geral vale a igualdade para qualquer caminho diferenciável $\lambda : (-\delta, \delta) \to U$, com $\lambda(0) = a$ e $\lambda'(0) = v$.

As n igualdades numéricas que exibem a diferenciabilidade das funções-coordenada f_i se resumem na seguinte igualdade entre vetores de $\mathbb{R}^n$:

$$f(a+v) - f(a) = f'(a) \cdot v + r(v), \ \textit{com} \ \lim_{v\to 0} \frac{r(v)}{\|v\|} = 0.$$

Por vezes, é mais adequado escrever esta condição da seguinte forma

$$f(a+v) - f(a) = f'(a) \cdot v + \rho(v) \cdot \|v\|, \ \textit{com} \ \lim_{v\to 0} \rho(v) = 0.$$

Onde $\rho(v) = \frac{r(v)}{\|v\|}$ para todo $v \neq 0$ tal que $a + v \in U$.

Essa relação caracteriza univocamente a derivada da aplicação f no seguinte sentido;

Proposição 3.57. *Se uma transformação linear* $T : \mathbb{R}^m \to \mathbb{R}^n$ *é tal que para* $a,\ a+v \in U$ *tem-se* $f(a+v) - f(a) = T \cdot v + r(v)$, *com* $\lim_{v\to 0} \dfrac{r(v)}{\|v\|} = 0$, *então* $T = f'(a)$.

Demonstração: Note que $f(a+tv) - f(a) = T \cdot tv + r(tv)$, assim,

$$f(a+tv) - f(a) = T \cdot tv + r(tv),$$

logo

$$\frac{f(a+tv) - f(a)}{t} = T \cdot v \pm \frac{r(tv)}{\|tv\|} \cdot \|v\|.$$

Fazendo $t \to 0$, temos

$$T \cdot v = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} = f'(a) \cdot v$$

Como vale para todo $v \neq 0$, concluímos que $T = f'(a)$. ■

Definição 3.58. Quando $f : U \to \mathbb{R}^n$ é diferenciável em todos os pontos de U, dizemos que f é diferenciável em U.

No caso da aplicação f diferenciável em U, fica definida uma aplicação $f' : U \to \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ que faz corresponder a cada $x \in U$ a transformação linear $f'(x) : \mathbb{R}^m \to \mathbb{R}^n$, em que $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ é o conjunto das transformações lineares de $\mathbb{R}^m$ em $\mathbb{R}^n$ e, por vezes, o identificaremos como $M_{(n\times m)}$ o conjunto das matrizes $n \times m$ ou como o espaço $\mathbb{R}^{nm}$.

Observação 3.59. Dizer que a aplicação derivada $f' : U \to \mathbb{R}^{nm}$ é contínua, equivale a afirmar a continuidade das suas nm funções-coordenada $\frac{\partial f_i}{\partial x_j} : U \to \mathbb{R}$, ou seja, $f \in C^1$.

Assim como para funções, aplicações $f : U \to \mathbb{R}^n$ de classe C^k são definidas por indução: diz-se que $f \in C^k$ quando f é diferenciável e sua derivada $f' : U \to \mathbb{R}^{nm}$ é de classe C^{k-1}. Se $f \in C^k$ para todo $k \in \mathbb{N}$, dizemos que $f \in C^\infty$ e portanto, $f' \in C^\infty$ também.

Observação 3.60. Geralmente, a forma mais simples de verificar que uma aplicação f é diferenciável, se resume em calcular diretamente as derivadas parciais $\frac{\partial f_i}{\partial x_j}(x)$, mostrar que elas dependem continuamente de x e usar o Teorema 3.11 que garante a diferenciabilidade a toda função de classe C^1.

Quase todas as aplicações diferenciáveis vistas em modelagem matemática e em ciência básica, são de classe C^1. No entanto, ocorre que as propriedades mais relevantes das aplicações C^1, resultam da relação que caracteriza sua diferenciabilidade. Por isso, esse conceito é tão importante.

Exemplo 3.61. Sejam $I \subset \mathbb{R}$ um intervalo aberto e $f : I \to \mathbb{R}^n$ um caminho diferenciável no ponto $a \in I$. Considerando f como uma aplicação, sua derivada no ponto a é a transformação linear $f'(a) : \mathbb{R} \to \mathbb{R}^n$ cuja matriz jacobiana tem por única coluna o vetor $v = \left(\frac{df_1}{dt}(a), ..., \frac{df_n}{dt}(a)\right)$, a qual vem a ser o vetor velocidade do caminho f no ponto a, indicado por $f'(a)$ anteriormente. Como $f'(a) : \mathbb{R} \to \mathbb{R}^n$ é uma transformação linear, então ela faz corresponder a cada "vetor" $t \in \mathbb{R}$, o vetor $t \cdot v \in \mathbb{R}^n$. Ou seja, $f'(a) \cdot t = t \cdot f'(a)$.

Exemplo 3.62. Se $T : \mathbb{R}^m \to \mathbb{R}^n$ é uma transformação linear, então T é diferenciável e $T'(a) = T$ para todo $a \in \mathbb{R}^m$. Ou seja, $T'(a) \cdot v = T \cdot v$, quaisquer que sejam $a, v \in \mathbb{R}^m$.

Perceba que;

$$\begin{aligned} T(a+v)-T(a) &= T(a)+T(v)-T(a)\\ &= T(v)\\ &= T(v)+r(v),\ \textit{onde } r(v)=0 \textit{ para todo } v\in\mathbb{R}^m. \end{aligned}$$

Assim, pela Proposição 3.57, temos que $T\cdot v = T'(a)\cdot v$, portanto, $T = T'(a)$.

Exemplo 3.63. Seja $B:\mathbb{R}^m\times\mathbb{R}^n\to\mathbb{R}^p$ uma aplicação bilinear, isto é, linear em cada uma de suas duas variáveis. Se escrevermos, para cada par de vetores $(e_i,\widetilde{e}_j)$ das bases canônicas de $\mathbb{R}^m$ e $\mathbb{R}^n$ respectivamente, $B(e_i,\widetilde{e}_j)=v_{ij}$ então para $x\in\mathbb{R}^m$ e $y\in\mathbb{R}^n$ teremos $B(x,y)=\sum_{i,j}x_iy_jv_{ij}$. Isto mostra que B é contínua, portanto assume seu valor máximo $|B|$ no compacto $S^{m-1}\times S^{n-1}$. Assim, para quaisquer $x\in\mathbb{R}^m$ e $y\in\mathbb{R}^n$ não-nulos vale; $\|B(x,y)\| = \left\|B\left(\frac{x}{\|x\|},\frac{y}{\|y\|}\right)\right\|\cdot\|x\|\cdot\|y\|\le|B|\cdot\|x\|\cdot\|y\|$. Considerando $x=0$ ou $y=0$, temos a desigualdade $\|B(x,y)\|\le|B|\cdot\|x\|\cdot\|y\|$ imediatamente, pois $B(0,y)=0=B(x,0)$. Mostremos agora, que toda aplicação bilinear B é diferenciável com $B'(x,y)=(u,v)=B(u,y)+B(x,v)$. De fato; se $x,u\in\mathbb{R}^m$ e $y,v\in\mathbb{R}^n$, pela bilinearidade de B, temos

$$\begin{aligned} B(x+u,y+v)-B(x,y) &= B(x+u,y)+B(x+u,v)-B(x,y)\\ &= B(x,y)+B(u,y)+B(x,v)+B(u,v)-B(x,y)\\ &= B(u,y)+B(x,v)+B(u,v). \end{aligned}$$

E como $\|(u,v)\|=\sqrt{\|u\|^2+\|v\|^2}\ge\|v\|$, então

$$0\le\left\|\frac{B(u,v)}{\|(u,v)\|}\right\|\le\frac{\|B(u,v)\|}{\|(u,v)\|}\le\frac{|B|\|u\|\cdot\|v\|}{\|u\|}=|B|\|v\|.$$

Logo, $\displaystyle\lim_{(u,v)\to0}\frac{B(u,v)}{\|(u,v)\|}=0$. Portanto, B é diferenciável.

Teorema 3.64. *(Regra da Cadeia) Sejam $U\subset\mathbb{R}^m$, $V\subset\mathbb{R}^n$ abertos e $f:U\to\mathbb{R}^n$, $g:V\to\mathbb{R}^p$ diferenciáveis nos pontos $a\in U$, $b=f(a)\in V$, com $f(U)\subset V$. Então $g\circ f:U\to\mathbb{R}^p$ é diferenciável no ponto a com*

$$(g\circ f)'(a)=g'(b)\cdot f'(a):\mathbb{R}^m\to\mathbb{R}^p.$$

Ou seja, a derivada da aplicação composta é a composta das derivadas.

Demonstração: Podemos escrever $f(a+v)=f(a)+f'(a)\cdot v+\rho(v)\cdot\|v\|$, com $\lim\limits_{v\to0}\rho(v)=0$ e $g(b+w)=g(b)+g'(b)\cdot w+\sigma(w)\cdot\|w\|$, com $\lim\limits_{w\to0}\sigma(w)=0$. Assim, $(g\circ f)(a+v)=$

$g(f(a) + f'(a) \cdot v + \rho(v) \cdot \|v\|)$. Tomando $w = f'(a) \cdot v + \rho(v) \cdot \|v\|$, temos;

$$\begin{aligned}(g \circ f)(a + v) &= g(b + w) \\ &= g(b) + g'(b) \cdot w + \sigma(w) \cdot \|w\| \\ &= g(b) + g'(b)(f'(a) \cdot v + \rho(v) \cdot \|v\|) + \sigma(w) \cdot \|w\| \\ &= g(b) + g'(b)(f'(a) \cdot v) + g'(b)\rho(v) \cdot \|v\| + \sigma(w) \cdot \|w\| \\ &= (g \circ f)(a) + (g'(b) \cdot f'(a)) \cdot v + C(v) \cdot \|v\|\end{aligned}$$

onde $C(v) = g'(b) \cdot \rho(v) + \sigma(w) \cdot \left\| f'(a) \cdot \frac{v}{\|v\|} + \rho(v) \right\|$. Se $v \to 0$, então $w \to 0$ e $f'(a) \cdot \frac{v}{\|v\|}$ é limitada. Portanto, $\lim\limits_{v \to 0} C(v) = 0$. ■

Observação 3.65. Se $f : U \to \mathbb{R}^n$ e $g : V \to \mathbb{R}^p$, onde $U \subset \mathbb{R}^n$ e $f(U) \subset V \subset \mathbb{R}^n$, são de classe C^k, então $g \circ f : U \to \mathbb{R}^p$ é de classe C^k.

Corolário 3.66. *(As Regras de Derivação) Sejam $f, g : U \to \mathbb{R}^n$ diferenciáveis no ponto $a \in U \subset \mathbb{R}^m$, α um número real e $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^p$ bilinear. Então*

i) $f + g : U \to \mathbb{R}^n$ *é diferenciável no ponto a, com* $(f + g)'(a) = f'(a) + g'(a)$.

ii) $\alpha \cdot f : U \to \mathbb{R}^n$ *é diferenciável no ponto a, com* $(\alpha \cdot f)'(a) = \alpha + f'(a)$.

iii) $B(f, g) : U \to \mathbb{R}^p$ *definida por* $B(f, g)(x) = B(f(x), g(x))$ *é diferenciável no ponto a, com* $[B(f, g)]'(a) \cdot v = B(f'(a) \cdot v, g(a)) + B(f(a), g'(a) \cdot v)$.

Demonstração:

i) Considerando a transformação $S : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ dada por $S(x, y) = x + y$, note que $f + g = S \circ (f, g)$ e como $S' = S$, pela Regra da Cadeia temos $(f + g)' = S \circ (f', g') = f' + g'$. Logo, $(f + g)'(a) = f'(a) + g'(a)$.

ii) Considerando a transformação $\mu_\alpha : \mathbb{R}^n \to \mathbb{R}^n$ dada por $\mu_\alpha(x) = \alpha \cdot x$, note que $\alpha \cdot f = \mu_\alpha \circ f$ e como $(\mu_\alpha)' = \mu_\alpha$, pela Regra da Cadeia temos $(\alpha \cdot f)' = (\mu_\alpha \circ f) = \mu_\alpha \circ f' = \alpha \cdot f'$. Logo, $(\mu_\alpha \cdot f)'(a) = \alpha \cdot f'(a)$.

iii) Como $B(f, g) = B \circ (f, g)$, para cada $v \in \mathbb{R}^m$ temos;

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

■

Observação 3.67. Uma aplicação bilinear $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^p$ pode e deve ser considerada como uma forma de multiplicar um elemento de $\mathbb{R}^n$ por outro, obtendo um produto em $\mathbb{R}^p$. Usando a notação multiplicativa $x \bullet y$ em vez de $B(x, y)$, o item $iii)$ da Regra de Derivação lê-se $(f \bullet g)' = f' \bullet g + f \bullet g'$, isto é, para todo $x \in U$ e todo $v \in \mathbb{R}^n$, $(f \bullet g)'(x) \cdot v = (f'(x) \cdot v) \bullet g(x) + f(x) \bullet (g'(x) \cdot v)$, em que $\cdot$ é a aplicação de uma transformação linear sobre um vetor.

Teorema 3.68. *(Desigualdade do Valor Médio) Seja $f : U \to \mathbb{R}^n$ 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 $\lambda : [0, 1] \to \mathbb{R}^n$, definido por $\lambda(t) = f(a + tv)$ é diferenciável e $\lambda'(t) = f'(a + tv) \cdot v$, logo;

$$\|\lambda'(t)\| \leq \|f'(a + tv)\| \cdot \|v\| \leq M \cdot \|v\|, \text{ para todo } t \in [0, 1].$$

Assim, pela Desigualdade do Valor Médio para caminhos, temos que existe $c \in (0, 1)$ tal que $\|\lambda(1) - \lambda(0)\| \leq \|\lambda'(c)\| \cdot (1 - 0)$, o que implica em

$$\|f(a + v) - f(a)\| \leq M \cdot \|v\| \cdot 1 = M \cdot \|v\|.$$

■

3.7 O Teorema da Função Inversa

Estamos interessados em aplicações bijetivas diferenciáveis, cuja inversa também é diferenciável.

Em Análise Real, vemos o Teorema de Darboux, que nos afirma que: "Seja $f[a, b] \to \mathbb{R}$ derivável. Se $f'(a) < d < f'(b)$, então existe $c \in (a, b)$ tal que $f'(c) = d$". Nele, bastaria considerar $f : I \to J$, onde $I \subset \mathbb{R}$ e $J = f(I)$ e supor que $f'(x) \neq 0$ para todo $x \in I$, para garantirmos que f possui inversa, basta-nos verificar que f é injetora, para isso, nos é suficiente mostrar que $f'(x) > 0$ para todo $x \in I$ ou $f'(x) < 0$ para todo $x \in I$. Supondo por contradição que existem x_1, x_2 de modo que $f'(x_1) < 0$ e $f'(x_2) > 0$, pelo Teorema de Darboux existe $c \in (x_1, x_2)$ tal que $f'(c) = 0$, o que é um absurdo.

Como visto no parágrafo acima, para garantirmos que a função $f : I \to J$ é um difeomorfismo, nos é suficiente mostrar que $f(I) = J$ e que $f'(x) \neq 0$ para todo $x \in I$. A seguir, veremos sob quais condições funções são difeomorfismos em dimensões maiores que 1.

Definição 3.69. Sejam $U \subset \mathbb{R}^m$ e $V \subset \mathbb{R}^n$ abertos. Uma aplicação $f : U \to V$ é dita um difeomorfismo entre U e V quando é uma bijeção diferenciável, cuja inversa

$g = f^{-1} : V \to U$ também é diferenciável.

Se $f : U \to V$ é um difeomorfismo, com $g = f^{-1} : V \to U$, então $g \circ f = I_U$ e $f \circ g = I_V$ resulta, pela Regra da Cadeia que; $g'(f(x)) \cdot f'(x) = I_{\mathbb{R}^m}$ e $f'(x) \cdot g'(f(x)) = I_{\mathbb{R}^n}$ para todo $x \in U$. Assim, $f'(x) : \mathbb{R}^m \to \mathbb{R}^n$ é um isomorfismo cujo inverso é $g'(f(x)) : \mathbb{R}^n \to \mathbb{R}^m$.

Exemplo 3.70. As aplicações $f, g : \mathbb{R}^2 \to \mathbb{R}^2$ dadas por $f(x) = 2x$ e $g(y) = \frac{1}{2} \cdot y$ são difeomorfismos.

Observe que, $(g \circ f)(x) = g(f(x)) = g(2x) = \frac{1}{2} \cdot 2x = x$ e $(f \circ g)(y) = f(g(y)) = f\left(\frac{1}{2} \cdot y\right) = 2 \cdot \frac{y}{2} = y$. Assim, f e g são homeomorfismos. Note ainda que para todo $v \in \mathbb{R}^2$ tal que $x + v \in \mathbb{R}^2$, por definição temos $f(x+v) - f(x) = f'(x) \cdot v + r(v)$ e portanto, $r(v) = 0$. O que implica que $\displaystyle\lim_{\|v\| \to 0} \frac{r(v)}{\|v\|} = 0$. Semelhantemente, para todo $s \in \mathbb{R}^2$ tal que $y + s \in \mathbb{R}^2$ temos que $g(y+s) - g(y) = g'(y) \cdot s + r(s)$ e consequentemente $r(s) = 0$. O que implica que $\displaystyle\lim_{\|s\| \to 0} \frac{r(s)}{\|s\|} = 0$. Logo, f e g são diferenciáveis, portanto, difeomorfos.

Observação 3.71. Nem todo homeomorfismo diferenciável é um difeomorfismo. Ou seja, nem todo homeomorfismo diferenciável possui inversa diferenciável.

Considere $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^3$. Como $f'(x) = 3x^2$ e $f'(0) = 0$, a função inversa de f que é $g(x) = \sqrt[3]{x}$ não é diferenciável no ponto $0 = f(0)$, pois

$$g'(0) = \lim_{h \to 0} \frac{g(0+h) - g(0)}{h} = \lim_{h \to 0} \frac{g(h) - 0}{h} = \lim_{h \to 0} \frac{\sqrt[3]{h}}{h} = \lim_{h \to 0} h^{\frac{-2}{3}} = \infty.$$

Logo, $g'(0)$ não existe.

Definição 3.72. Seja $U \subset \mathbb{R}^m$ um aberto. Uma aplicação diferenciável $f : U \to \mathbb{R}^m$, é dita um difeomorfismo local, quando para cada $x \in U$ existe uma bola aberta $B = B(x; \delta) \subset U$ tal que f aplica B difeomorficamente sobre um aberto V contendo $f(x)$.

Segue então que se $f : U \to \mathbb{R}^m$ é um difeomorfismo local, então $f'(x) : \mathbb{R}^m \to \mathbb{R}^m$ é um isomorfismo, para todo $x \in U$.

Veremos adiante que a recíproca é válida quando $f \in C^1$, no Teorema da Aplicação Inversa.

Resulta da última definição que um difeomorfismo local $f : U \to \mathbb{R}^m$ é uma aplicação aberta, ou seja, a imagem $f(A)$ de qualquer aberto $A \subset U$ é um subconjunto aberto de $\mathbb{R}^m$. De fato, tomando para cada $x \in A$ uma bola aberta $B_x \subset A$, centrada em x de modo que f seja um difeomorfismo de B_x sobre um aberto $V_x \subset \mathbb{R}^m$, então $A = \cup_{x \in A} B_x$ e $f(A) = f(\cup_{x \in A} B_x) = \cup_{x \in A} f(B_x) = \cup_{x \in A} V_x$ é uma união de abertos, portanto, $f(A)$ é aberto.

Observação 3.73. O difeomorfismo local $f : U \to \mathbb{R}^m$ é um difeomorfismo global de U sobre o aberto $V = f(U) \subset \mathbb{R}^m$ se, e somente se, é uma aplicação injetiva.

Assumindo que $f : U \to \mathbb{R}^M$ é um difeomorfismo local e injetiva, vamos mostrar que f é um difeomorfismo global. Note que se f é injetiva e $f|_{f(U)}$ é sobrejetiva, então $f : V \to U$ é um homeomorfismo, visto que admite inversa contínua. Vamos mostrar que essa inversa $g = f^{-1}$ é diferenciável. Seja $b \in V$, vamos mostrar que g é diferenciável em b. Por f ser um difeomorfismo local, existe $B = B(a;\delta) \subset U$ tal que $f|_B$ é um difeomorfismo sobre um aberto $\hat{V}$ com $a = f(b) \in \hat{V}$. Note que para $g|_{\hat{V}}$, g continua sendo a inversa de f e pela unicidade da inversa, $g = f^{-1}|_{\hat{V}}$. Assim, g é diferenciável em b. A recíproca é obtida de maneira bem similar, usando novamente a unicidade da função inversa.

Lema 3.74. *A aplicação* $inv : \mathcal{L}_{inv}(\mathbb{R}^m) \to \mathcal{L}_{inv}(\mathbb{R}^m)$*, definida por* $inv(T) = T^{-1}$ *é uma aplicação de classe* C^∞*.*

Demonstração: Inicialmente vamos verificar que $\partial inv/\partial V : U \to M$ é contínua. Seja $U = M_{inv}(m \times m)$ e $M = M(m \times m)$, perceba que

$$\frac{\partial inv}{\partial V}(X) = f'(X) \cdot V = -X^{-1}VX^{-1}.$$

Sejam $\varphi_V : M \times M \to M$ e $\beta : M \to M \times M$, definidas por $\varphi_V(X,Y) = XVY$ e $\beta(Z) = (Z,Z)$. Então $\varphi_V, \beta \in C^\infty$, mas $(-\varphi_V) \circ \beta \circ (inv)(X) = -X^{-1}VX^{-1}$. Assim

$$\partial inv/\partial V = (-\varphi_V) \circ \beta \circ inv. \tag{3.14}$$

A demonstração segue por indução. Como inv é diferenciável, $inv \in C^0$ e a igualdade 3.14 implica que $\partial inv/\partial V \in C^0$, logo $(inv)' \in C^0$, portanto $inv \in C^1$. Supondo que $inv \in C^k$, a mesma igualdade 3.14, implica que $inv \in C^{k+1}$. Assim $inv \in C^k$ para todo $k \in \mathbb{N}$, logo $inv \in C^\infty$. ■

Teorema 3.75. *Se o difeomorfismo* $f : U \to V$ *é de classe* C^k $(k \geq 1)$*, então seu inverso* $g = f^{-1} : V \to U$ *também é de classe* C^k*.*

Demonstração: Observe que dado $y \in V$, existe um único $x \in U$ tal que $f(x) = y$ ou $g(y) = x$. Agora, temos $g'(y) = [f'(x)]^{-1} = [f'(f^{-1}(y))]^{-1}$. Uma vez que $g'(y)$ é uma transformação linear de $\mathbb{R}^m$ e $\mathbb{R}^m$, ou seja, $g'(y) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m) = \mathbb{R}^{m^2}$, podemos escrever g' como a seguinte composição

$$g' = (inv) \circ f' \circ f^{-1}, \tag{3.15}$$

onde inv é o operador que leva um operador invertível T em T^{-1}. Sabendo que $inv \in C^\infty$, além disso, se $f \in C^k$, temos que $f' \in C^{k-1}$.

- Se $k = 1$, por (3.15), segue que $g' \in C^0 \Rightarrow g \in C^1$.

- Suponha que para todo $h \in C^{k-1}$, tenhamos que $h^{-1} \in C^{k-1}$. Como $f \in C^k$ em particular $f \in C^{k-1}$, logo pela hipótese de indução, $f^{-1} \in C^{k-1}$ e por (3.15), obtemos que $g' \in C^{k-1}$ e, portanto, $g \in C^k$ pelo Princípio de Indução Finita.

■

Teorema 3.76. *Seja $f : U \to \mathbb{R}^n$ de classe C^1 no aberto $U \subset \mathbb{R}^m$. Se, para algum $a \in U$, a derivada $f'(a) : \mathbb{R}^m \to \mathbb{R}^n$ 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: Perceba que $u \mapsto \|f'(a) \cdot u\|$ é positiva em todos os pontos u da esfera unitária S^{m-1}, que é compacta. Pelo Corolário 2.19, existe $c > 0$ tal que $\|f'(a) \cdot u\| \geq 2c$ para todo $u \in S^{m-1}$. Pela linearidade, segue que $\|f'(a) \cdot v\| \geq 2c \cdot \|v\|$ para todo $v \in \mathbb{R}^m$. Além disso, para todo $x \in U$, escrevamos $r(x) = f(x) - f(a) - f'(a)(x - a)$. Então, para $x, y \in U$ quaisquer, temos $f(x) - f(y) = f'(a).(x - y) + r(x) - r(y)$. Pela desigualdade triangular, segue que;

$$\begin{aligned} \|f(x) - f(y)\| &\geq \|f'(a) \cdot (x - y)\| - \|r(x) - r(y)\| \\ &\geq 2c \cdot \|x - y\| - \|r(x) - r(y)\|. \end{aligned}$$

Notamos que a aplicação r é de classe C^1, onde $r(a) = 0$ e $r'(a) = 0$. Por r' ser contínua, existe $\delta > 0$ tal que $\|x-a\| < \delta$ implica que $x \in U$ e $\|r'(x)\| < c$. Aplicando a desigualdade do Valor Médio em r no conjunto convexo $B = B(a;\delta)$, temos que se $x, y \in B$, então $\|r(x) - r(y)\| \leq c\|x - y\|$. Como consequência, $x, y \in B$ implica em $\|f(x) - f(y)\| \geq 2c\|x - y\| - c\|x - y\| = c\|x - y\|$. ■

Teorema 3.77. *(Diferenciabilidade do Homeomorfismo Inverso) Seja $f : U \to V$ um homeomorfismo de classe C^1 entre os abertos $U, V \subset \mathbb{R}^m$. Se para algum $x \in U$, a derivada $f'(x) : \mathbb{R}^m \to \mathbb{R}^m$ é um operador invertível, então o homeomorfismo inverso $g = f^{-1} : V \to U$ é diferenciável no ponto $f(x)$, com $g'(f(x)) = [f'(x)]^{-1}$.*

Demonstração: Sejam $x \in U$ e $x + v \in U$, escrevamos $f(x) = y$ e $f(x + v) = y + w$. Então, $w = f(x+v)-f(x) = f'(x)\cdot v+r(v)$, $\lim\limits_{v\to 0} r(v)/\|v\| = 0$ e $v = g(f(x+v))-g(f(x)) = g(y + w) - g(y)$. Para provarmos que $f'(x)^{-1}$ é a derivada de g no ponto y, escrevemos

$$s(w) = g(y + w) - g(y) - f'(x)^{-1} \cdot w, \tag{3.16}$$

e basta-nos mostrar que $\lim\limits_{w\to 0} \dfrac{s(w)}{\|w\|} = 0$. Substituindo as expressões de v e w na equação

(3.16), temos $v = f'(x)^{-1}[f'(x) \cdot v + r(v)] + s(w)$. Assim

$$\begin{aligned} v &= f'(x)^{-1} \cdot f'(x) \cdot v + f'(x)^{-1} \cdot r(v) + s(w) \\ &= I \cdot v + f'(x)^{-1} \cdot r(v) + s(w) \\ &= v + f'(x)^{-1} \cdot r(v) + s(w). \end{aligned}$$

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

$$\frac{s(w)}{\|w\|} = -f'(x)^{-1} \cdot \frac{r(v)}{\|v\|} \cdot \frac{\|v\|}{\|w\|}. \tag{3.17}$$

Substituindo a expressão de w na equação (3.17), temos

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

Quando $w \to 0$, temos que $v \to 0$ por g ser contínua, visto que f é um homeomorfismo. Assim, $\frac{r(v)}{\|v\|} \to 0$. Além disso, como $f \in C^1$ no aberto $U \subset \mathbb{R}^m$ e $f'(x)^{-1}$ é injetiva, pelo Teorema 3.76, existem $\delta > 0$ e $c > 0$ tais que $\|v\| < \delta$ implica $\|f(x+v) - f(x)\| \geq c\|v\|$, portanto

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

Logo, $\lim\limits_{w \to 0} \dfrac{s(w)}{\|w\|} = 0$ e, portanto, g é diferenciável em $f(x)$ e $g'(f(x)) = [f'(x)]^{-1}$. ■

Corolário 3.78. *Se $f : U \to V$ é um homeomorfismo de classe C^k cuja derivada $f'(x) : \mathbb{R}^m \to \mathbb{R}^m$ é invertível para todo $x \in U$, então seu inverso $g = f^{-1} : V \to U$ é de classe C^k.*

Demonstração: De fato, a derivada $g' : V \to \mathcal{L}(\mathbb{R}^m)$, dada por $g'(y) = f'(g(y))^{-1}$ para cada $y \in V$, pode ser escrita como $g' = inv \circ f' \circ g$, onde a aplicação inv, de classe C^∞ é a inversão de transformações lineares bijetivas e $f' \in C^{k-1}$. Admitindo, por indução, que $g \in C^{k-1}$, resulta que $g' \in C^{k-1}$, logo $g \in C^k$. ■

Lema 3.79. *Sejam $U \subset \mathbb{R}^m$ aberto e $g : U \to \mathbb{R}^n$ diferenciável no ponto $a \in U$, com $g'(a) : \mathbb{R}^m \to \mathbb{R}^n$ sobrejetiva. Se a é um ponto de mínimo local de $\|g(x)\|$, com $x \in U$, então $g(a) = 0$.*

Demonstração: Se a é um ponto de mínimo local para $\|g(x)\|$, a também será um ponto de mínimo local para $\varphi : U \to \mathbb{R}$, dada por $\varphi(x) = \langle g(x), g(x)\rangle$. Note que $\varphi'(x) = \langle g'(x), g(x)\rangle + \langle g(x), g'(x)\rangle = 2\langle g'(x), g(x)\rangle$. Então, $\varphi'(a) = 2\langle g'(a), g(a)\rangle$. Como $\|g(a)\| \leq \|g(x)\|$ para todo x em U suficientemente próximo de a, temos que $\varphi(a) \leq \varphi(x)$, para todo x em U suficientemente próximo de a. Logo, a é um mínimo local para φ. Por φ ser diferenciável, a é ponto crítico, então $\varphi'(a) = 0$. Então, para todo $v \in \mathbb{R}^m$, temos

$$\begin{aligned}\varphi'(a) \cdot v &= 2\langle g'(a) \cdot v, g(a)\rangle \\ 0 \cdot v &= 2\langle g'(a) \cdot v, g(a)\rangle \\ 0 &= 2\langle g'(a) \cdot v, g(a)\rangle.\end{aligned}$$

Por $g'(a)$ ser sobrejetiva, ela não pode ser a matriz nula. Escolhendo v_0 de modo que $g'(a) \cdot v_0 = g(a)$, temos que;

$$0 = \varphi'(a) \cdot v_0 = 2\langle g(a), g(a)\rangle = 2\|g(a)\|^2.$$

Portanto, $g(a) = 0$. ■

Teorema 3.80. *(Teorema da Função Inversa) Seja $f : U \to \mathbb{R}^m$ de classe C^k $(k \geq 1)$ no aberto $U \subset \mathbb{R}^m$. Se $a \in U$ é tal que $f'(a) : \mathbb{R}^m \to \mathbb{R}^m$ é invertível, então existe uma bola aberta $B = B(a;\delta) \subset U$ tal que a restrição $f|_B$ é um difeomorfismo, cuja inversa é de classe C^k, sobre um aberto $V \ni f(a)$.*

Demonstração:

Por U ser aberto, existe $\delta_1 > 0$ tal que $B(a;\delta_1) \subset U$ e, pelo Teorema 3.76, existe $\delta_2 > 0$ de modo que $f|_{B(a;\delta_2)}$ é injetiva. Tomando $\delta_3 = min\{\delta_1, \delta_2\}$ e $\delta = \frac{\delta_3}{2}$ obtemos que $f : B_1 \to f(B_1)$ é bijetiva, onde $B_1 = B[a;\delta] = \overline{B}$. Por f ser contínua, usando o Teorema 2.42, obtemos que $f : B_1 \to f(B_1)$ é um homeomorfismo de classe C^1 e como $f'(a)$ é invertível por hipótese, segue $f'(x)$ é um operador invertível para todo $x \in B$ "próximo" de a. Mostremos então, que $f(B) \subset \mathbb{R}^m$ é aberto. Seja $p \in B$ tal que $f(p) = q$, considere $S = S[a;\delta]$ que é a fronteira de $\overline{B}$. Como $f|_{\overline{B}}$ é injetiva, $q = f(p) \notin f(S)$ pois se $q \in f(S)$ existiria $x_1 \in S$ tal que $f(x_1) = q$, então $f(p) = q = f(x_1)$, mas como f é injetiva, $p = x_1$, o que é um absurdo pois p pertence ao interior da bola e x_1 está na esfera. Portanto, existe $\varepsilon > 0$ tal que $\|f(x) - q\| \geq 2\varepsilon$, para todo $x \in S$, pois $f(S)$ é compacto. Agora, afirmamos que $B(q;\varepsilon) \subset f(B)$. Com efeito, seja $y \in B(q;\varepsilon)$, considerando $g(x) = f(x) - y$, temos que o mínimo de $\|g(x)\|$ não é atingido num ponto $x \in S$ quando x varia no compacto $\overline{B}$, pois se $x_1 \in S$, teríamos $h(x_1) \leq h(x)$ para todo $x \in B$. Mas, perceba que $\|f(x_1) - y\| \geq h(x_1) = \|f(x_1) - y\| = \|f(x_1) - q + q - y\| \geq \|f(x_1) - q\| - \|q - y\| > 2\varepsilon - \varepsilon = \varepsilon$. Enquanto que se $x = p$, teríamos que $\|f(p) - y\| = \|q - y\| = \|y - q\| < \varepsilon$. Portanto, $\varepsilon > h(x) \geq h(x_1) > \varepsilon$, logo $\varepsilon > \varepsilon$, o que é um absurdo. Assim, o mínimo de $\|g(x)\| = \|f(x) - y\|$ é atingido num ponto $x_0 \in B$. Diante disso, perceba que x_0 é ponto de mínimo de $\|g(x)\|$ e $f'(x_0)$ é sobrejetiva, então pelo Lema anterior $g(x_0) = 0$, ou seja, $f(x_0) - y = 0$, portanto $f(x_0) = y$. Assim $y \in f(B)$, então $B(q;\varepsilon) \subset f(B)$. Tendo feito uso do Teorema 3.77, concluímos que $f : B \to f(B)$ é um difeomorfismo, e mais, pelo Corolário 3.75 tal difeomorfismo é de classe C^k. ■

Corolário 3.81. *Seja $a \in U$ um ponto crítico da função $f : U \to \mathbb{R}$ de classe C^2 no aberto $U \subset \mathbb{R}^n$. Se a matriz Hessiana $Hf(a) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right]$ é invertível, então existe um aberto V, com $a \in V \subset C$, no qual não há outros pontos críticos de f.*

Demonstração: Com efeito, a matriz Hessiana $Hf(x)$ é, para todo $x \in U$, a matriz Jacobiana da aplicação $F : U \to \mathbb{R}^n$, $F(x) = \nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), ..., \frac{\partial f}{\partial x_n}(x)\right)$. Como $Hf(a)$ é invertível, F é injetiva numa vizinhança $V \ni a$, então $F(x) \neq F(a)$, ou seja, $\nabla f(x) \neq 0$ para toda $x \in V \backslash \{a\}$. ■

Quando $\nabla f(a) = 0$ e $Hf(a)$ é invertível, dizemos que a é um ponto crítico não-degenerado da função f. Assim, pontos críticos não-degerenados são os pontos críticos isolados.

3.8 O Teorema da Função Implícita

Nesta seção os pontos do espaço $\mathbb{R}^{m+n}$ serão representados sob a forma $z = (x, y)$, onde $x \in \mathbb{R}^m$ e $y \in \mathbb{R}^n$.

Definição 3.82. Um difeomorfismo $h : U \to V$, entre abertos $U, V \subset \mathbb{R}^{m+n}$ é chamado de vertical quando for do tipo $h(x, y) = (x, h_2(x, y))$. Ou seja, quando deixar a coordenada x invariante.

Observação 3.83. O inverso de um difeomorfismo vertical também é vertical.

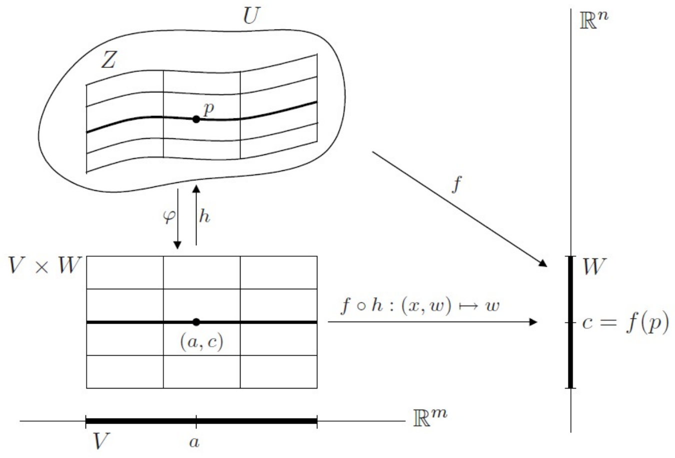
Usualmente um difeomorfismo $\varphi : U \to V$ é interpretado como uma transformação geométrica que aplica diferenciavelmente o conjunto U sobre o conjunto V, de forma invertível. Contudo, às vezes é conveniente olhar para φ como uma mudança de coordenadas, em que as coordenadas dos ponto $x \in U$ passam a ser aquelas da sua imagem $y = \varphi(x) \in V$.

Nessa perspectiva, o teorema seguinte nos diz que se a derivada de uma aplicação f, de classe C^k, é sobrejetiva num ponto p, então é possível obter de maneira simples, um sistema de coordenadas válido numa vizinhança aberta Z de p, tal que, em termos dessas novas coordenadas, a aplicação f assume a expressão;

$$(x_1, ..., x_m, w_1, ..., w_n) \mapsto (w_1, ..., w_n).$$

Teorema 3.84. *(Forma Local das Submersões) Seja $f = (f_1, ..., f_n)$ uma aplicação de classe C^k $(k \geq 1)$ de um aberto $U \subset \mathbb{R}^{m+n}$ em $\mathbb{R}^m$. Se num ponto $p = (a, b) \in U$, a matriz $\left[\frac{\partial f_i}{\partial y_j}(p)\right]$, $(i, j = 1, ..., n)$ é invertível então existem abertos Z, V e W, nos quais $p \in Z$ em $\mathbb{R}^{m+n}$, $a \in V$ em $\mathbb{R}^m$, $c = f(p) \in W$ em $\mathbb{R}^n$ e um difeomorfismo vertical $h : V \times W \to Z$, de classe C^k, tal que $f(h(x, w)) = w$ para todo $x \in V$ e todo $w \in W$.*

Figura 3.4: Representação da Forma Local das Submersões.



Fonte: LIMA, 2016, p. 118

Demonstração: Seja $\varphi : U \to \mathbb{R}^{m+n}$ 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 invertível no ponto $p = (a, b)$. Pelo Teorema da Aplicação Inversa, φ é um difeomorfismo de um aberto $Z \in p$ 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 $h : V \times W \to Z$ é da forma $h(x, w) = (x, h_2(x, w))$. Então, para qualquer $(x, w) \in V \times W$ temos

$$\begin{aligned}(x, w) &= \varphi(h(x, w)) \\ &= \varphi(x, h_2(x, w)) \\ &= (x, f(x, h_2(x, w))) \\ &= (x, f(h(x, w)))).\end{aligned}$$

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

Dada $f : U \to \mathbb{R}^n$, de classe C^k no aberto $U \subset \mathbb{R}^{m+n}$, a matriz da sua derivada $f'(p) : \mathbb{R}^{m+n} \to \mathbb{R}^n$ tem n linhas e $m + n$ colunas, que é a matriz Jacobiana $Jf(p)$.

Observação 3.85. A transformação linear $f'(p)$ é sobrejetiva quando é possível escolher n colunas de modo que a matriz $n \times n$ resultante, seja invertível. No teorema anterior, a fim de simplificar a notação, foram escolhidas as n últimas colunas.

Definição 3.86. Seja $f : U \to \mathbb{R}^n$ com $U \subset \mathbb{R}^{m+n}$, dizemos que f é uma submersão se f

possui derivada sobrejetiva $f'(z) : \mathbb{R}^{m+n} \to \mathbb{R}^n$ em todo ponto $z \in U$.

Corolário 3.87. *Seja $f : U \to \mathbb{R}^n$ 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 $h : V \times W \to Z$ de classe C^k, tais que $f(h(x,w)) = w$, para todos $x \in V$ e $w \in W$.*

Demonstração: Como f é uma submersão, $f'(z) : \mathbb{R}^{m+n} \to \mathbb{R}^n$ é sobrejetiva, então n das $m+n$ colunas da matriz Jacobiana $Jf(z)$ são linearmente independentes, logo formam uma matriz Jacobiana invertível $n \times n$. Se essas forem as últimas colunas, o Corolário é precisamente o Teorema 3.84. Se não forem as últimas, modificamos um pouco sua demonstração, permutando inicialmente as coordenadas $\mathbb{R}^{m+n}$ de modo que as n colunas linearmente independentes de $Jf(z)$ sejam as últimas, agora. ■

Teorema 3.88. *(Teorema das Funções Implícitas) Seja $f = (f_1, ..., f_n) : U \to \mathbb{R}^n$ de classe C^k $(k \geq 1)$ no aberto $U \subset \mathbb{R}^{m+n}$. Suponhamos que, no ponto $p = (a,b) \in U$, com $f(p) = c$, a matriz $n \times n$;*

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

seja invertível. Então $Z \subset U$, aberto contendo p, $V \subset \mathbb{R}^m$, aberto contendo a, e $\xi : V \to \mathbb{R}^n$ de classe C^k, com $\xi(a) = b$, com a seguinte propriedade:

$$[(x,y) \in Z \text{ e } f(x,y) = c] \Leftrightarrow [x \in V \text{ e } y = \xi(x)].$$

A equivalência acima significa que $f^{-1}(c) \cap Z$ é o gráfico de ξ, isto é, $f^{-1}(c) \cap Z = \{(x, \xi(x)); x \in V\} = G(\xi)$.

Demonstração: Defina;

$$\begin{aligned} \varphi : U &\longrightarrow \mathbb{R}^{m+n} \\ (x,y) &\longmapsto \varphi(x,y) = (x, f(x,y)). \end{aligned}$$

Por $f \in C^k$, temos que $\varphi \in C^k$ e

$$J\varphi(p) = \begin{bmatrix} I & 0 \\ A & B \end{bmatrix}, \text{ onde } A = \left[\frac{\partial f_i}{\partial x_l}(p)\right] \text{ e } B = \left[\frac{\partial f_i}{\partial x_j}(p)\right],$$

com $i, j = 1, ..., n$ e $l = 1, ..., m$. Desta forma, $detJ\varphi(p) = det(I) \cdot det(B) \neq 0$, logo $J\varphi(p)$ é invertível. Assim, pelo Teorema da Função Inversa, existe $\delta > 0$ e $\varphi(B_\delta)$ tal que $\varphi : B_\delta \to \varphi(B_\delta)$ é um difeomorfismo, com $B_\delta = B(p;\delta)$. Sendo $V \subset \mathbb{R}^m$ e $W \subset \mathbb{R}^n$ abertos, denotando $Z = B_\delta$ e escrevendo $\varphi(B_\delta) = V \times W$, considere $h : V \times W \to Z$ o difeomorfismo inverso de φ. Assim, $h(x,y) = (x, h_2(x,y))$. Perceba que se $(x,w) \in V \times W$, então $(x,w) = \varphi(h(x,w)) = \varphi(x, h_2(x,w)) = (x, f(x, h_2(x,w)))$. Logo, $f(x, h_2(x,w)) = w$

para todo $(x,w) \in V \times W$. Deste modo, defina $\xi : V \to \mathbb{R}^n$; $\xi(x) = h_2(x,c)$ em que $f(p) = c$. Assim, $\xi \in C^k$, pois h_2 é de classe C^k. Agora, observe que se $(x,y) \in Z$, então $x \in V$ e $(x,y) = h(x,w) = (x, h_2(x,w))$, para algum $w \in W$. Como $f(p) = c$, temos que $\varphi(p) = (a, f(p)) = (a,c) \in V \times W$, então $c \in W$. Deste modo, se $x \in V$, temos $(x,c) \in V \times W$ e $c = f(h(x,c)) = f(x, h_2(x,c)) = f(x, \xi(x))$. Também $c = f(x,y) = f(x, h_2(x,w)) = f(x,\xi(x))$, assim $\xi(x) = y$. Portanto, se $(x,y) \in f^{-1}(c) \cap Z$, tem-se $(x,y) \in G(\xi)$. Reciprocamente se $x \in V$, sendo $y = \xi(x)$, temos $y = h_2(x,c)$ então $f(x,y) = f(x, h_2(x,c)) = f(h(x,c)) = c$. Logo, $(x,y) \in f^{-1}(c) \cap Z$. ■

Corolário 3.89. *Seja $f : U \to \mathbb{R}^n$ de classe C^k no aberto $U \subset \mathbb{R}^{m+n}$. Se, no ponto $p \in U$, com $f(p) = c$, a derivada $f'(p) : \mathbb{R}^{m+n} \to \mathbb{R}^n$ é sobrejetiva então existe um aberto $Z \subset U$, com $p \in Z$, tal que $f^{-1}(c) \cap Z$ é o gráfico de uma aplicação $\xi : V \to \mathbb{R}^n$, de classe C^k num aberto $V \subset \mathbb{R}^m$.*

A abordagem clássica do Teorema das Funções Implícitas "Se $f_1, ..., f_n$ são funções reais de $m+n$ variáveis, k vezes continuamente diferenciáveis, e $p = (a_1, ..., a_m, b_1, ..., b_n)$ é uma solução particular do sistema de equações,

$$\begin{aligned} f_1(x_1, ..., x_m, y_1, ..., y_n) &= c_1 \\ f_2(x_1, ..., x_m, y_1, ..., y_n) &= c_2 \\ &\vdots \\ f_n(x_1, ..., x_m, y_1, ..., y_n) &= c_n, \end{aligned}$$

onde a matriz $n \times n$ é invertível, então as equações acima definem, de maneira única, na vizinhança do ponto p em $\mathbb{R}^{m+n}$, as variáveis $y_1, ..., y_n$ como funções de classe C^k das variáveis $x_1, ..., x_m : y_1 = \xi(x_1, ..., x_m), ..., y_n = \xi(x_1, ..., x_m)$". Escrevendo $x = (x_1, ..., x_m)$ e $\xi(x) = (\xi_1(x), ..., \xi_n(x))$, tem-se para cada $i = 1, ..., n$, com x em uma vizinhança de $a = (a_1, ..., a_m)$:

$$f_i(x, \xi_1(x), ..., \xi_n(x)) = c_i, \ ou \ f_i(x, \xi(x)) = c_i.$$

Derivando cada uma dessas n identidades em relação a x_j, temos;

$$\frac{\partial f_i}{\partial x_j} + \sum_{k=1}^{n} \frac{\partial f_i}{\partial y_k} \cdot \frac{\partial \xi_k}{\partial x_j} = 0, \ j = 1, ..., m.$$

Em termos matriciais, isto significa que;

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial f_n}{\partial x_j} \end{bmatrix} = - \begin{bmatrix} \dfrac{\partial f_1}{\partial y_1} & \cdots & \dfrac{\partial f_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial y_1} & \cdots & \dfrac{\partial f_n}{\partial y_n} \end{bmatrix} \begin{bmatrix} \dfrac{\partial \xi_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial \xi_n}{\partial x_j} \end{bmatrix}.$$

Ou seja;

$$\begin{bmatrix} \dfrac{\partial \xi_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial \xi_n}{\partial x_j} \end{bmatrix} = - \begin{bmatrix} \dfrac{\partial f_1}{\partial y_1} & \cdots & \dfrac{\partial f_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial y_1} & \cdots & \dfrac{\partial f_n}{\partial y_n} \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{\partial f_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial f_n}{\partial x_j} \end{bmatrix}.$$

Isto exibe as derivadas parciais $\frac{\partial \xi_i}{\partial x_j}$ a partir de $f_1, ..., f_n$, sem ser necessário conhecer explicitamente as funções $\xi_1, ..., \xi_n$.

Sob o ponto de vista da Álgebra Linear intríseca, a fim de mostrar como a derivada $\xi'(x) : \mathbb{R}^m \to \mathbb{R}^n$ pode ser calculada quando se conhece f mas não ξ explicitamente, é preciso entender o conceito da derivada parcial como transformação linear.

Definição 3.90. As transformações lineares $\frac{\partial f}{\partial x}(z) : \mathbb{R}^m \to \mathbb{R}^n$ e $\frac{\partial f}{\partial y}(z) : \mathbb{R}^n \to \mathbb{R}^m$, cujas matrizes nas bases canônicas dos espaços Euclidianos em questão, são $\left[\frac{\partial f_i}{\partial x_j}(z)\right] \in M_{(n\times m)}$ e $\left[\frac{\partial f_i}{\partial y_j}(z)\right] \in M_{(n\times n)}$, são chamadas as derivadas parciais de f no ponto z, relativamente a decomposição $\mathbb{R}^{m+n} = \mathbb{R}^m \bigoplus \mathbb{R}^n$, obtida ao escrevermos cada $z \in \mathbb{R}^{m+n}$ sob a forma $z = (x_1, ..., x_m, y_1, ..., y_n)$. Assim, $\frac{\partial f}{\partial x}(z)$ é a restrição da transformação linear $f'(z) : \mathbb{R}^{m+n} \to \mathbb{R}^n$ ao subespaço $\mathbb{R}^m \subset \mathbb{R}^{m+n}$ formado pelos vetores $(x, 0)$ e $\frac{\partial f}{\partial y}(z)$ é a restrição de $f'(z)$ ao subespaço $\mathbb{R}^n$ que consiste nos vetores da forma $(0, y)$. Para todo vetor $w = (u, v) \in \mathbb{R}^{m+n}$, tem-se

$$f'(z) \cdot w = \frac{\partial f}{\partial x}(z) \cdot u + \frac{\partial f}{\partial y}(z) \cdot v.$$

Usando essas derivadas parciais, a Regra da Cadeia nos permite concluir, a partir da identidade $f(x, \xi(x)) = c$ para todo $x \in V$, que

$$\frac{\partial f}{\partial x}(z) + \frac{\partial f}{\partial y}(z) \cdot \xi'(x) = 0, \ \textit{com} \ z = (x, \xi(x)).$$

Logo

$$\xi'(x) = -\left[\frac{\partial f}{\partial y}(z)\right]^{-1} \cdot \frac{\partial f}{\partial x}(z).$$

Note que a hipótese do Teorema das Funções Implícitas assegura que a transformação linear $\frac{\partial f}{\partial y}(z) : \mathbb{R}^n \to \mathbb{R}^n$ é invertível para todo z na vizinhança de p.

Capítulo 4

Superfícies Regulares

As superfícies regulares são definidas como conjuntos e não como aplicações. Sem entrar em pormenores, uma superfície em $\mathbb{R}^3$ é obtida tomando-se pedaços do plano, deformando-os e colocando-os entre si, de modo que a figura resultante não apresente pontos, arestas ou auto-interseções, e que tenha sentido falar em plano tangente nos pontos desta figura. Neste capítulo, a ideia é definir um conjunto que seja, em certo sentido, bi-dimensional e que seja suficientemente suave de modo que as noções usuais de Cálculo possam ser estendidas a esse tal conjunto. Aqui, usamos como referência principal [5].

4.1 Superfícies Regulares e Valores Regulares

Definição 4.1. Um subconjunto $S \subset \mathbb{R}^3$ é uma superfície regular se, para cada $p \in S$, existe uma vizinhança V de p em $\mathbb{R}^3$ e uma aplicação $\psi : U \to V \cap S$ de um aberto U de $\mathbb{R}^2$ sobre $V \cap S \subset \mathbb{R}^3$ tal que;

1. ψ é diferenciável;

   Ou seja, se escrevemos $\psi(u, v) = (\psi_1(u, v), \psi_2(u, v), \psi_3(u, v))$; $(u, v) \in U$, as funções $\psi_1(u, v), \psi_2(u, v), \psi_3(u, v)$ têm derivadas parciais contínuas de todas as ordens em U.

2. ψ é um homeomorfismo;

3. (Condição de regularidade) Para todo $q \in U$, a derivada $\psi'(q) : \mathbb{R}^2 \to \mathbb{R}^3$ é injetiva.

A aplicação ψ é chamada uma parametrização ou um sistema de coordenadas locais em uma vizinhança de p. A vizinhança $V \cap Z$ de p em S é chamada uma vizinhança coordenada.

A matriz da aplicação linear $\psi'(q)$ é dada por; $\psi'(q) = \begin{bmatrix} \dfrac{\partial\psi_1}{\partial u} & \dfrac{\partial\psi_1}{\partial v} \\ \dfrac{\partial\psi_2}{\partial u} & \dfrac{\partial\psi_2}{\partial v} \\ \dfrac{\partial\psi_3}{\partial u} & \dfrac{\partial\psi_3}{\partial v} \end{bmatrix}$.

Observação 4.2. A condição 3 pode ser expressa exigindo-se que os dois vetores coluna desta matriz sejam linearmente independentes, ou equivalentemente, que o produto vetorial $\frac{\partial\psi_1}{\partial u} \wedge \frac{\partial\psi_1}{\partial v} \neq 0$, ou ainda, que um dos menores de ordem 2 da matriz $\psi'(q)$, isto é, que um dos determinantes Jacobianos;

$$\frac{\partial(\psi_1, \psi_2)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial\psi_1}{\partial u} & \dfrac{\partial\psi_1}{\partial v} \\ \dfrac{\partial\psi_2}{\partial u} & \dfrac{\partial\psi_2}{\partial v} \end{vmatrix}, \quad \frac{\partial(\psi_2, \psi_3)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial\psi_2}{\partial u} & \dfrac{\partial\psi_2}{\partial v} \\ \dfrac{\partial\psi_3}{\partial u} & \dfrac{\partial\psi_3}{\partial v} \end{vmatrix}, \quad \frac{\partial(\psi_1, \psi_3)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial\psi_1}{\partial u} & \dfrac{\partial\psi_1}{\partial v} \\ \dfrac{\partial\psi_3}{\partial u} & \dfrac{\partial\psi_3}{\partial v} \end{vmatrix},$$

seja diferente de zero em q.

Perceba que a condição 1 é bem natural se esperamos fazer alguma geometria diferencial sobre S. Já a injetividade na condição 2, tem como objetivo excluir a possibilidade de auto-interseções em superfícies regulares. E isso é necessário se queremos falar sobre plano tangente em um ponto p de S. A continuidade da inversa ainda na condição 2 tem um propósito mais sutil, o qual poderá ser entendido com os resultados posteriores. Por fim, a condição 3 garante a existência de um "plano tangente" em todos os pontos de S.

Exemplo 4.3. Vamos mostrar que a esfera unitária $S^2 = \{(x, y, z) \in \mathbb{R}^3; x^2 + y^2 + z^2 = 1\}$ é uma superfície regular.

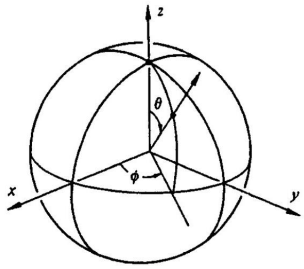
Inicialmente, verificaremos que a aplicação $\psi^1 : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ dada por $\psi^1(x, y) = (x, y, \sqrt{1 - (x^2 + y^2)})$; $(x, y) \in U$, onde $\mathbb{R}^2 = \{(x, y, z) \in \mathbb{R}^3; z = 0\}$ e $U = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 < 1\}$ é uma parametrização de S^2. Note que $\psi^1(U)$ é a parte "aberta" de S^2 acima do plano xy. Como $x^2 + y^2 < 1$, a função $\sqrt{1 - (x^2 + y^2)}$ tem derivadas parciais contínuas de todas as ordens. Assim, ψ^1 é diferenciável e a condição 1 é satisfeita. Verificamos a condição 3 facilmente, visto que;

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 - 0 = 1.$$

Para verificarmos a condição 2, observamos que ψ^1 é bijetiva e que $(\psi^1)^{-1}$ é a restrição da projeção contínua $\pi(x, y, z) = (x, y)$ ao conjunto $\psi^1(U)$. Portanto, $(\psi^1)^{-1}$ é contínua em $\psi^1(U)$. Agora vamos cobrir a esfera inteira com parametrizações similares. Definimos $\psi^2 : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ por $\psi^2(x, y) = (x, y, -\sqrt{1 - (x^2 + y^2)})$; $(x, y) \in U$, verificamos que ψ^2 é uma parametrização, e observamos que $\psi^1 \cup \psi^2$ cobre a esfera, exceto o equador $\{(x, y, z) \in \mathbb{R}^3; x^2 + y^2 = 1; z = 0\}$. Utilizando os planos xz e zy, definimos as seguintes parametrizações; $\psi^3(x,z) = (x, \sqrt{1-(x^2+z^2)}, z)$, $\psi^4(x,z) = (x, -\sqrt{1-(x^2+z^2)}, z)$, $\psi^5(y,z) = (\sqrt{1-(y^2+z^2)}, y, z)$ e $\psi^6(y,z) = (-\sqrt{1-(y^2+z^2)}, y, z)$, que juntamente com ψ^1 e ψ^2, cobrem inteiramente S^2. Assim, mostramos que S^2 é uma superfície regular.

Para a maioria da aplicações, é conveniente relacionar parametrizações com as coordenadas geográficas em S^2. Sejam $V = \{(\theta, \varphi); 0 < \theta < \pi,\ 0 < \varphi < 2\pi\}$ e $\psi : V \to \mathbb{R}^3$ dada por $\psi(\theta, \varphi) = (sen(\theta)cos(\varphi), sen(\theta)sen(\varphi), cos(\theta))$. É claro que $\psi(V) \subset S^2$. Vamos mostrar que ψ é uma parametrização de S^2. Comumente dizemos que θ é a colatitude (complemento da latitude) e φ é a longitude.

Figura 4.1: Colatitude e Longitude de uma Superfície.



Fonte: CARMO, 2010, p. 66

ψ é diferenciável, visto que as funções $sen(\theta)cos(\varphi), sen(\theta)sen(\varphi), cos(\theta)$ têm derivadas parciais contínuas de todas as ordens. Além disso, para que os determinantes jacobianos;

$$\frac{\partial(x,y)}{\partial(\theta,\varphi)} = cos(\theta)sen(\theta)$$
$$\frac{\partial(y,z)}{\partial(\theta,\varphi)} = sen^2(\theta)cos(\theta)$$
$$\frac{\partial(x,z)}{\partial(\theta,\varphi)} = -sen^2(\theta)sen(\varphi),$$

se anulem simultaneamente, é preciso que $\theta = \pi$ e $\varphi = \pi$. Mas isso não ocorre em V, visto que $0 < \theta < \pi$, portanto as condições 1 e 3 são satisfeitas. Em seguida, observamos que dado $(x,y,z) \in S^2 - C$, onde C é o semi-círculo $C = \{(x,y,z) \in S^2;\ y = 0,\ x \geq 0\}$, θ fica determinado de maneira única por $\theta = cos^{-1}(z)$, uma vez que $0 < \theta < \pi$. Conhecendo o valor de θ, determinamos $sen(\varphi)$ e $cos(\varphi)$ a partir de $x = sen(\theta)\cos(\varphi)$, $y = sen(\theta)sen(\varphi)$, e isto determina φ de maneira única $(0 < \varphi < 2\pi)$. Portanto ψ tem uma inversa ψ^{-1}. Para completar a verificação da condição 2, deveríamos mostrar que ψ^{-1} é contínua, contudo, veremos adiante que esta prova não é necessária quando já se sabe que S é uma superfície regular.

Notamos que $\psi(V)$ omite apenas um semi-círculo de S^2 (incluindo os dois polos) e que S^2 pode ser coberta por vizinhanças coordenadas de duas parametrizações deste tipo.

O exemplo anterior nos mostrou que decidir se um dado subconjunto de $\mathbb{R}^3$ é uma superfície regular, pode ser uma tarefa cansativa, visto que precisamos verificar os três itens. A fim de simplificar, veremos duas proposições a seguir.

Proposição 4.4. *Se $f : U \to \mathbb{R}$ é uma função diferenciável em um conjunto aberto U de $\mathbb{R}^2$, então o gráfico de f, isto é, o subconjunto de $\mathbb{R}^3$ dado por $(x, y, f(x, y))$ para $(x, y) \in U$, é uma superfície regular.*

Demonstração: Bata-nos mostrar que a aplicação $\psi : U \to \mathbb{R}^3$ dada por $\psi(u, v) = (u, v, f(u, v))$ é uma parametrização do gráfico. Perceba que a condição 1 verificamos facilmente, visto que f é diferenciável, então ψ também é diferenciável. Note ainda que $\frac{\partial(x,y)}{\partial(u,v)} = 1$, portanto a condição 3 é satisfeita. Por fim, cada ponto (x, y, z) do gráfico é a imagem por ψ de um único ponto $(u, v) = (x, y) \in U$. Por consequência, ψ é bijetiva, e como ψ^{-1} é a restrição ao gráfico de f da projeção contínua de $\mathbb{R}^3$ sobre o plano xy, ψ^{-1} é contínua. ■

Definição 4.5. Um ponto de $\mathbb{R}^m$ é dito um valor regular de f quando ele não é um valor crítico. Reveja o que é um valor crítico na Definição 3.24.

Observação 4.6. Um ponto $p = (x_0, y_0) \in U$ é crítico se $f'(p) = 0$, ou seja, se a derivada de f no ponto p leva todos os vetores em $\mathbb{R}^2$ no vetor nulo. Consideramos por definição que qualquer $a \notin f(U)$ é um valor regular de f.

Proposição 4.7. *Se $f : U \subset \mathbb{R}^3 \to \mathbb{R}$ é uma função diferenciável e $a \in f(U)$ é um valor regular de f, então $f^{-1}(a)$ é uma superfície regular em $\mathbb{R}^3$.*

Demonstração: Seja $p = (x_0, y_0, z_0)$ um ponto de $f^{-1}(a)$. Por a ser um valor regular de f, podemos admitir, trocando os nomes dos eixos coordenados se necessário, que $f_z \neq 0$ em p, já que a não é ponto crítico. Definindo uma aplicação $F : U \to \mathbb{R}^3$ dada por $F(x, y, z) = (x, y, f(x, y, z))$ e indicando por (u, v, t) as coordenadas de um ponto de $\mathbb{R}^3$ onde F toma seus valores, perceba que a derivada de F em p é dada por;

$$F'(p) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ f_x & f_y & f_z \end{bmatrix}.$$

Assim, $detF'(p) \neq 0$ e, pelo Teorema da Função Implícita, existem $Z \subset U$; $p \in Z$ e $V \subset \mathbb{R}^2$ com $(x_0, y_0) \in V$ e $\xi : V \to \mathbb{R}$ diferenciável tal que $\xi(x_0, y_0) = z_0$ satisfaz $G(\xi) = f^{-1}(a) \cap Z$. Ou seja, $f^{-1}(a)$ é localmente o gráfico de uma função diferenciável, portanto, pela Proposição 4.4, $f^{-1}(a)$ é uma superfície diferenciável. ■

Exemplo 4.8. O elipsóide formado pelo conjunto de pontos (x, y, z) de $\mathbb{R}^3$ satisfazendo $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$, é uma superfície regular.

De fato, defina $f : \mathbb{R}^3 \to \mathbb{R}$ dada por $f(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1$. Perceba que f é uma função diferenciável, que 0 é um valor regular de f e que $f^{-1}(0)$ é o elipsóide em questão. Visto que as derivadas parciais $f_x = \frac{2x}{a^2}, f_y = \frac{2y}{b^2}$ *e* $f_z = \frac{2z}{c^2}$ se anulam simultaneamente apenas no ponto $(0,0,0)$, que não pertence a $f^{-1}(0)$. Assim pela Proposição 4.7, o elipsoide é uma superfície diferenciável

Definição 4.9. Uma superfície $S \subset \mathbb{R}^3$ é conexa quando é conexa por caminhos.

Note que os exemplos de superfícies regulares que foram apresentados até agora, são conjuntos conexos de $\mathbb{R}^3$. Mas, a definição de superfície regular não apresenta nenhuma restrição quanto a conexidade das superfícies. Veremos então, que superfícies regulares dadas pela Proposição anterior podem não ser conexas.

Exemplo 4.10. O parabolóide de duas folhas $-x^2 - y^2 + z^2 = 1$ é uma superfície regular.

De fato, já que é dada por $S = f^{-1}(0)$, onde 0 é um valor regular de $f(x,y,z) = -x^2 - y^2 + z^2 - 1$. Perceba que a superfície S não é conexa, pois dados dois pontos em duas folhas distintas ($z > 0$ *e* $z < 0$) não é possível ligá-las por uma curva contínua $\alpha(t) = (x(t), y(t), z(t))$ contida na superfície. Pois se fosse possível, $z(t)$ mudaria de sinal, e para algum t_0 teríamos $z(t_0) = 0$. Ou seja, $\alpha(t_0) \neq S$.

Proposição 4.11. *Se $f : S \subset \mathbb{R}^3 \to \mathbb{R}$ é uma função contínua que não se anula, definida sobre uma superfície conexa, então f não muda de sinal em S.*

Demonstração: Suponha, por contradição, que $f(p) > 0$ e $f(q) < 0$ para dois pontos $p, q \in S$. Por S ser conexa, existe uma curva contínua $\alpha : [a,b] \to S$, tal que $\alpha(a) = p$ e $\alpha(b) = q$. Aplicando o Teorema do Valor Intermediário à função contínua $f \circ \alpha : [a,b] \to \mathbb{R}$, temos que existe $C \in (a,b)$ tal que $f \circ \alpha(c) = 0$, isto é, f é zero em $\alpha(c)$, o que é um absurdo. ∎

Exemplo 4.12. O toro T é a "superfície" gerada pela rotação de um círculo S^1 de raio r em torno de uma reta pertencente ao plano do círculo e a uma distância $a > r$ do centro do círculo.

Seja S^1 o círculo do plano yz centrado no ponto $(0,a,0)$. Então S^1 é dado por $(y-a)^2 + z^2 = r^2$ e os pontos do conjunto T, obtidos pela rotação deste cículo em torno do eixo O_z satisfazem a equação $z^2 = r^2 - (\sqrt{x^2+y^2} - a)^2$. Consequentemente T é a imagem inversa de r^2 pela função $f(x,y,z) = z^2 + (\sqrt{x^2+y^2} - a)^2$. Essa função é diferenciável para $(x,y) \neq (0,0)$, e como $\frac{\partial f}{\partial z} = 2z$, $\frac{\partial f}{\partial y} = \frac{2y(\sqrt{x^2+y^2}-a)}{\sqrt{x^2+y^2}}$, $\frac{\partial f}{\partial x} = \frac{2x(\sqrt{x^2+y^2}-a)}{\sqrt{x^2+y^2}}$, r^2 é um valor regular de f. Portanto o toro T é uma superfície regular.

A Proposição 4.4 nos diz que o gráfico de uma função diferenciável é uma superfície regular. Veremos a sua recíproca localmente com a proposição a seguir.

Proposição 4.13. *Sejam $S \subset \mathbb{R}^3$ uma superfície regular e $p \in S$. Então existe uma vizinhança V de p em S tal que V é o gráfico de uma função diferenciável que tem uma das seguintes formas: $z = f(x,y)$, $y = g(x,z)$, $x = h(y,z)$.*

Demonstração: Seja $\psi : U \subset \mathbb{R}^2 \to S$ uma parametrização de S em p, em que $\psi(u,v) = (\psi_1(u,v), \psi_2(u,v), \psi_3(u,v)); (u,v) \in U$. Pela condição de regularidade da definição de superfície regular, um dos determinantes Jacobianos $\frac{\partial(\psi_1,\psi_2)}{\partial(u,v)}$, $\frac{\partial(\psi_2,\psi_3)}{\partial(u,v)}$, $\frac{\partial(\psi_1,\psi_3)}{\partial(u,v)}$ não se anula em $\psi^{-1}(p) = q$. Assim, supondo que $\frac{\partial(\psi_1,\psi_2)}{\partial(u,v)}(q) \neq 0$, considere a aplicação $\pi \circ \psi : U \to \mathbb{R}^2$, onde π é a projeção $\pi(x,y,z) = (x,y)$. Logo, $(\pi \circ \psi)(u,v) = (\psi_1(u,v), \psi_2(u,v))$ e por $\frac{\partial(\psi_1,\psi_2)}{\partial(u,v)}(q)$ não se anular, podemos aplicar o Teorema da Função Inversa para garantir que existem as vizinhanças V_1 de q e V_2 de $(\pi\circ\psi)(q)$ tais que $\pi\circ\psi$ aplique V_1 difeomorficamente sobre V_2. Portanto, π restrita a $\psi(V_1) = V$. Além disso, a função $(\pi \circ \psi)^{-1} : V_2 \to V_1$ é diferenciável, assim observando que $(\pi \circ \psi)^{-1}(x,y) = \psi^{-1}(\pi^{-1}(x,y)) = \psi^{-1}(x,y,z(x,y))$. De onde temos que $z : V_2 \to \mathbb{R}$ é uma função diferenciável, concluindo então que V é o gráfico de uma função diferenciável. De maneira análoga, obtemos $y = g(x,z)$ e $x = h(y,z)$, supondo que $\frac{\partial(y,z)}{\partial(u,v)} \neq 0$ e $\frac{\partial(x,z)}{\partial(u,v)} \neq 0$, respectivamente. ■

Veremos adiante que quando já sabemos que S é uma superfície regular e temos ψ como candidato a uma parametrização de S, então não precisamos verificar que ψ^{-1} é contínua, desde que as demais condições da definição de superfície regular sejam satisfeitas.

Proposição 4.14. *Seja $p \in S$ um ponto de uma superfície regular S e seja $\psi : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ uma aplicação com $p \in \psi(U)$ tal que as condições 1 e 3 da definição de superfície regular sejam satisfeitas. Suponha que ψ seja bijetiva. Então ψ^{-1} é contínua.*

Demonstração: Seja, $\psi(u,v) = (\psi_1(u,v), \psi_2(u,v), \psi_3(u,v)); (u,v) \in U$ e $q \in U$, pelas condições 1 e 3 da definição de superfície regular, podemos admitir, trocando os eixos coordenados de $\mathbb{R}^3$ se necessário, que $\frac{\partial(\psi_1,\psi_2)}{\partial(u,v)}(q) \neq 0$. Considere $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ a projeção $\pi(x,y,z) = (x,y)$. A partir do Teorema da Função Inversa, obtemos V_1 de q em U e V_2 de $(\pi \circ \psi)(q)$ em $\mathbb{R}^2$ tais que $\pi \circ \psi$ aplica V_1 difeomorficamente sobre V_2. Como por hipótese ψ é bijetiva, então restrita a $\psi(V_1)$, $\psi^{-1} = (\pi \circ \psi)^{-1} \circ \pi$. Como a composição é feita com aplicações contínuas, segue que ψ^{-1} também é contínua. ■

Exemplo 4.15. O cone de uma folha C, dado por $z = \sqrt{x^2 + y^2}; (x,y) \in \mathbb{R}^2$, não é uma superfície regular.

Perceba que não podemos chegar a esta conclusão apenas pelo fato da parametrização "natural" $(x,y) \to (x,y,\sqrt{x^2+y^2}); (x,y) \in \mathbb{R}^2$ não ser diferenciável, pois poderiam existir outras parametrizações que cumprissem todas as condições da definição de superfície regular. Pela Proposição 4.13, se C fosse uma superfície regular, seria uma vizinhança do $(0,0,0) \in C$, o gráfico de uma função diferenciável tendo uma das três formas; $y = h(x,z)$, $x = g(y,z)$, $z = f(x,y)$. Como as projeções de C sobre os planos xz e yz não são injetivas, descartamos as duas primeiras formas. Assim, $z = f(x,y)$ teria que coincidir, em uma vizinhança do $(0,0,0)$, com $z = \sqrt{x^2+y^2}$. Mas como z não é diferenciável em $(0,0)$, visto que $\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2+y^2}}$ e $\frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2+y^2}}$. Portanto, C não é uma superfície regular.

Exemplo 4.16. Uma parametrização para o toro T (do exemplo 4.12) pode ser dada por $\psi(u,v) = ((rcos(u)+a)cos(v), (rcos(u)+a)sen(v), rsen(u))$, onde $0 < u < 2\pi$, $0 < v < 2\pi$.

Note que ψ é diferenciável visto que suas componentes têm derivadas parciais contínuas de todas as ordens. Além disso,

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} -rsen(u)cos(v) & -rcos(u)sen(v) - asen(v) \\ -rsen(u)sen(v) & rcos(u)cos(v) + acos(v) \end{vmatrix} = sen(u)(-r^2cos(u) - ra), \tag{4.1}$$

e;

$$\frac{\partial(x,z)}{\partial(u,v)} = \begin{vmatrix} -rsen(u)cos(v) & -rcos(u)sen(v) - asen(v) \\ rcos(u) & 0 \end{vmatrix} = cos(u)sen(v)(r^2cos(u)+ra). \tag{4.2}$$

Perceba que para que para o determinante Jacobiano (4.1) se anular, é necessário que $u = \pi$, mas $(-r^2cos(u) - ra)$ nunca se anula, visto que $a > r$. Enquanto que para o determinante Jacobiano(4.2) se anular, necessariamente $v = \pi$ ou $u = \frac{3\pi}{2}$. Portanto, os determinantes são se anulam simultaneamente, satisfazendo assim, a condição de regularidade. Além disso, como vimos anteriormente que T é uma superfície regular, pela Proposição 4.14, para verificarmos a segunda condição de superfície regular, basta-nos mostrar que ψ é bijetiva. Observamos que $sen(u) = \frac{z}{r}$ e se $\sqrt{x^2+y^2} \leq a$, então $\frac{\pi}{2} \leq u \leq \frac{3\pi}{2}$, se $\sqrt{x^2+y^2} \geq a$, então ou $0 < u \leq \frac{\pi}{2}$ ou $\frac{3\pi}{2} \leq u \leq 2\pi$. Assim, dado (x,y,z), isto determina u, $0 < u < 2\pi$ de maneira única. Conhecendo u, x e y, determinamos $cos(v)$ e $sen(v)$. Determinando v de maneira única, $0 < v < 2\pi$. Portanto ψ é bijetiva.

Referências Bibliográficas

[1] LIMA, Elon Lages. **Análise Real - Funções de n variáveis**, vol 2: 6.ed. Rio de Janeiro. Editora Associação Instituto Nacional de Matemática Pura e Aplicada, 2016.

[2] LIMA, Elon Lages. **Curso de Análise**, vol 2: 11.ed. Rio de Janeiro: Editora Associação Instituto Nacional de Matemática Pura e Aplicada, 2010.

[3] LIMA, Elon Lages. **Análise Real - Funções de uma variável**, vol 1: 8.ed. Rio de Janeiro. Coleção Matemática Universitária, IMPA, 2006.

[4] STEWART, James. **Cálculo**, vol 2: 7.ed. São Paulo: Editora Cengage Learning, 2013.

[5] CARMO, Manfredo Perdigão do. **Geometria Diferencial de Curvas e Superfícies**, 4.ed. Rio de Janeiro: Editora Sociedade Brasileira de Matemática, 2010.